

საქართველოს ტექნიკური უნივერსიტეტი

ოთარ თავდიშვილი

კომპიუტერული ხეღვა

I ნაწილი

დამტკიცებულია სახელმძღვანელოდ
სტუ-ს სარედაქციო-საგამომცემლო
საბჭოს მიერ

თბილისი
2013

უაკ 004.9

კომპიუტერული ხედვის წარმოდგენილი სახელმძღვანელო შეიქმნა საქართველოს ტექნიკური უნივერსიტეტის ინფორმატიკისა და მართვის სისტემების ფაკულტეტის ხელოვნური ინტელექტის დეპარტამენტის სტუდენტებისთვის წაკითხული ლექციების საფუძველზე. მასში განხილულია ორგანოზომილებიანი გამოსახულებების ციფრული დამუშავებისა და ანალიზის ძირითადი საკითხები.

სახელმძღვანელო განკუთვნილია ბაკალავრიატისა და მაგისტრატურის სტუდენტებისთვის, ის ასევე საინტერესო იქნება კომპიუტერული ხედვის საკითხებით დაინტერესებულ მკითხველთა ფართო წრისათვის.

ავტორისაგან

კომპიუტერული ხედვის წარმოდგენილ სახელმძღვანელოში განხილულია 2D ციფრული გამოსახულებების დამუშავებისა და ანალიზისთვის საჭირო აუცილებელი თეორიული საკითხები და მათი პრაქტიკაში გამოყენების კონკრეტული მაგალითები. კომპიუტერულ ხედვას ფართო გამოყენება აქვს ბიოლოგიაში, დედამიწის ზონდირების ამოცანებში, მანქანურ ხედვაში, სამედიცინო დიაგნოსტიკაში, სასოფლო-სამეურნეო ამოცანებში, კარტოგრაფიასა და, საზოგადოდ, სცენათა ანალიზის ამოცანებში.

წარმოდგენილი სახელმძღვანელო ამ დარგის პირველი ქართულენოვანი გამოცემის მცდელობაა. მას საფუძვლად დაედო ინფორმატიკისა და მართვის სისტემების ფაკულტეტის ხელოვნური ინტელექტის დეპარტამენტის სტუდენტებისთვის წაკითხული ლექციების კურსი.

მიუხედავად დიდი მცდელობისა სახელმძღვანელო ვერ იქნება დაზღვეული ხარვეზებისაგან. ავტორი სიამოვნებით მიიღებს ყველა საქმიან შენიშვნას, რომელიც ხელს შეუწყობს წიგნის დახვეწას და გაითვალისწინებს შემდეგი გამოცემისთვის.

სახელმძღვანელო განკუთვნილია ინფორმატიკისა და მართვის სისტემების ფაკულტეტის ბაკალავრიატისა და მაგისტრატურის სტუდენტებისთვის, ასევე ამ დარგში მომუშავე სპეციალისტებისთვის.

I თავი. შესავალი

1.1. კომპიუტერული ხედვის საგნის არსი და მიზანი

თვალი გრძნობის ორგანოებს შორის ყველაზე უფრო ინფორმაციულ ორგანოს წარმოადგენს. ის გვაწვდის ყველაზე დიდი მოცულობის ინფორმაციას გარესამყაროს შესახებ და შესაძლებლობას გვაძლევს დავამყაროთ ურთიერთობა ჩვენს ირგვლივ არსებულ საგნებთან უშუალო ფიზიკური კონტაქტის გარეშე. მისი საშუალებით ხდება ობიექტების ადგილმდებარეობის და ურთიერთგანლაგების განსაზღვრა სივრცეში.

ამასთან ერთად თვალი გრძნობათა ორგანოებს შორის ყველაზე უფრო რთული ორგანოა და მის შესახებ არსებული ცოდნა, ძირითადად, ეხება მხედველობითი ინფორმაციის დამუშავების იმ ეტაპებს, რომლებიც უშუალოდ დაკავშირებული სენსორებიდან მიღებულ მონაცემებთან. ჯერჯერობით წარუმატებელი აღმოჩნდა მცდელობა შექმნილიყო ისეთი მანქანები, რომლებსაც შეეძლებოდა “დანახვა”. უკვე შემუშავებული მხედველობითი სისტემებიდან განსაკუთრებით წარმატებული აღმოჩნდა ის სისტემები, რომელიც გამოიყენება სამრეწველო სფეროში, მაგალითად, კონვეიერიდან დეტალების ასაღებად განკუთვნილი რობოტის ხელის სამართი სისტემები.

გამოთვლითი მანქანების აპარატურული და პროგრამული უზრუნველყოფის განვითარება ემსახურება პრაქტიკაში მათი მრავალმხრივი და ეფექტური გამოყენების მიზანს. ეს კი თავის მხრივ ხელს უწყობს ხელოვნური ინტელექტის იმ დარგის განვითარებას, რომელსაც მანქანური აღქმა ეწოდება. კერძოდ, ყველაზე დიდი ყურადღება ექცევა იმ ამოცანებს, რომლებშიც მონაცემები წარმოდგენილია გამოსახულებების სახით.

საზოგადოდ, გამოსახულების სახით წარმოდგენილი მონაცემების დამუშავება იყოფა ორ მიმართულებად: კომპიუტერული გრაფიკა და კომპიუტერული ხედვა. ისინი წარმოადგენს კომპიუტერული მეცნიერების (ინფორმატიკის) მნიშვნელოვან სფეროებს.

კომპიუტერული გრაფიკის ამოცანა ის არის, რომ არაგამომსახველობით სახეში (რიცხვითი მონაცემები) წარმოდგენილი საწყისი მონაცემებიდან მოახდინოს გამოსახულების აღწარმოება (მიღება). ეს ნიშნავს, რომ ასეთი სისტემის შესასვლელს წარმოადგენს რიცხვითი მონაცემები, ხოლო გამოსასვლელს კი – გამოსახულება. რიცხვითი მონაცემების ვიზუალურ სახედ წარმოდგენის მარტივი მაგალითებია: 1) ფუნქციათა და ექსპერიმენტულ მონაცემთა საფუძველზე აგებული გრაფიკები; 2) მომხმარებლის მიერ მითითებულ წერტილებს შორის ეკრანზე გავლებული წირები; 3) დღეს უკვე ძალიან პოპულარული მანქანური თამაშები. მანქანური გრაფიკის მიმართულებას მიეკუთვნება აგრეთვე მანქანური ფერწერა და მულტიფილმების მანქანური სინთეზი.

კომპიუტერული ხედვის მიზანია რეალური სამყაროს სცენების და ობიექტების შესახებ სასარგებლო დასკვნების ფორმირება გადამწოდებიდან მიღებული გამოსახულებების ანალიზის საფუძველზე. ეს გულისხმობს თვით ობიექტების გარკვეული სახით აღწერას. შესაბამისად კომპიუტერული ხედვის საგანს წარმოადგენს სცენის აღწერის ფორმირება მისი გამოსახულების საფუძველზე.

კომპიუტერული ხედვა მოიცავს შემდეგ სამ დიდმნიშვნელოვან ამოცანას:

- გამოსახულებათა დამუშავება;
- სახეთა კლასიფიკაცია ანუ ამოცნობა;

- **სცენათა ანალიზი.**

გამოსახულებათა დამუშავების სისტემა შეიძლება წარმოვადგინოთ შემდეგი სქემის სახით (სურ. 1.1).



სურ. 1.1

სისტემის გამოსასვლელზე ახალი გამოსახულება მიიღება შესასვლელი გამოსახულების ხმაურისაგან გაწმენდის, კონტრასტირების, საზღვრების გამოკვეთის შედეგად. ასეთი დამუშავების შედეგი უნდა იყოს ის, რომ მიღებული გამოსახულება გარდაიქმნას ანალიზისთვის უფრო მოხერხებულ სახეში, ანუ უფრო ადვილად ინტერპრეტირებადი გახდეს. კონკრეტულად ეს გულისხმობს შემდეგი ამოცანების გადაწყვეტას: გამოსახულების გაუმჯობესება, რაც გულისხმობს გეომეტრიული დამახინჯებების კორექციას, ვიზუალური ხარისხის გაუმჯობესებას, გამოსახულების გაუმჯობესებას და აღდგენას, ეფექტურ კოდირებას. მაგალითად, სურათები, რომლებიც დაყოვნების შედეგადაა მიღებული, ასევე გადარეცხილი სურათები შეიძლება გაუმჯობესდეს კონტრასტის გაზრდის მეთოდების გამოყენებით. გამოსახულება, რომელზედაც გვაქვს სიკაშკაშეების დიდი დიაპაზონი, შეიძლება დაყვანილ იქნეს გამოსახულებაზე, სადაც სიკაშკაშეების მხოლოდ ორი დონეა (შავი და თეთრი). ასეთი გარდაქმნის შედეგად გამოსახულებაზე მიღებული ობიექტების სილუეტებიდან შეიძლება მივიღოთ ობიექტთა ფორმები (ობიექტების მომგლები კონტურები). ასევე შეიძლება მივიღოთ სულ ახალი გამოსახულება სხვა გამოსახულებათა კრებუ-

ლიდან, ანუ მოვახდინოთ გამოსახულებათა სინთეზი, ვთქვათ, როგორც ეს ტომოგრაფიაშია, სადაც ადამიანის სხეულის განივი კვეთების 2D გამოსახულებების სინთეზის შედეგად ხდება სხეულის სრული 3D სურათის მიღება. ზემოაღნიშნული ამოცანები მიეკუთვნება გამოსახულების **დამუშავების დაბალ დონეს**.

გამოსახულების **დამუშავების საშუალო დონეზე** სისტემის შესასვლელს ისევ გამოსახულება წარმოადგენს, მაგრამ გამოსასვლელი უკვე არა გამოსახულება, არამედ გამოსახულებიდან გამოყოფილი სხვადასხვა ატრიბუტებია. ესენია: ობიექტების კიდეები, კონტურები და ცალკეული ობიექტების ინდივიდუალობები – თავისებურებები.

სახეთა კლასიფიკაციის ანუ ამოცნობის ძირითად ამოცანას წარმოადგენს სახის “მიკუთვნება” რომელიმე წინასწარ ცნობილი კლასისათვის. გარკვეული აზრით ის მანქანური გრაფიკის შებრუნებულ ამოცანას წარმოადგენს. ამოცნობის პროცედურის განხორციელების შედეგად გამოსახულება გარდაიქმნება რაღაც აბსტრაქტული სახით წარმოდგენილ აღწერაში: რიცხვების კრებულში, სიმბოლოების ჯაჭვში, ან გრაფში. ასეთ დროს გამოსახულებაზე არსებული ობიექტები აღიწერება მათი მახასიათებელი ნიშან-თვისებების გაზომვის შედეგად მიღებული რიცხვების ერთობლიობების (სიმაღლე, სიგრძე, წონა, ფართობი, პერიმეტრი, ფერი და ა.შ.) სახით. მართალია, კლასიფიკატორის შესასვლელს გამოსახულება არ წარმოადგენს, მაგრამ სახეთა კლასიფიკაციის მეთოდები ფართოდ გამოიყენება მხედველობით სისტემებში გამოსახულებათა ანალიზისას (სურ. 1.2).

გამოსახულებათა ამოცნობის გამოყენების მარტივ მაგალითს წარმოადგენს ფოსტაში კონვერტების დახარისხების პროცესი. ამ დროს სპეციალური მოწყობილობა სწავლობს კონ-

ვერტზე დაწერილი საფოსტო ინდექსის შემადგენელ ციფრებს და ახდენს მათ იდენტიფიკაციას.



სურ. 1.2

სცენათა ანალიზის ამოცანაში გამოსახულების წინასწარი დამუშავების შედეგად მიღებული მარტივი, დაბალი დონის სიმბოლური აღწერიდან (კონტურული სურათები) ხდება გადასვლა უფრო რთული, მაღალი დონის აღწერისაკენ (ობიექტები, მათი ფორმები, ობიექტების ურთიერთგანლაგება სივრცეში, ორიენტაცია) (სურ. 1.3).

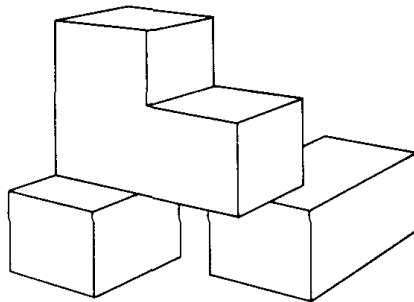
ზემოაღნიშნულ სახეთა კლასიფიკაციის (ანუ ამოცნობის) და სცენათა ანალიზის ამოცანები, რომელთა მიზანია გარკვეული დასკვნების გაკეთება ამოცნობილი ობიექტების შესახებ, დაწყებული გამოსახულების ანალიზიდან და დამთავრებული ხედვასთან დაკავშირებული შემეცნებითი პროცესებით, მიეკუთვნება დამუშავების მაღალ დონეს.



სურ. 1.3

საილუსტრაციოდ განვიხილოთ კონტურულ სახეში წარმოდგენილი სურათის ინტერპრეტაციის ამოცანა (სურ. 1.4).

სურათზე წარმოდგენილ გამოსახულებაზე ერთმანეთზე მოთავსებული რამდენიმე მრავალწახნაგას აღწერა მოცემულია წრფის მონაკვეთების სიმრავლის სახით (ესკიზური აღწერა). მანამ, სანამ გადავალთ უფრო რთული აღწერისაკენ, საჭიროა ამ მონაკვეთებით შემოსაზღვრული იმ არეების განსაზღვრა, რომლებიც ერთსა და იმავე ობიექტს მიეკუთვნება. გარდა ამისა, საჭიროა აგრეთვე განისაზღვროს ობიექტების ურთიერთგანლაგება. ასეთი ინფორმაციის ცოდნა მოგვცემს შესაძლებლობას გადავიდეთ გამოსახულების მარტივი აღწერიდან (წრფის მონაკვეთების, წრეწირის რკალების დონე) უფრო რთული აღწერისაკენ (ობიექტების ცალკეული არეების (ნაწილების) დონე).

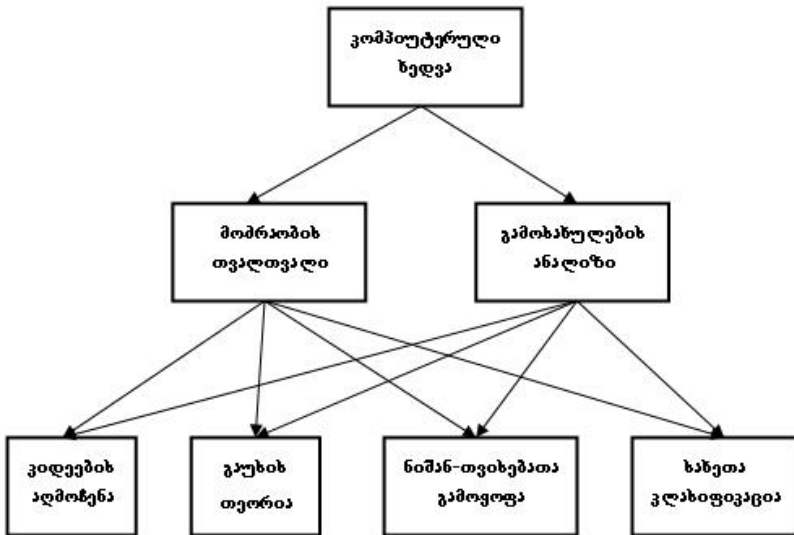


სურ. 14

სხვა სიტყვებით რომ ვთქვათ, გამოსახულების წინასწარი დამუშავების ეტაპზე ხდება ესკიზის (გამოსახულების კონტურული სურათის) ფორმირება, რის შედეგადაც მიიღება გამოსახულების დეტალური, მაგრამ მოუწესრიგებელი აღწერა. შემდეგ ამ აღწერიდან ხდება უფრო ეკონომიური, სტრუქტურირებული აღწერის ფორმირება, რაც უკვე გარკვეული გადაწყვეტილების მიღების შესაძლებლობას იძლევა. პირველ ეტაპს ეწო-

დება გამოსახულების ანალიზი, ხოლო შემდგომ დამუშავებას კი სცენის ანალიზი. ამგვარად, გამოსახულების ანალიზი იწყება გამოსახულებიდან, ხოლო სცენის ანალიზი კი – ესკიზიდან. სცენის ანალიზის ამოცანის მიზანია გამოსახულების სიმბოლური აღწერის მიღება.

კომპიუტერული ხედვის ის მნიშვნელოვანი საკვანძო საკითხები, რომლებიც ამ საგნის ძირითად შინაარსს წარმოადგენს, მოცემულია ბლოკ-სქემის სახით 1.5 სურათზე.



სურ. 1.5

განვიხილოთ ეს საკითხები მოკლედ, რათა გავერკვეთ იმაში, თუ რა კონკრეტული ამოცანების გადაწყვეტას ისახავენ ისინი მიზნად.

კომპიუტერული ხედვის ორ მთავარ საკითხს წარმოადგენს: მობრაობის თვალთვალი (motion tracking) და გამოსახულების ანალიზი (image analysis).

მოძრაობის თვალთვალის საშუალებით ხდება მოძრავი ობიექტების აღმოჩენა დინამიკურ სცენებზე. ეს საკითხი კომპიუტერულ ხედვაში საკვანძო ეტაპია, რადგან ის წარმოადგენს გამოსახულების წარმატებული წინასწარი დამუშავების და სეგმენტაციის ეტაპების შედეგს. ის ძლიერ ზეგავლენას ახდენს გამოსახულების შემდგომი ანალიზისა და ამოცნობის ამოცანაზე. ციფრული გამოსახულებების დამუშავების ისეთ ამოცანაში, როგორცაა სატრანსპორტო სცენის ანალიზი, ცხადია, შეუძლებელია რაიმე სატრანსპორტო მონაცემების აღმოჩენა მანამ, სანამ გამოსახულებაზე თვით ტრანსპორტი არ იქნება აღმოჩენილი.

გამოსახულების ანალიზი იკვლევს იმ მონაცემებს, რომლებიც კომპიუტერული ხედვის ამოცანების გადასაწყვეტად წარმოდგენილი არიან გამოსახულების სახით. ეს საკითხი მოიცავს ორ ქვეთემას: ა) **ნიშან-თვისებათა გამოყოფა** (feature extraction) და ბ) **სახეთა კლასიფიკაცია** (pattern classification). ობიექტის კიდეების აღმოჩენა და კონტურების გამოყოფა წარმოადგენს ნიშან-თვისებათა გამოყოფის ორ ეტაპს, ხოლო ობიექტების ამოცნობა და ობიექტების შეპირისპირება კი – სახეთა კლასიფიკაციის.

კიდის აღმოჩენა (edge detection) გამოყოფს და ლოკალიზებას უკეთებს გამოსახულების იმ წერტილებს (პიქსელებს), რომელთა მიდამოში (მეზობლობაში) ადგილი აქვს სიკაშკაშის მკვეთრ ცვლილებას. ამასთან, უფრო მაღალი დონის პროცედურების შესრულება, როგორცაა ობიექტის **კონტურების გამოყოფა** და ამოცნობა, მნიშვნელოვანწილადაა განპირობებული გამოყოფილი კიდეების სიზუსტითა და დასრულებულობით. ობიექტის კონტური თავისთავში მოიცავს ობიექტის ფორმისთვის დამახასიათებელ საჭირო ინფორმაციას. მოცემულ 2D გამოსახულებაზე არსებული ობიექტის კონტურის გამოყოფა წა-

რმოადგენს იმ მიდგომის მნიშვნელოვან ნაბიჯს, რომელიც ეფუძნება ნიშან-თვისებაზე დამყარებული ობიექტების ამოცნობას.

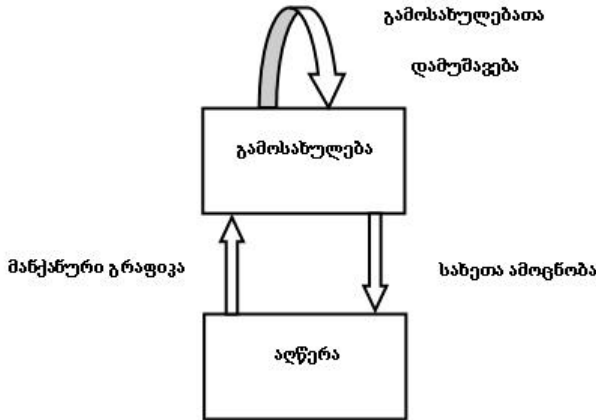
გაუსის თეორია (Gaussian theory) გამოიყენება გამოსახულებაზე არსებული ობიექტების კიდების აღმოსაჩენად და უზრუნველყოფს არა მარტო მათ კარგ აღმოჩენას, არამედ ასევე ზუსტ ლოკალიზაციასაც. გამოსახულებაზე გაუსის თეორიის გამოყენების საფუძველს წარმოადგენს გამოსახულების სიგნალის მრავალშკალიანი (multi-scale) წარმოდგენა. დამტკიცებულ იქნა, რომ გაუსის ბირთვი არის ერთადერთი ბირთვი, რომელსაც შეუძლია მრავალშკალიან გამოსახულებაზე დაამუშაოს ბევრი სპეციფიური მახასიათებელი.

მიუხედავად იმისა, რომ კომპიუტერული გრაფიკა და კომპიუტერული ხედვა ერთმანეთის საპირისპირო ამოცანებს წყვეტენ, მათ შორის მაინც არსებობს კავშირი – თანაკვეთა. მაგალითად, მათი საერთო ინტერესის ამოცანას წარმოადგენს კომპიუტერში გამოსახულებების წარმოდგენის ხასიათი (მონაცემთა სტრუქტურები, შენახვა და ძიება, და ა.შ.). ურთიერთკავშირი კომპიუტერულ გრაფიკას, გამოსახულების დამუშავებასა და სახეთა ამოცნობას შორის სქემატურად ნაჩვენებია 1.6 სურათზე.

კომპიუტერული ხედვა გამოიყენება ადამიანის მოღვაწეობის მრავალ სფეროში, როგორცაა დოკუმენტების დამუშავება, დისტანციური ზონდირების სისტემები, რადიოლოგია, მიკროსკოპია, რობოტების სამრეწველო მართვა და კონტროლი.

კომპიუტერული ხედვის ბევრ ამოცანაში ინფორმაციის სწრაფი დამუშავება წარმოადგენს ერთ-ერთ მთავარ მოთხოვნას. ტიპური ხედვის სისტემები წარმოადგენს სისტემებს, რომლებიც რეალურ დროში ფუნქციონირებს და, შესაბამისად, გამოსახულების ან გამოსახულებების მიმდევრობის (კადრების

მიმდევრობის) დამუშავება უნდა ხდებოდეს ძალიან მცირე დროში. აქედან გამომდინარე, კომპიუტერული ხედვის სისტემებში კომპიუტერების სწრაფქმედებას და მესხიერების მოცულობას ენიჭება ძალიან დიდი მნიშვნელობა.



სურ. 1.6

წარმოდგენილ სახელმძღვანელოში განხილული იქნება კომპიუტერული ხედვის როგორც ფუნდამენტური, ასევე გამოყენებითი საკითხები, მათ შორის შემდეგი უმნიშვნელოვანესი თემები:

სცენის აღქმა – განიხილება გადამწოდების საშუალებით რეალური სცენების გამოსახულებების მიღების ხერხები; გამოსახულებაზე რეალური ობიექტების თვისებების წარმოდგენის საშუალებები, მაგალითად, მასალის, ფორმის, განათებულობის და სივრცითი მიმართებების.

ინფორმაციის კოდირება – გამოსახულების საშუალებით დგინდება ის ინფორმაცია, რაც სამგანზომილებიანი სამყაროს გაგების შესაძლებლობას იძლევა. კერძოდ, ინფორმაცია ობიექტ-

ბის იდენტიფიკაციისათვის, მათი გეომეტრიული და ტექსტურული თვისებებისა და მოძრაობის შესახებ.

წარმოდგენა – განისაზღვრება თუ როგორი სახით შეიძლება წარმოდგენილ იქნეს კომპიუტერული ხედვის სისტემაში ობიექტების, მათი ნაწილების, თვისებებისა და ურთიერთობების აღწერა.

აღგორითმები – განიხილება გამოსახულების დამუშავების არსებული მეთოდები და რეალური ობიექტებისა და სცენების აღწერის ხერხები.

1.2. ციფრული გამოსახულების სტრუქტურა

ციფრული გამოსახულებები შეიძლება წარმოდგენილ იქნეს სხვადასხვა სახით, მაგალითად, ნახატის, ტექსტური გვერდის, ადამიანის პირისახის, კატალოგში წარმოდგენილი გასაყიდი საქონლის სახით. მათი გამოყენების სფერო ძალიან ფართოა, კერძოდ მედიცინაში, ბიოლოგიაში, დეტალების დამზადებისას, დედამიწის ზედაპირის ზონდირებისას და ა.შ.

ციფრული გამოსახულება შედგება სტრიქონებისა და სვეტების ფიქსირებული რაოდენობისაგან. თითოეული ამ სტრიქონისა და სვეტის გადაკვეთაში მოთავსებულია გამოსახულების ელემენტი – ე.წ. პიქსელი (pixel). ეს სიტყვა წარმოადგენს ორი სიტყვის – “გამოსახულების ელემენტი” (picture element), შემოკლების შედეგს. პიქსელი შეიძლება წარმოვიდგინოთ როგორც წერტილი ციფრულ გამოსახულებაზე, რომელშიც სიკაშკაშე ღებულობს სხვადასხვა დისკრეტულ მნიშვნელობებს. ხშირად ეს არის მნიშვნელობები 0-დან 255-ის ჩათვლით. კოდირების სქემიდან გამომდინარე 0 შეიძლება შეესაბამებოდეს სიკაშკაშის ყველაზე მცირე მნიშვნელობას

(შავს), ხოლო 255 – ყველაზე დიდ მნიშვნელობას (თეთრს), ან პირიქით.

ფერად გამოსახულებებში ყოველ პიქსელში გვაქვს სამი რიცხვი: ერთი წითელი, ერთი ცისფერი და ერთი მწვანე ფერისთვის. ყველაზე უფრო ხშირად ციფრული გამოსახულების ასახვა ხდება მონიტორზე, რომელიც შეიძლება წარმოვიდგინოთ როგორც მესხიერების მქონე სატელევიზიო ეკრანი. ფერადი გამოსახულების შემთხვევაში პიქსელის გამონათებისათვის საჭიროა ენერჯის გადაცემა ეკრანის სხვადასხვა ლუმინესცენციური მასალისგან დამზადებული სამი მეზობელი უბნისათვის.

გამოსახულებათა ციფრული დამუშავებისა და ანალიზის სფეროში, საზოგადოდ, სხვადასხვა პარამეტრებისთვის მიღებულია გარკვეული აღნიშვნები. ეს აღნიშვნები შეესაბამება როგორც ვიდეო მონაცემებთან დაკავშირებულ სტანდარტებს, ასევე ალგორითმებს და ა.შ. ზოგიერთი ასეთი პარამეტრისთვის პრაქტიკაში გამოყენებული მნიშვნელობები მოცემულია 1.1 ცხრილში.

ცხ. 1.1

პარამეტრი	აღნიშვნა	რიცხვითი მნიშვნელობა
სტრიქონები	N	256, 512, 525, 625, 1024, 1035
ხვეტები	M	256, 512, 768, 1024, 1320
სიკაშკაშის დონეები	Q	2, 64, 256, 1024, 4096, 16384

პრაქტიკაში საკმაოდ ხშირად საქმე გვაქვს $N = M = 2^K$, ($K = 8, 9, 10$) განზომილების ციფრულ გამოსახულებებთან. რაც შეეხება განსხვავებული ნახევარტონების რაოდენობას გამოსა-

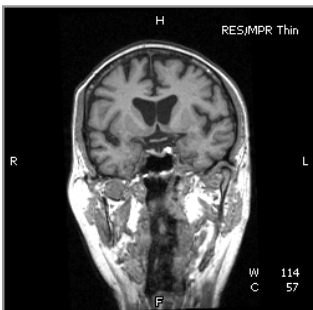
ხულებზე, ის განისაზღვრება 2-ის ხარისხით, $Q = 2^B$, სადაც B არის ბიტების რაოდენობა სიკაშკაშის დონეების ორობითი წარმოდგენისას. როცა $B > 1$, ჩვენ გვაქვს ნახევარტონური გამოსახულება, ხოლო როცა $B = 1$ – ბინარული ანუ შავ-თეთრი გამოსახულება.

1.3. ციფრული გამოსახულებების გამოყენების მაგალითები

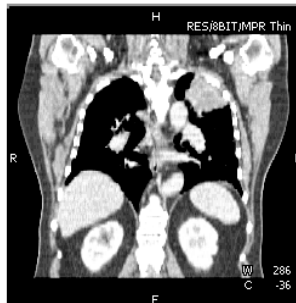
ციფრული გამოსახულებების გამოყენების სფერო ძალიან ფართოა. კომპიუტერების მექანიკების გაზრდამ და ტელეკომუნიკაციური არხების გამტარუნარიანობის ზრდამ შესაძლებელი გახადა გამოსახულებების მონაცემთა ბაზების შემუშავება. დიდი რაოდენობის გამოსახულებებთან ეფექტური ურთიერთობისთვის საჭიროა ისეთი მოხერხებული საშუალებების შემუშავება, რომლებიც მათ მომხმარებლისთვის ადვილად მისაწვდომს გახდის. მონაცემთა ბაზებთან ურთიერთობის უკვე არსებული ჩვეულებრივი ტექნოლოგიები შეიძლება გამოყენებულ იქნეს ისეთ გამოსახულებებთან წვდომისთვის, რომელთაც აქვთ ტექსტური ჭდეები. ასევე მნიშვნელოვანია მონაცემთა ბაზებში გამოსახულებების მოძიება მათი შინაარსიდან გამომდინარე. მაგალითად, მუზეუმის მონაცემთა ბაზაში შენახული სურათებიდან ისეთი სურათების შერჩევა, რომლებსაც ფერების ერთნაირი პალიტრა აქვს, ან ისეთი პეიზაჟების შერჩევა, რომლებზედაც ადამიანი არ არის გამოსახული. ასევე არქიტექტორს შეუძლია მის მიერ მითითებული გამოსახულება-ნიმუშის მიხედვით მონაცემთა ბაზაში მოიძიოს ყველა ის შენობა, რომელსაც მისთვის საინტერესო შენობის მსგავსი ფასადი აქვს.

გამოსახულებების საინსპექციო კონტროლი შეიძლება გამოყენებულ იქნეს მათზე არსებული დეტალების ვარგისიანობ-

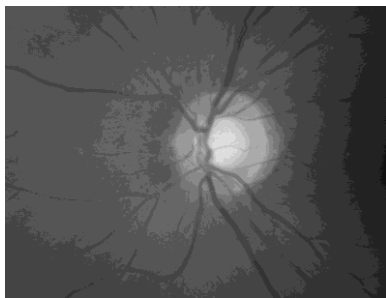
ის დასადგენად. მაგალითად, თუ დეტალზე გაკეთებულია საჭანტიკე ხვრელები, მაშინ ასეთი დეტალის გამოსახულებაზე შესაძლებელია მათი რაოდენობის დათვლა და ამით დეტალის შემდგომი გამოყენების შესაძლებლობის დადგენა; სამედიცინო გამოსახულებების, მაგალითად, (ა) თავის ტვინის, (ბ) შინაგანი ორგანოების, (გ) თვალის ანალიზი შესაძლებლობას იძლევა აღმოჩენილ იქნეს სისხლძარღვების დაზიანებები, სხვადასხვა ანომალიები, სიმსივნეები (სურ. 1.7).



ა



ბ



ბ

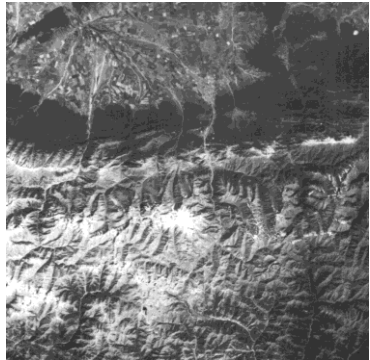
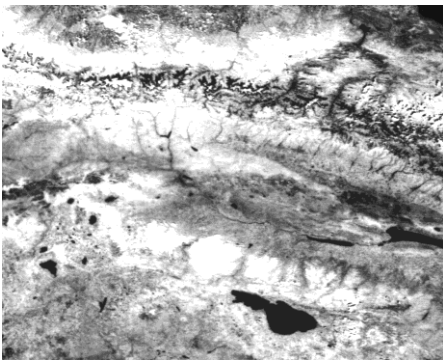
სურ. 1.7

ინფორმაციულ სისტემებში გავრცელებულ ამოცანას წარმოადგენს ნაბეჭდ ტექსტებში არსებული ინფორმაციის

გარდაქმნა ციფრულ სახეში. მაგალითად, ინტერნეტის საშუალებით საჭიროა რაიმე ძველი წიგნის პოვნა, ან რომელიმე დეტალის ნახაზის გარდაქმნა ფაილში, სადაც ჩაწერილი იქნება მონაცემები ამ დეტალის გეომეტრიის შესახებ, რომელიც თავის მხრივ შემდგომში გამოყენებული იქნება ამ დეტალის დასამზადებლად პროგრამულ ჩარხზე.

ხელოვნური თანამგზავრების საშუალებით რეგულარულად ხდება დედამიწის ზედაპირის დიდი ნაწილის სკანირება და შემდეგ მისი გადმოცემა დედამიწაზე ციფრული გამოსახულების სახით (სურ. 1.8).

ამ გამოსახულებების შემდგომი დამუშავება ბევრ სასარგებლო ინფორმაციას იძლევა. მაგალითად, მდინარის წყალგამყოფზე არსებული თოვლის რაოდენობა შეიძლება იყოს ძალიან მნიშვნელოვანი კაშხლებში წყლის რაოდენობის რეგულირებისთვის, წყალმომარაგებისთვის, გარემოს დაცვისთვის და ა.შ.



სურ. 1.8

მრავალ საწარმოო პროცესში არსებობს ეტაპები, რომლებიც ითხოვს ცალკეული დეტალების გადაადგილებას კონვეიერზე. ამ დროს საჭირო ხდება დეტალების მოთავსება სხვა-

დასხვა მექანიზმებში, მათი კონტროლი და ა.შ. ასეთი ოპერაციები ძალიან შრომატევადია ან სახიფათოა სიცოცხლისათვის. ამიტომ, მიზანშეწონილია ხედვის სისტემის მქონე რობოტიზებული კომპლექსის გამოყენება. რობოტის ხედვის სისტემა ახდენს დეტალების ამოცნობას, არკვევს თუ რომელი დეტალები ჩანს კონკრეტულად და როგორ არიან ისინი ორიენტირებული და განლაგებული მუშა სივრცეში. ყოველი დეტალის ამოცნობის პროცესში გამოიყენება პროექტირების ავტომატიზებულ სისტემაში ფორმირებული სამგანზომილებიანი მოდელები. რობოტის ხედვის სისტემა ახდენს შესასვლელი გამოსახულების შედარებას იმ კომპიუტერულ გამოსახულებებთან, რომლებიც გენერირებულია მოდელების მონაცემებისა და სივრცეში ობიექტის მდებარეობის მონაცემების საფუძველზე. ასეთი შედარების შედეგად მიღებული ცუდი დამთხვევები იგნორირდება. ამის შემდეგ აირჩევა საუკეთესო დამთხვევა. დეტალის ამოცნობის შემდეგ რობოტის მართვის სისტემა გადასცემს მანიპულატორს ინფორმაციას იმის შესახებ, თუ როგორ უნდა ჩაავლოს ხელი დეტალს და სად მოათავსოს იგი.

14. ოპერაციები გამოსახულებებზე

ოპერაციები, რომლებიც გამოიყენება გამოსახულებათა დამუშავებისათვის სხვადასხვაა იმისდა მიხედვით, თუ დამუშავების რომელ ეტაპზე ვიმყოფებით, რა ტიპის შედეგი გვსურს მივიღოთ, გამოთვლების რა რესურსებია საჭირო მოცემული ოპერაციის შესასრულებლად და სხვ. კერძოდ, ზოგიერთი ოპერაცია გამოიყენება გამოსახულების ხარისხის გასაუმჯობესებლად ადამიანის მიერ მათი კარგი ვიზუალური აღქმის მიზნით, სხვები კი იმ ინფორმაციის მისაღებად, რომელიც საჭიროა გამოსახულების ავტომატური დამუშავებისათვის; ზოგიერთი ოპე-

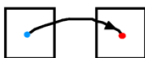
რაციის გამოყენებით კომპიუტერული ხედვის სისტემის გამოსასვლელზე მიიღება ახალი გამოსახულებები, ხოლო სხვა ოპერაციები კი გამოსასვლელი მონაცემების სახით იძლევა აღწერებს არაგრაფიკულ სახეში და სხვ.

ციფრული გამოსახულების მიმართ გამოყენებული ოპერაციები გამოსახულების ერთი სახიდან მეორეში გარდასაქმნელად ან რაღაც სხვა სახით წარმოსადგენად შეიძლება დაიყოს სამ კატეგორიად:

- ა) ოპერაციები გამოსახულების ელემენტზე – პიქსელზე;
- ბ) ოპერაციები მოცემული პიქსელის მიდამოზე – ლოკალური ოპერაციები;
- გ) ოპერაციები მთელ გამოსახულებაზე – გლობალური ოპერაციები.

განვიხილოთ ოპერაციათა თითოეული ამ კატეგორიის მიერ შესრულებული გარდაქმნის არსი $N \times N$ განზომილების ციფრული გამოსახულების და მოცემული პიქსელის $P \times P$ განზომილების მიდამოს (მეზობლობის) შემთხვევაში:

- ა) ოპერაციები გამოსახულების ელემენტზე – პიქსელზე:
გამოსასვლელი სიდიდის მნიშვნელობა განსაზღვრულ კოორდინატში დამოკიდებულია მხოლოდ შესასვლელი სიდიდის იმავე კოორდინატში მნიშვნელობაზე (სურ. 1.9).



სურ. 1.9

- ბ) ოპერაციები მოცემული პიქსელის მიდამოზე – ლოკალური ოპერაციები:

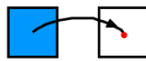
გამოსასვლელი სიდიდის მნიშვნელობა განსაზღვრულ კოორდინატში დამოკიდებულია პიქსელის მიდამოში შემავალი შესასვლელი P^2 რაოდენობის სიდიდეების მნიშვნელობებზე იმავე კოორდინატში (სურ. 1.10).



სურ. 1.10

გ) ოპერაციები მთელ გამოსახულებაზე – გლობალური ოპერაციები:

გამოსასვლელი სიდიდის მნიშვნელობა განსაზღვრულ კოორდინატში დამოკიდებულია შესასვლელი გამოსახულების ყველა ელემენტის – N^2 რაოდენობის პიქსელების მნიშვნელობაზე (სურ. 1.11).



სურ. 1.11

განვიხილოთ ზოგიერთი ოპერაციის გამოყენების მაგალითები.

1.4.1. პიქსელების მოდიფიკაცია მცირე მიდამოებში

პიქსელებში სიკაშკაშის მნიშვნელობები შეიძლება შეიცვალოს მათ მეზობლად განლაგებულ უახლოეს პიქსელებთან ურთიერთკავშირის გათვალისწინებით, მაგალითად, მეზობელი სტრიქონების ან სვეტების პიქსელებით. ხშირად ბინარულ გამოსახულებაზე გამოიყოფა მცირე არეები, სადაც გვხვდება 0-ებისა და 1-ების იზოლირებული მნიშვნელობები. ასეთ შემთხვევაში არეში არსებულ მეზობლებთან თანხვედრის მიზნით ახდე-

ნენ მათ შეცვლას საწინააღმდეგო მნიშვნელობით. ასეთი ოპერაციის მიზანი შეიძლება იყოს უწყვეტი გამოსახულების დისკრეტიზაციის დროს წარმოშობილი ხმაურისგან გამოსახულების გაწმენდა ან გამოსახულების გამარტივება, მაგალითად, მასზე არსებული არაარსებითი წვრილ-წვრილი არეების იგნორირების მიზნით; ან ნაბეჭდი ფურცლის გამოსახულებაზე არსებული დეფექტების მოსაცილებლად. პიქსელების მნიშვნელობების მოდიფიკაციის კიდევ ერთ მაგალითს წარმოადგენს ოპერაცია, რომლის საშუალებითაც ხდება ობიექტის სასაზღვრო პიქსელების მნიშვნელობების შეცვლა ფონური პიქსელების მნიშვნელობებით. მართალია, ასეთი ოპერაციის შედეგად თვით ობიექტის ზომები მცირდება, მაგრამ სამაგიეროდ მისი საზღვრები უფრო მკაფიო და ცხადი ხდება.

14.2. გამოსახულების ხარისხის გაუმჯობესება

ზოგიერთი ოპერაცია მთელ გამოსახულებაზე ერთი და იგივე დამუშავებას ახორციელებს. შესაძლოა, რომ გამოსახულება იყოს ძალიან მუქ ტონში, მაგალითად, მაქსიმალური სიკაშკაშის მნიშვნელობა ტოლი იყოს 100-ის. აქედან გამომდინარე, გამოსახულების ეკრანზე გამოტანის დროს მისი ვიზუალური ხარისხის გაუმჯობესების მიზნით შესაძლებელია სიკაშკაშის ყველა მნიშვნელობა გაუზარდოთ ორჯერ. ასევე მთელი გამოსახულების ხმაურისგან გასაწმენდად ან არაარსებითი დეტალების მოსაშორებლად ყოველი პიქსელში სიკაშკაშის მნიშვნელობა შეიძლება შეცვალათ მის 3×3 ფანჯარაში არსებულ 9 პიქსელში სიკაშკაშის მნიშვნელობების საშუალო მნიშვნელობით. მეორე მხრივ, გამოსახულებაზე შესაძლებელია განხორციელდეს ობიექტების საზღვრების კონტრასტის გაზრდა,

თუ ყოველი პიქსელის მნიშვნელობას შევცვლით მასსა და მის მეზობლებს შორის კონტრასტის მნიშვნელობით.

14.3. რამდენიმე გამოსახულების კომბინაცია

გამოსახულების ფორმირება შესაძლებელია ორი საწყისი გამოსახულების შეკრებით ან გამოკლებით. გამოსახულებათა გამოკლებას იყენებენ, როცა აინტერესებთ ცვლილების აღმოჩენა დროის რაიმე ინტერვალის განმავლობაში. მაგალითად, გამოსახულებათა გამოკლებით შესაძლებელია მოძრავი ობიექტის საზღვრის შენარჩუნება, თუმცა არა იდეალური სახით. ამავე დროს, იმის გამო, რომ პიქსელებისთვის უარყოფითი მნიშვნელობები არ გამოიყენება, გამოსასვლელ გამოსახულებაში ყველა ცვლილება არ იქნება წარმოდგენილი. ასევე ახალი გამოსახულების ფორმირება შესაძლებელია რაიმე გამოსახულებაზე სხვა გამოსახულების “დამატებით”, რის შედეგადაც პირველისგან განსხვავებული კომბინაცია მიიღება.

1.5. გამოსახულების სიმბოლური (არაგრაფიკული) აღწერის ფორმირება

როგორც ზემოთ აღვნიშნეთ, კომპიუტერული ხედვის ამოცანის საბოლოო მიზანს წარმოადგენს სცენის გამოსახულების სიმბოლური (არაგრაფიკული) აღწერა, ანუ მონაცემთა ისეთი სტრუქტურის მიღება, რომელიც გამოსახულებას არ წარმოადგენს. შესაბამისად არსებობს გამოსახულების ანალიზის ოპერაციები (ე.წ. მაღალი დონის ოპერაციები), რომლებითაც ამ ამოცანის გადაწყვეტა ხდება. გამოსახულების არაგრაფიკული აღწერის მაგალითს წარმოადგენს ბაქტერიების გამოსახულების ანალიზის შედეგად მიღებული მონაცემები სხვადასხვა ტიპ-

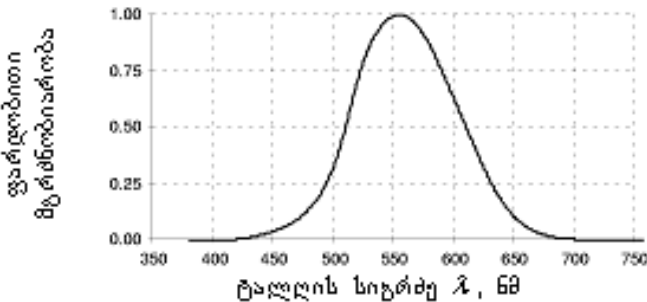
ის ბაქტერიათა რაოდენობის, მათი ფართობების, სიმძიმის ცენტრების მდებარეობის შესახებ; ქალაქის ორ სხვადასხვა ნაწილს შორის სატრანსპორტო ნაკადის მოცულობის განსაზღვრა ვიდეომონაცემებზე ჩატარებული გამოთვლების საფუძველზე; ჟურნალიდან აღებული ტექსტის სკანირების შედეგად მიღებული გრაფიკული გამოსახულებიდან (გარკვეული ოპერაციების ჩატარების შედეგად) მონაცემთა ისეთი სტრუქტურის მიღება, რომელიც შეიცავს ამოცნობილი ტექსტის ცალკეულ ნაწილებს და სურათებს ტექსტიდან.

II თავი. გამოსახულების ფორმირება და წარმოდგენა

2.1. მხედველობის ფსიქოფიზიკური თვისებები

გამოსახულების ფორმირების სისტემის გამოსასვლელზე ჩვეულებრივად ვლენულობთ ფოტოსურათს ან ეკრანის გამოსახულებას, რომელიც შემდეგ უკვე აღიქმება ჩვენ მიერ. აქედან გამომდინარე, ასეთი სისტემების ეფექტური შემუშავებისთვის ძალზე მნიშვნელოვანია ადამიანის მხედველობის სისტემის მექანიზმების ცოდნა.

სინათლე წარმოადგენს სხივურ ენერგიას, რომელიც მოქმედებს მხედველობის ორგანოებზე და საშუალებას აძლევს მათ განახორციელონ მხედველობითი ფუნქციები. ცნობილია, რომ ადამიანისათვის ხილული სინათლე წარმოადგენს ელექტრომაგნიტური გამოსხივების ფორმას, რომელსაც უჭირავს სპექტრის ვიწრო ზოლი ტალღის სიგრძით $\approx (350 \div 780)$ ნანომეტრი(ნმ) (სურ. 2.1).



სურ. 2.1

წარმოდგენილ გრაფიკზე ნაჩვენებია ადამიანის მხედველობის სისტემის სპექტრალური მგრძობიარობის დამოკიდებულება დაცემული სინათლის ტალღის სიგრძისგან. სინათლის წყაროს მიერ გამოსხივებული სინათლე ხასიათდება გამოსხივების ინტენსივობით. ადამიანის თვალს ის ხვდება ან უშუალოდ, ან რომელიმე ობიექტიდან არეკვლის შემდეგ. სინათლის გამოსხივების წყარო ხასიათდება **სპექტრალური სიმკვრივით**, რომელიც განიმარტება როგორც ენერჯის გამოსხივების სიჩქარე ტალღის სიგრძის ერთეულოვან ინტერვალში.

გამოსახულების ფორმირების სისტემებში გამოყენებულია სხვადასხვა ტიპის სინათლის წყაროების გამოსხივება; კერძოდ, მზის, ვოლფრამის ლამპის ვარვარების, სინათლის გამომსხივებელი დიოდის, ჰელიუმ-ნეონის ლაზერის და ა.შ. ამათგან თითოეულს შესაბამისი გამოსხივების სპექტრალური სიმკვრივები გააჩნია. თითოეული ამ წყაროდან წამოსულ გამოსხივებას განსხვავებული შეფერილობა და სხვადასხვა სიძლიერის სიკაშკაშე აქვს და შესაბამისად სხვადასხვანაირად აღიქმება დამკვირვებლის მიერ. არსებობს სინათლის შეგრძნების სამი ძირითადი მახასიათებელი: **განათებულობა (illumination)**, **ფერის ტონი (hue)** და **გაჯერებულობა (saturation)**. თუ მაგალითად გვაქვს ერთნაირი სპექტრალური სიმკვრივის მქონე ორი გამოსხივების წყარო, მაშინ მეტი ინტენსივობის მქონე წყაროდან წამოსული სინათლის სხივი ადამიანის მიერ აღიქმება როგორც უფრო კაშკაშა. მეორე მხრივ, არსებობს მაგალითები, როცა საგანი, რომლის ყველა წერტილი ერთნაირი ინტენსივობით ასხივებს, არ აღიქმება როგორც თანაბრად განათებული. აქედან გამომდინარე, სინათლის ინტენსივობა არ შეიძლება აღებულ იქნეს განათებულობის რაოდენობრივ მახასიათებლად.

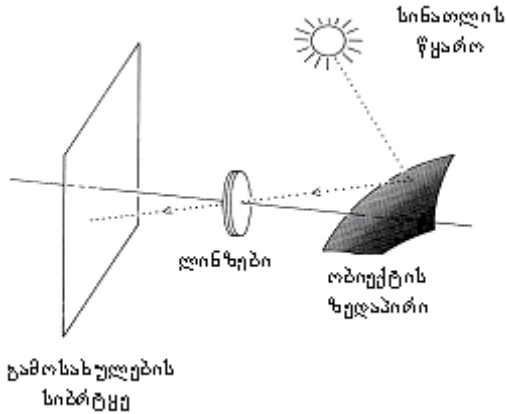
ნიშანს, რითაც ერთი ფერი მეორისგან განსხვავდება, **ფერის ტონი** ეწოდება. ამავე დროს ბუნებაში ვხვდებით ისეთ ფე-

რებსაც, რომლებიც არ ჩანს პრიზმაში სინათლის გავლის შედეგად წარმოშობილ ცისარტყელაში. მაგალითად, ასეთ ცისარტყელაში არ ჩანს მეწამული ფერი, რომელიც მიიღება თანაბარი რაოდენობის წვრილზოლიანი წითელი და ლურჯი ფერების შერევის შედეგად. ამიტომ ტალღის სიგრძე არ წარმოადგენს ფერის ადეკვატურ საზომს.

სინათლის შეგრძნების მესამე მახასიათებელს წარმოადგენს მისი **გაჯერებულობა**. ეს მახასიათებელი შესაძლებლობას იძლევა, რომ ერთმანეთისგან განვასხვაოთ სპექტრალური ფერი ისეთივე ფერის ტონის მქონე პასტელური მკრთალი ფერისაგან. თავისი არსით გაჯერებულობა აღწერს ფერის “სითეთრეს”. რაოდენობრივად გაჯერებულობა შეიძლება გაიზომოს პროცენტებში, რისთვისაც მოცემული ფერი უნდა შედარდეს სპექტრალურ ფერს ფერთა გრაფიკზე. მიუხედავად ამისა, გაჯერებულობა არ გამოიყენება როგორც სინათლის რაოდენობრივი მახასიათებელი.

სამგანზომილებიან გარესამყაროში ჩვენს ირგვლივ არსებული ობიექტების ვიზუალურად აღქმადი სცენა წარმოადგენს გამოსახულების ფორმირების სისტემის (ოპტიკური, ფოტოგრაფიული ან სატელევიზიო) მიერ შექმნილ 2D გამოსახულებას, რომელიც ფიქსირდება გარკვეულ ჩამწერ გარემოში (სურ. 2.2).

ტერმინი “გამოსახულება” აღნიშნავს ბრტყელ ობიექტს, რომლის სახე იცვლება წერტილიდან წერტილამდე. შავ-თეთრი გამოსახულების შემთხვევაში ეს ცვლილება შეიძლება აღიწეროს ერთადერთი პარამეტრით, რომელიც შეესაბამება ობიექტის მოცემული წერტილიდან არეკვლილ და დამკვირვებლის თვალის ბადურაზე დაცემულ სინათლის რაოდენობას. ფერადი გამოსახულების შემთხვევაში კი საჭიროა სამი პარამეტრი, რომლებიც შეესაბამება წითელი-მწვანე-ლურჯი (RGB) ფერების შესაბამის სინათლის რაოდენობას.



სურ. 2.2

გარესამყაროს ობიექტებიდან არეკვლილი სინათლის სხივი ადამიანის ორივე თვალის ბადურაზე დაცემის შედეგად ახდენს ორი გამოსახულების ფორმირებას. ამ ორი გამოსახულების შერწყმის შედეგად ადამიანი აღიქვამს გარესამყაროს სამგანზომილებიან სახეში. ზემოთქმულიდან გამომდინარე, გარესამყაროს სცენის აღქმის ძირითად კომპონენტებს წარმოადგენენ:

- გარესამყაროს სცენა ობიექტებით;
- სცენის განათებულობა;
- სცენიდან არეკვლილი სინათლის სხივების აღქმა.

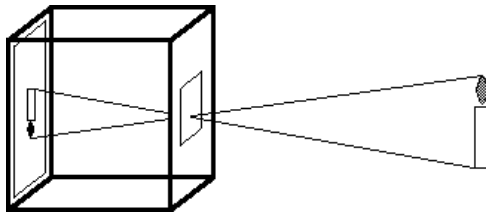
ამ თავში ჩვენ გავეცნობით 2 და 3-განზომილებიანი სცენების ციფრული გამოსახულებების ფორმირების ხერხებს გადამწოდების გამოყენებით. გადამწოდებით შესაძლებელია სხვადასხვა სახის გამოსხივების აღქმა. კერძოდ, გარდა ხილული სპექტრის აღმქმნელი გადამწოდებისა, არსებობს გადამწო-

დები, რომელთა მგრძობიარობა მეტია, ვიდრე 780 ნმ (ინფრა-წითელი), და ასევე გადამწოდები, რომლებიც მგრძობიარეა მოკლელტალლიანი რენტგენის სხივებისა და რადიოდიამაზონის გრძელტალლიანი ტალღების მიმართ.

ორგანზომილებიანი ციფრული გამოსახულება წარმოადგენს სინათლის ინტენსივობის მნიშვნელობების მასივს. მაგრამ ამავე დროს ორგანზომილებიანი გამოსახულება ასევე შეიძლება წარმოადგენდეს სამგანზომილებიანი სცენის პროექციასაც.

2.2. გამოსახულების ფორმირების მოდელი

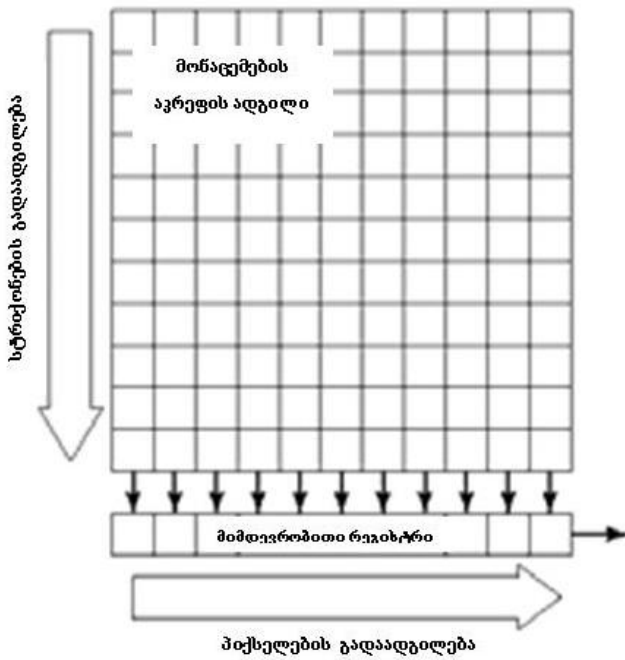
გამოსახულების ფორმირების გეომეტრიული მოდელის როლში შეიძლება განვიხილოთ სამგანზომილებიანი სცენის ყოველი წერტილის პროექცირება (კამერის ობიექტივის ცენტრის გავლით) გამოსახულების სიბრტყეზე. ასეთ დროს გამოსახულების წერტილის სიკაშკაშე დაკავშირებულია სამგანზომილებიან ზედაპირზე მდებარე წერტილის გამოსხივების ინტენსივობასთან. ასეთი მოდელის ფიზიკური რეალიზაცია შესაძლებელია ხერყელიანი ულინზო ობიექტივის მქონე კამერის, ე.წ. ობსკურის კამერის (pinhole camera) მეშვეობით. ობსკურის კამერა წარმოადგენს ყუთის მსგავსი კორპუსის მქონე კამერას პატარა ხერყელით და არის ულინზო კამერის უმარტივესი მოდელი (სურ. 2.3).



სურ. 2.3

ამასთან, რაც უფრო მცირეა ხვრელი, მით უკეთესია ფოკუსი, ანუ მით მკვეთრია გამოსახულება და სინათლის ნაკლები ენერგიაა საჭირო გამოსახულების მისაღებად სცენის ნებისმიერი წერტილიდან.

გამოსახულების ფორმირებისთვის გამოიყენება ასევე მუხტ-კავშირიან პრინციპზე დამყარებული ხელსაწყოები (მკხ) – კამერები (charge-coupled device, CCD), რომლებიც 35 მმ-იანი ფოტოფირის მქონე ფოტოკამერის მსგავსია და აქვთ ლინზური ობიექტივი (სურ. 2.4). განსხვავებით ფოტოკამერისაგან, მკხ-ში გამოსახულების სიბრტყეში ფოტოფირის ნაცვლად მოთავსებულია პატარა მყარტანიანი უჯრედების მატრიცა, რომელიც სინათლის ენერგიას გარდაქმნის ელექტრულ მუხტში.



სურ. 2.4

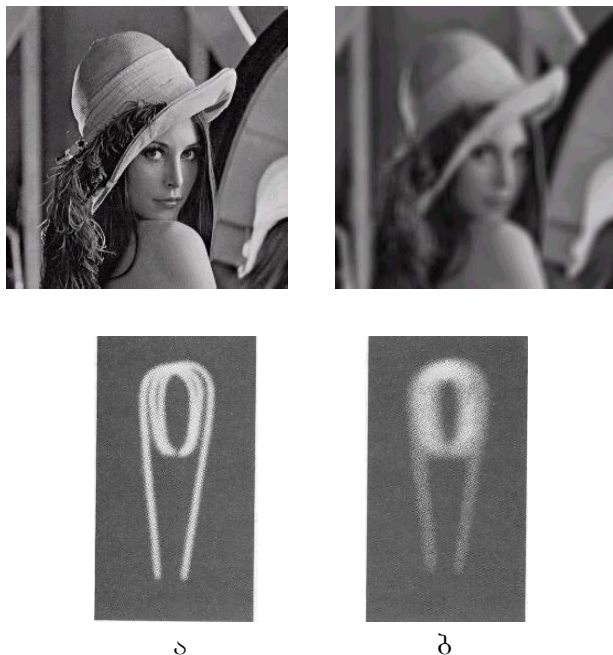
დასაწყისში ყველა უჯრედი სუფთავდება მასში დაგროვებული მუხტებისაგან, ხოლო შემდეგ კი სინათლის ენერჯის ზემოქმედების შედეგად უჯრედებში იწყება მუხტების დაგროვება. დაგროვების დროის მართვა და გამოსახულების კომპიუტერში გადაწერის პროცესი იმართება კომპიუტერის მეშვეობით.

ზემოთქმულიდან გამომდინარე საჭიროა მხედველობაში მივიღოთ ორი რამ:

- ცნობილია, რომ ობიექტივი წარმოადგენს სინათლის შემკრებს. სინათლის წყაროდან გამოსახულების წერტილებზე დაცემული სინათლის ტალღა ვრცელდება სამგანზომილებიანი წერტილიდან გამომავალი სინათლის სხივების მიერ შექმნილი კონუსის შიგნით. ასეთივე კონუსები აქვს სცენის ყველა წერტილს. მაგრამ იმის გამო, რომ ღინზები არაიდეალურია, ასევე სინათლის გამოსხივების დამოკიდებულების გამო ტალღის სიგრძისგან და სხვადასხვა მოვლენების გამო, სინათლის კონუსის მიერ გამოსახულების სიბრტყეზე ჩნდება გადარეცხილი ლაქა, რომელსაც **გაბნევის წრე** (circle of confusion) ეწოდება;
- მკვ-კამერა აგებულია სასრული რაოდენობის ელემენტებისგან და არა უსასრულოდ მცირე წერტილებისგან. ამის გამო, მისი ყოველი უჯრედის რეაქცია განპირობებულია სამგანზომილებიანი ზედაპირის გარკვეული წერტილის მიდამოდან წამოსული და მასზე დაცემული სხივებით.

ამ ორი ეფექტის შედეგად ადგილი აქვს **გამოსახულების გადარეცხვას** (blurring) (სურ. 2.5, (ა) საწყისი გამოსახულება, (ბ)

გადარეცხილი გამოსახულება). ამავე დროს აღნიშნული ეფექტები განსაზღვრავენ გამოსახულებაზე არსებული უმცირესი დეტალების სიმკვეთრესა და ზომებს.



სურ. 2.5

მკხ-კამერის უჯრედების მასივი იწარმოება მიკროსკემის სახით, ზომით $\approx (1 \times 1)$ სმ. პიქსელის გვერდების ზომები მასივში, რომელიც შეიცავს (512×512) პიქსელს, დაახლოებით ტოლია 0.001 დუიმიხა.

არსებობს მკხ-კამერის უჯრედების სიბრტყეზე (ან სტრიქონზე) განლაგების სხვა ხერხებიც. მაგალითად ისეთ ამოცანებში, სადაც საჭიროა გაიზომოს ობიექტის მხოლოდ სიგანე ან გვერდს, რომ მივიღოთ გამოსახულება და მოვასხინოდ მისი

რომელიმე გრძელი ფრაგმენტის კონტროლი, შეიძლება უჯრედების მასივის წრფივად განლაგება. არსებობს წრფივი გადამწოდები, რომელიც შეიცავს სტრიქონში 1000-დან 5000-მდე პიქსელს. ფერადი გამოსახულების ან ნაბეჭდი ტექსტის ციფრული გამოსახულების მისაღებად გამოიყენება წრფივი გადამწოდის მსგავსი პლანშეტური სკანერები. მათში ქაღალდის ფურცლის გამოსახულების ფორმირება ხდება გადაადგილებადი ოპტიკური სისტემის მეშვეობით. წრიული ტიპის შკალების, მაგალითად საათების ან სპიდომეტრების კონტროლისათვის გამოიყენება წრიული ტიპის გადამწოდები.

გადამწოდების საშუალებით მიღებული გამოსახულების აღქმისას ჩნდება რამდენიმე პრობლემა. გამოსახულების ფორმირების იდეალიზებული მოდელი წარმოადგენს ფიზიკური სამყაროს მხოლოდ მიახლოებით მოდელს. რეალურად ადგილი აქვს სხვადასხვა სახის დამახინჯებებს, რომელთა ჯამური ეფექტი ვლინდება როგორც გეომეტრიულ, ასევე სიკაშკაშესთან დაკავშირებულ მახასიათებლებში.

ა) გეომეტრიული დამახინჯებები

გამოსახულების ფორმირებისას გეომეტრიული დამახინჯების მიზეზი შეიძლება სხვადასხვა იყოს. მაგალითად, ის შეიძლება განპირობებული იყოს ოპტიკური სისტემის ობიექტივის დამზადებისას წარმოშობილი დეფექტებით, რომლებიც არ უზრუნველყოფს სინათლის სხივების ზუსტ გარდატეხას.

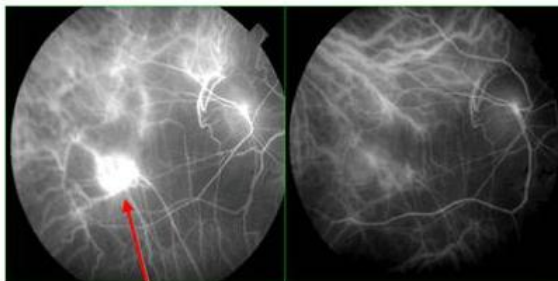
ბ) დისპერსია

სინათლის სხივები ნივთიერებაში გასვლისას გარდატეხდება და განიბნევა. ასეთი დეფექტები თან სდევს აეროკოსმოსურ ფოტოსურათებს. ამის მიზეზი კი ის არის, რომ წყლის ორთქლ-

ისა და ტემპერატურული გრადიენტების გამო ატმოსფერო ლინზის თვისებას იძენს.

გ) ბლუმინგი – ჭარბი სიკაშკაშე

იმის გამო, რომ მკხ-ს დისკრეტული უჯრედები ერთმანეთისგან იზოლირებული არ არის, შესაძლებელია ერთ უჯრედში დაგროვილი მუხტების გადადინება მეზობელ უჯრედებში. **ბლუმინგი** (blooming) დაკავშირებულია ისეთ შემთხვევასთან, როცა ადგილი აქვს მუხტის გადადინებას ბრტყელი გამოსახულების გაკაშკაშებულ არეში. ამის გამო გამოსახულებაზე ჩნდება კაშკაშა “ყვავილი”, რომლის ზომები ჩვეულებრივზე უფრო დიდია (სურ. 2.6, (ა) ბლუმინგი, (ბ) ექსპონირებული გამოსახულება).



ბლუმინგი

ექსპონირებული

გამოსახულება

ა

ბ

სურ. 2.6

დ) მკხ მატრიცის არაერთგვაროვნება

წარმოების პროცესში წარმოშობილი დეფექტების გამო მკხ მატრიცის ელემენტების მგრძობიარობა შესაძლოა ერთმანეთისგან განსხვავდებოდეს. შესაბამისად, მკხ მატრიცის თანაბარი განათებისას შეიძლება განსხვავებული იყოს სხვადასხვა

უჯრედის რეაქციაც. ასეთ შემთხვევაში ინტენსივობის ზუსტი მნიშვნელობის განსაზღვრისათვის შეიძლება საჭირო გახდეს ყოველი პიქსელისთვის თითო – მასშტაბის $s[i, j]$ მამრავლები-სა და $t[i, j]$ წანაცვლებების მთელი მასივის განსაზღვრა. ეს მიიღწევა თუ მკვ მატრიცას თანაბრად გავანათებთ და შემდეგ მოვახდენთ მის კალიბრაციას. ასეთ შემთხვევაში მატრიცის ელემენტში (პიქსელში) ინტენსივობის სიდიდე $I_2[i, j]$ შეიძლება გამოთვლილ იქნეს შემდეგნაირად

$$I_2[i, j] = s[i, j]I_1[i, j] + t[i, j],$$

სადაც $I_1[i, j]$ საწყისი ინტენსივობაა. მკვ მატრიცაზე უკიდურეს შემთხვევაში შეიძლება იყოს რამდენიმე მკვდარი წერტილი, რომელიც საერთოდ არ გენერირებს გამოსასვლელ სიგნალს. ეს დეფექტები მუდავნდება მატრიცის შემოწმებისას და მათ აღმოსაფხვრელად გამოიყენება პროგრამული საშუალებები. ასეთ დროს მკვდარი წერტილის გამოსასვლელ სიგნალს ენიჭება მისი მეზობელი უჯრედების გამოსასვლელი სიგნალების გასაშუალოებული მნიშვნელობა.

ე) ჩამოჭრა და ციკლური დაბრუნება

ანალოგური ციფრული გარდაქმნისას ინტენსივობის მნიშვნელობებმა შეიძლება ძალიან დიდ სიდიდეებს მიაღწიოს. ამიტომ ხდება მათი ჩამოჭრა მაქსიმალურ მნიშვნელობამდე, ან მაღალი ბაიტების ჩამოჭრა, რაც გულისხმობს მათ ციკლურ დაბრუნებას ინტენსივობის მნიშვნელობების დიაპაზონის დასაწყისში. ნახევარტონური გამოსახულების შემთხვევაში ციკლური დაბრუნება გამოსახულებაზე გვაძლევს ნათელ არეს მუქი ცენტრით. ფერადი გამოსახულების შემთხვევაში კი შეიძლება ადგილი ჰქონდეს ფერის შეცვლას.

ვ) ქრომატული დისტორსია

ცნობილია, რომ სხეულიდან არეკვლილი ან სინათლის წყაროდან წამოსული სხვადასხვა სიგრძის სინათლის ტალღები ლინზაში სხვადასხვანაირად გარდატყდება. ამის გამო სცენის ერთი და იგივე უბნიდან წამოსული სხვადასხვა ტალღის სიგრძის მქონე სინათლის ტალღების ფოკუსირება შეიძლება მოხდეს მგრძობიარე გადამწოდის ზედაპირის არა ერთსა და იგივე წერტილში (პიქსელში), არამედ გამოსახულების განსხვავებულ პიქსელებში. მაგალითად, თუ სცენის პერიფერიაზე გვაქვს უბანი, სადაც ადგილი აქვს ინტენსივობის მკვეთრ შავთეთრ გადასვლას, მაშინ გამოსახულებაზე ადგილი ექნება ინტენსივობის არა მკვეთრ, არამედ თანდათანობით (წრფივ) გადასვლას, რომელიც რამდენიმე პიქსელს მოიცავს (სურ. 2.7).



სურ. 2.7

2.3. კომპიუტერში გამოსახულების შეყვანის ტექნიკური საშუალებები

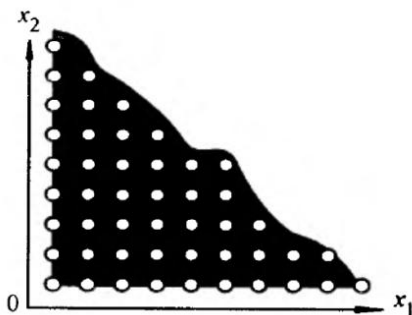
ოპტიკური გამოსახულებების კომპიუტერის მეხსიერებაში შეყვანისათვის და მიღებული შედეგების ვიზუალიზაციისათვის არსებობს სხვადასხვა ტექნიკური საშუალებები. ისინი ერთმანეთისგან განსხვავდება როგორც ელექტრომაგნიტური გამოსხივების აღქმის პრინციპებით, ასევე ელექტრომექანიკური აგებულებით. კერძოდ, ვიზუალიზაციისათვის გამოიყენება მაღალი გარჩევისუნარიანობის მქონე ფერადი დისპლეები და ასახვის სხვა საშუალებები.

გამოსახულების შეყვანა კომპიუტერის მეხსიერებაში ხორციელდება ვიდეოგადამწოდების მეშვეობით, რომლებსაც გამოსახულებაზე არსებული სიკაშკაშის ოპტიკური განაწილება გადაჰყავს ჯერ ელექტრულ სიგნალში და შემდეგ ხდება მისი კოდირება. იმის გამო, რომ გამოსახულება წარმოადგენს ორი სივრცითი ცვლადის ფუნქციას, ხოლო ელექტრული სიგნალი კი ერთი ცვლადის, კერძოდ, დროის ფუნქციას, ამიტომ შესაბამისი გარდაქმნის განხორციელებისთვის იყენებენ გაშლას. მაგალითად, სატელევიზიო კამერის გამოყენებისას გამოსახულების წაკითხვა ხდება სტრიქონ-სტრიქონ. ამასთან ყოველი სტრიქონის ფარგლებში სიკაშკაშის დამოკიდებულება სივრცით x კოორდინატზე გარდაიქმნება ელექტრული სიგნალის ამპლიტუდის t დროზე დამოკიდებულების პროპორციულად. გადასვლა ერთი სტრიქონის ბოლოდან მომდევნოს დასაწყისზე ხდება მყისიერად.

ვიდეოგადამწოდების სახით ფართო გამოყენება ჰპოვეს ასევე ფოტოდიოდების მატრიცებმა და მუხტ-კავშირიანი პრინციპის გამოყენებით აგებულმა ხელსაწყოებმა. მატრიცული ტიპის ვიდეოგადამწოდების გამოყენებისას იქმნება შთაბეჭდილე-

ბა, თითქოს გამოსახულებას ვაკვირდებით მრავალი გამჭვირვალე უჯრედისაგან შედგენილი ეკრანის უკან (მიღმა). მატრიცული ვიდეოგადამწოდის ფრაგმენტი ნაჩვენებია 2.8 სურათზე. ასეთი ტიპის თანამედროვე ვიდეოგადამწოდებში უჯრედების რაოდენობა ტოლია ან აღემატება (1024×1024).

დიდი ზომის გამოსახულების, მაგალითად (500×500), სადაც თითოეულ პიქსელში სიკაშკაშის მნიშვნელობა ჩაიწერება ერთი ბაიტი განზომილების მქონე მასივით, კომპიუტერში შეყვანისათვის საჭირო იქნება მეხსიერება, რომელიც დაიტევს მასივს განზომილებით 250 000 ბაიტი. ეს საკმაოდ დიდი მასივია და, შესაბამისად, მოითხოვს კამერისგან დიდ მეხსიერებას. ამის გამო არსებობს სპეციალური პლატები, რომლებიც მოთავსებულია კომპიუტერში და მონაწილეობას იღებს კამერის მართვაში. დღეს უკვე თანამედროვე მწარმოებლები გეთავაზობენ კამერებს, რომელთაც რამდენიმე გამოსახულების შენახვა შეუძლია თვით კამერის შიგნით და რომლებიც, საჭიროებისამებრ, ნებისმიერ მომენტში შეიძლება შეყვანილ იქნეს კამერიდან კომპიუტერში.



სურ. 2.8

გამოსახულების მიღება ასევე შესაძლებელია ვიდეოკამერით, რომელიც საშუალებას იძლევა მოხდეს გამოსახულებების მიმდევრობის ჩაწერა ტიპური სიხშირით 30კადრი/წმ. ეს ადამიანს საშუალებას აძლევს აღიქვას არა მარტო ობიექტის სივრცითი მახასიათებლები ცალკეული კადრების სახით, არამედ ობიექტი მოძრაობაში, რაღაც გარკვეული დროის განმავლობაში. იმისათვის, რომ მოძრაობის აღქმა იყოს უწყვეტი, ადამიანი იყენებს სიხშირეს 60 ნახევარკადრი/წმ. ახევარკადრების მიმდევრობა წარმოადგენს გამოსახულების ყველა ლუწი და კენტი სტრიქონებისგან შედგენილ და ერთმანეთის მიყოლებით რიგრიგობით მოძრავ პორციების მიმდევრობას. ამავე დროს პარალელურად ხდება ხმის ჩაწერაც. ვიდეოკამერებს, რომლებიც გამოიყენება გამოსახულებების მისაღებად მანქანური დამუშავებისათვის, შესაძლებლობა აქვს, რომ გამოსახულების ჩაწერა მოახდინოს გამოყენებითი ამოცანისათვის საჭირო სიჩქარით. შესაბამისად მათში არ გამოიყენება ნახევარკადრების ტექნოლოგია. ვიდეომიმდევრობის კადრების ერთმანეთისაგან გამოყოფა ხდება ჭდეების საშუალებით. მონაცემთა დიდი მასივის შემთხვევაში კი მისი მოცულობის შესამცირებლად იყენებენ გამოსახულებების შეკუმშვის მეთოდს. სოფლიოში ყველაზე მეტად გამოყენებული ვიდეო სისტემებია NTSC, PAL და SECAM.

2.4. გამოსახულების მათემატიკური მოდელი

გამოსახულებათა კომპიუტერული დამუშავება და ანალიზი თავისთავად მოითხოვს გამოსახულების აღმწერი მათემატიკური მოდელის აგებას. გამოსახულებების შესახებ საუბრისას ერთმანეთისგან უნდა განვასხვაოთ ორი ცნება – უწყვეტი და ციფრული გამოსახულებები.

2.4.1. უწყვეტი გამოსახულების აღწერა

(ა) დეტერმინისტული მიდგომა

სინათლის წყაროს მიერ გამოსხივებული ენერგიის განაწილება სივრცული x და y კოორდინატების, t დროის და λ ტალღის სიგრძეების მიხედვით აღიწერება ფუნქციით $C(x, y, t, \lambda)$. ასევე ცნობილია, რომ გამოსხივების ენერგია პროპორციულია ელექტრული ველის ამპლიტუდის კვადრატის და ამის გამო ნამდვილ დადებით რიცხვს წარმოადგენს. ამავე დროს ოპტიკური სისტემის მიერ გამოსახულების ფორმირებისას გამოსახულების მაქსიმალური სიკაშკაშე და ზომები შეზღუდულია მისი მაფორმირებელი და მაფიქსირებელი გარემოს გარკვეული თვისებების (სინათლის მგრძნობიარე ფირის გაჯერებულობის ან კინესკოპის ლუმინოფორის გადახურების) და ზომების გამო.

თეორიულად უწყვეტი გამოსახულება ფორმირდება იდეალური ოპტიკური სისტემის მიერ, ანუ ისეთის, რომელსაც უსასრულო სიზუსტე გააჩნია. შედეგად მიიღება ორგანზომილებიანი უწყვეტი გამოსახულება $f(x, y)$.

განსაზღვრება 2.1. *უწყვეტი გამოსახულება – ეს არის ორგანზომილებიანი $f(x, y)$ გამოსახულება, რომელიც ხასიათდება წარმოდგენის უსასრულო სიზუსტით როგორც სივრცითი x და y კოორდინატების, ასევე ინტენსივობის მნიშვნელობების მიმართ სივრცის ნებისმიერ (x, y) წერტილში.*

ასეთი გამოსახულების შემთხვევაში (X, Y) სიბრტყის რომელიმე (x, y) წერტილში ინტენსივობა წარმოადგენს ობიექტის წერტილების გამოსხივების ინტენსივობის ფუნქციას (picture

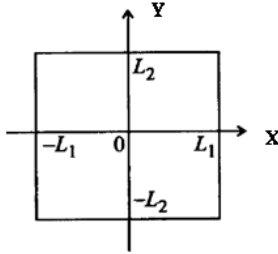
function) ობიექტის სიბრტყეში არსებული შესაბამისი (x_1, y_1) წერტილის მიდამოდან. შავ-თეთრი გამოსახულება ფორმალურად აღიწერება როგორც ორი ნამდვილი ცვლადის ნამდვილი ფუნქცია. ჩვეულებრივად ასეთ ფუნქციად გამოიყენება ინტენსივობის ფუნქცია $f(x, y)$, რომელიც გამოსახულების სიბრტყის ნებისმიერ (x, y) წერტილში (პიქსელში) ღებულობს ნამდვილ მნიშვნელობას და განსაზღვრავს ინტენსივობის სიდიდეს ამ წერტილში. ზემოთქმულიდან გამომდინარე, იმის გამო, რომ გამოსახულებიდან არეკვლილი და დამკვირვებლამდე მიღწეული სინათლის რაოდენობა სასრულია და თან არაუარყოფითი, ამიტომ ფუნქცია, რომელიც აღწერს გამოსახულებას, შესაბამისად, შემოსაზღვრულია და არაუარყოფითია. აქედან გამომდინარე შეიძლება ვთქვათ, რომ რეალური გამოსახულებისთვის არსებობს ინტენსივობის ზედა ზღვარი (გამოსახულების მაქსიმალური ინტენსივობა) M , ანუ გამოსახულება წარმოიდგინება ასეთი სახის ფუნქციით

$$0 \leq f(x, y) \leq M, \forall (x, y) \in R^2.$$

ამასთან, იმის გამო, რომ გამოსახულების ზომები შეზღუდულია თვით ოპტიკური მაფორმირებელი სისტემის მახასიათებლებით და ასევე გამოსახულების მარეგისტრირებელი გარემოს ზომებით, ამიტომ ვთვლით, რომ ყველა გამოსახულება ნულისგან განსხვავებულია მხოლოდ გარკვეულ მართკუთხა

$$-L_1 \leq x \leq L_1, -L_2 \leq y \leq L_2$$

არეში (სურ. 2.9).



სურ. 2.9

გამოსახულების ინტენსივობის მნიშვნელობას რომელიმე (x, y) წერტილში (პიქსელში) ასევე უწოდებენ **ნახევარტონის დონეს** ან **სიკაშკაშეს**.

ზემოთქმულიდან გამომდინარე, რეალური უწყვეტი $f(x, y)$ გამოსახულება ეს არის სასრულ მართკუთხა არეში მოცემული ორი სივრცითი ცვლადის შემოსაზღვრული ფუნქცია.

(ბ) სტატისტიკური მიდგომა

გამოსახულების ფორმალური აღწერისთვის გარდა დეტერმინისტული მიდგომისა, რომელიც გულისხმობს ისეთი პროცესების აღწერას, რომელთა მიმდინარეობაც ცალსახადაა განსაზღვრული, არსებობს ფიზიკური სისტემის (მაგალითად, სიგნალის) დროსა და სივრცეში ისეთი ცვლილებები, რომელთა ქცევის ზუსტი პროგნოზირება შეუძლებელია. სწორედ ასეთი სისტემის (სიგნალის) ქცევის აღწერისთვის შემოტანილ იქნა შემთხვევითი პროცესის ცნება.

ყოველი ექსპერიმენტის (ცდის) შედეგს წარმოადგენს დროის რაღაც ინტერვალზე განსაზღვრული შემთხვევითი ფუნქცია. თუ f ერთი ცვლადის შემთხვევითი ფუნქციაა, მაშინ ლაპარაკობენ შემთხვევით პროცესზე, ხოლო თუ ორი ან მეტი ცვლადის – შემთხვევით ველზე. გამოსახულებების აღწერისა-

თვის ხშირად იყენებენ შემთხვევითი ორგანზომილებიანი მიმდევრობების მათემატიკურ მოდელებს. ამავე დროს, უნდა აღინიშნოს, რომ შემთხვევითი სიგნალის ყოველი ცალკეული რეალიზაცია უკვე დეტერმინისტულ ფუნქციას წარმოადგენს. შემთხვევითი პროცესის თავისებურებები ვლინდება რეალიზაციათა ერთობლიობის ან მთელი ანსამბლის შესწავლისას. იმის გამო, რომ ეს ანსამბლი ალბათურია, შესაბამისად შემთხვევითი პროცესის მახასიათებლებიც ალბათურია. ამიტომ, შემთხვევითი პროცესის რეალიზაციების ინდივიდუალური თავისებების აღწერისთვის უნდა გამოვიყენოთ ალბათობის თეორიის ცნებები.

სიგნალის ქცევა, კერძოდ კი მისი ამპლიტუდების განაწილება, მჭიდროდაა დაკავშირებული ალბათობათა განაწილების ცნებასთან. გამოსახულების მოცემული არისთვის, რომელიც შეიძლება შესასვლელ გამოსახულებასაც წარმოადგენდეს, შეგვიძლია განვსაზღვროთ ამ არეში სიკაშკაშის განაწილების ფუნქცია და სიმკვრივის ფუნქცია.

განსაზღვრება 2.2. სიკაშკაშის განაწილების ფუნქცია
 $P(a)$ წარმოადგენს იმის ალბათობას, რომ არიდან აღებული სიკაშკაშის მნიშვნელობა ნაკლები ან ტოლი იქნება სიკაშკაშის მოცემულ a მნიშვნელობაზე, ამასთან, როცა a იზრდება $-\infty$ -დან ∞ -კენ, $P(a)$ იზრდება 0 -დან 1 -მდე. განაწილების $P(a)$ ფუნქცია არის არაკლებადი a -ს მიმართ და შესაბამისად $dP/da \geq 0$.

განსაზღვრება 2.3. სიკაშკაშის სიმკვრივის ფუნქცია წარმოადგენს იმის ალბათობას, რომ არიდან აღებული $P(a)$ განაწილების ფუნქციის მქონე სიკაშკაშის მნიშვნელობა მოხვდება

$(a, a + \delta a)$ ინტერვალში, რაც გამოისახება როგორც $p(a)\delta a$, სადაც $p(a)$ - სიმკვრივის ფუნქციაა

$$p(a)\delta a = \left(\frac{dP(a)}{da}\right)\delta a.$$

განაწილების $P(a)$ ფუნქციის არაკლებადობის გამო სიმკვრივის ფუნქცია უნდა აკმაყოფილებდეს შემდეგ პირობებს:

$$p(a) \geq 0, \text{ და } \int_{-\infty}^{\infty} p(a)da = 1. \quad (2.1)$$

ციფრული გამოსახულებისთვის, სადაც სიკაშკაშის მნიშვნელობები მთელ რიცხვებს წარმოადგენს, δa განიმარტება როგორც სიკაშკაშის მნიშვნელობათა ინტერვალის სიგანე. ჩვენ ვთვლით, რომ ინტერვალის სიგანე მუდმივია. ხშირად სიკაშკაშის *სიმკვრივის* ფუნქციას აფასებენ არეში (გამოსახულებაზე) არსებული სიკაშკაშის თითოეული მნიშვნელობის საერთო რაოდენობის გამოთვლით $h[a]$ ჰისტოგრამის აგებისას. ამის შემდეგ შეიძლება ჰისტოგრამის ისეთნაირი **ნორმირება**, რომ მის ქვეშ არსებული ფართობის სიდიდე ტოლი იყოს 1-ის, ანუ აკმაყოფილებდეს (2.1) პირობას. ასეთ შემთხვევაში არის ნორმირებული სიკაშკაშის სიმკვრივის $p[a]$ ფუნქცია ტოლია არეში სიკაშკაშის a მნიშვნელობების მქონე პიქსელების საერთო Λ რაოდენობის ნორმირებული მნიშვნელობის:

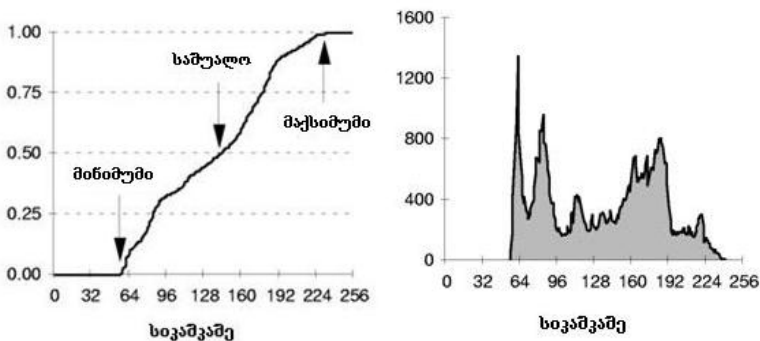
$$p[a] = \frac{1}{\Lambda} h[a], \text{ სადაც } \Lambda = \sum_a h[a]$$

სიკაშკაშის განაწილების ფუნქციას 2.10,ა სურათზე მოცემული გამოსახულებისთვის აქვს 2.10,ბ სურათზე წარმოდგენილი სახე. ხოლო 2.10,გ სურათზე წარმოდგენილია სიკაშკაშის არა-

ნორმირებული ჰისტოგრამა, რომელიც შეფასებული სიკაშკაშის სიმკვრივის ფუნქციის პროპორციულია. სიმაღლე ამ ჰისტოგრამაზე შეესაბამება გამოსახულებაზე სიკაშკაშის მოცემული მნიშვნელობის მქონე პიქსელების საერთო რაოდენობას. როგორც არისთვის გაზომილი განაწილების ფუნქცია, ასევე ჰისტოგრამა წარმოადგენს არის სტატისტიკურ მახასიათებლებს.



ა



ბ

ბ

სურ. 2.10

ამავე დროს, საჭიროა აღინიშნოს, რომ ნორმირებულ როგორც $P[a]$, ასევე $p[a]$ უნდა შევხედოთ როგორც ჭეშმარიტი

$(P(a)$ და $p(a)$) განაწილებების შეფასებებს, როცა ისინი გამოთვლილია სპეციფიკური არისთვის. შესაბამისად, გამოსახულების აღწერის ჩვენ მიერ წარმოდგენილ მოდელში როგორც გამოსახულება, ასევე სპეციფიკური არე განიხილება როგორც გამოსახულების და არის მაფორმირებელი სხვადასხვა შემთხვევითი პროცესების ერთი რეალიზაცია.

იგივე კონტექსტში განიხილება ქვემომოყვანილი სტატისტიკებიც. კერძოდ:

საშუალო:

უწყვეტი გამოსახულების R არის საშუალო სიკაშკაშე m_a განისაზღვრება როგორც არეში შემავალი ყველა (Λ) პიქსელის სიკაშკაშეების ამოკრებითი საშუალო

$$m_a = \frac{1}{\Lambda} \sum_{(i,j) \in R} a[i,j].$$

აღტერნატიულად, ციფრული გამოსახულების შემთხვევაში თუ გამოვიყენებთ არანორმირებულ სიკაშკაშის ჰისტოგრამას სიკაშკაშის a დისკრეტული მნიშვნელობებით, $h(a) = \Lambda \cdot p(a)$, მაშინ გვექნება

$$m_a = \frac{1}{\Lambda} \sum_a a \cdot h[a].$$

როგორც ვხედავთ, ციფრულ გამოსახულებაზე საშუალო სიკაშკაშე m_a წარმოადგენს უწყვეტი გამოსახულების სიკაშკაშის განაწილების ფუნქციის საშუალო μ_a სიკაშკაშის შეფასებას.

სტანდარტული გადახრა:

თუ R არე შეიცავს Λ რაოდენობის პიქსელებს, მაშინ სიკაშკაშის მნიშვნელობების s_a სტანდარტული გადახრის ჩაუნაცვლებელ შეფასებას ეწოდება **ამონაკრების სტანდარტული გადახრა** და იგი მოიცემა ასეთნაირად

$$\begin{aligned} s_a &= \sqrt{\frac{1}{\Lambda-1} \sum_{(i,j) \in R} (a[i,j] - m_a)^2} = \\ &= \sqrt{\frac{\sum_{(i,j) \in R} a^2[i,j] - \Lambda m_a^2}{\Lambda-1}}. \end{aligned}$$

თუ წარმოდგენილ გამოსახულებაში გამოვიყენებთ ჰისტოგრამის ზემომოყვანილ გამოსახულებას, მივიღებთ

$$s_a = \sqrt{\frac{\sum_a a^2 \cdot h[a] - \Lambda m_a^2}{\Lambda-1}}.$$

s_a სტანდარტული გადახრა წარმოადგენს უწყვეტ გამოსახულებაზე სიკაშკაშის განაწილების σ_a სტანდარტული გადახრის შეფასებას.

მოდა

განაწილების **მოდა** წარმოადგენს სიკაშკაშის მნიშვნელობათა შორის იმ მნიშვნელობას, რომელიც ყველაზე უფრო ხშირად გვხვდება გამოსახულებაზე.

2.4.2 დისკრეტული გამოსახულების აღწერა

იმის გამო, რომ გამოსახულება შესაძლებელია წარმოდგენილ იქნეს უწყვეტი ფუნქციის სახით, შესაბამისად კომპიუტერში

ფუნქციის წარმოდგენის ნებისმიერი ხერხი ზოგადად გამოსაღები იქნება ასევე გამოსახულების წარმოდგენისათვის. ზოგიერთი მიდგომა თავიდანვე შეიძლება უგულებელყოფილ იქნეს, მაგალითად, ინტენსივობის ფუნქციის წარმოდგენა მრავალწევრის სახით. ცნობილია, რომ ამ უკანასკნელის წარმოდგენა კომპიუტერში ხდება მისი კოეფიციენტების კრებულის სახით. მაგრამ ინტენსივობის ფუნქციების მხოლოდ მცირე რაოდენობა წარმოადგენს დაბალი ხარისხის მრავალწევრებს. საზოგადოდ, რეალურად გამოყენებულ ინტენსივობის ფუნქციებს რთული ანალიზური სახე აქვს. ამიტომ კომპიუტერში მათი წარმოდგენისათვის ხდება მათი მნიშვნელობების გაზომვა ანუ ანათვლების აღება (X, Y) სიბრტყის საკმაოდ დიდი რაოდენობის წერტილებში და შემდგომ შენახვა მეხსიერებაში. ამგვარად, ციფრული გამოსახულება კომპიუტერში წარმოდგინება დისკრეტულ მნიშვნელობათა ორგანზომილებიანი მასივის ან სხვა სიტყვებით, ცხრილის სახით. ამისათვის გამოსახულების უწყვეტი სივრცე და ინტენსივობის უწყვეტი დიაპაზონი იკვანტება დისკრეტული მნიშვნელობების კრებულად და კომპიუტერში გვაქვს ორგანზომილებიანი სტრუქტურის სახით წარმოდგენილი გამოსახულება.

განსაზღვრება 2.4. *ციფრული გამოსახულება – ეს არის ორგანზომილებიანი დისკრეტული მასივის სახით წარმოდგენილი $I[i, j]$, $i = \overline{1, N}$, $j = \overline{1, M}$ გამოსახულება, რომლის თითოეული $[i, j]$ ელემენტის მნიშვნელობა წარმოადგენს ინტენსივობის მნიშვნელობას აღებულს შეზღუდული სიზუსტით.*

როგორც უკვე აღვნიშნეთ, ორგანზომილებიანი გადამწოდის შესასვლელზე მიწოდებული უწყვეტი $f(x, y)$ გამოსახულება წარმოადგენს სხვადასხვა ცვლადზე დამოკიდებულ $2D$ ფი-

ზიკურ სიგნალს, მათ შორის ფერზე (λ), სიღრმეზე (z), დროზე (t). ჩვენ მიერ უმეტეს შემთხვევებში განხილული იქნება $2D$ მონოქრომატული, სტატიკური გამოსახულებები.

ჩვეულებრივად ყოველ პიქსელში ინტენსივობის მნიშვნელობად გამოყენებულია ერთბაიტიანი რიცხვები, ანუ გამოსახულებაზე გვექნება სიკაშკაშის მნიშვნელობები 0-დან 255-ის ჩათვლით, ხოლო ფერადი გამოსახულებების შემთხვევაში ყოველ პიქსელში ინტენსივობის წარმოსადგენად საჭიროა სამი მნიშვნელობა, თითო-თითო სამი ძირითადი ფერისთვის (RGB). ზოგჯერ სამედიცინო ამოცანებისთვის გამოიყენება 10 ბიტიანი წარმოდგენა, რომელიც ინტენსივობისთვის გვაძლევს ნახევარტონების 1024 მნიშვნელობას (0 – 1023).

არსებობს სხვადასხვა ტიპის გამოსახულებები:

- **ნახევარტონური** – ეს არის მონოქრომატული (შავ-თეთრი) ციფრული გამოსახულება $I[i, j]$, რომელშიც ყოველ პიქსელს შეესაბამება ინტენსივობის ერთი მნიშვნელობა.
- **მრავალსპექტრალური (მულტისპექტრალური)** – ეს არის ორგანზომილებიანი გამოსახულება $M[i, j]$, რომელშიც ყოველ სივრცით წერტილს შეესაბამება მნიშვნელობათა ვექტორი. ფერადი გამოსახულების შემთხვევაში კი ვექტორი სამგანზომილებიანია.
- **ბინარული** – ეს არის ციფრული გამოსახულება, რომლის პიქსელების მნიშვნელობებია 0 და 1.
- **მარკირებული** – ეს არის ციფრული გამოსახულება $L[i, j]$, რომელშიც პიქსელის მნიშვნელობად აიღება მნიშვნელობები სასრული ალფაბეტიდან. მოცემული პიქსელისათვის რაიმე სიმბოლოს მინიჭება ხდება მის მი-

მართ რაიმე გადაწყვეტილების მიღების შედეგად (მაგალითად, კლასიფიკაციის ან სეგმენტაციის).

ციფრული გამოსახულებების დამუშავებისა და ანალიზის სისტემის შესასვლელს შეიძლება წარმოდგენდეს არა მარტო პიქსელების მნიშვნელობების რიცხვითი მასივი, არამედ სხვა ტიპის რიცხვითი (ან ვექტორული) მასივები, მაგალითად, გამოსახულების ფილტრაციის ან სხვადასხვა გარდაქმნების შედეგები; ასევე “სიმბოლური” გამოსახულებები, სადაც პიქსელების მნიშვნელობებს წარმოდგენს არა რიცხვითი სიდიდეები, არამედ ჭდეები, მაგალითად, ზღურბლთან შედარების შედეგად მიღებული ბინარული გამოსახულებები, ბმული კომპონენტების მარკირების შედეგად მიღებული მასივები და ა.შ. გამოსახულების ანალიზისას მონაცემთა სტრუქტურის სახით გამოიყენება გეომეტრიული წარმოდგენები და გრაფები. ამ დროს გამოყოფილი ნიშან-თვისებები და არეები შეიძლება წარმოდგენილი იყოს როგორც ბინარული გამოსახულებების სახით, ასევე უფრო კომპაქტური სახითაც, მაგალითად, კოორდინატების კრებულით, საზღვრების კოდებით. მონაცემთა გეომეტრიული სტრუქტურების საშუალებით შესაძლებელია ამოსაცნობი სამგანზომილებიანი ობიექტების წარმოდგენაც.

ჩვეულებრივად რიცხვითი მასივის სახით წარმოდგენილ შესასვლელ მონაცემებზე სრულდება შემდეგი ტიპის ოპერაციები:

- **ოპერაციები წერტილებზე და ლოკალური ოპერაციები** (სიკაშკაშის გარდაქმნები და ა.შ.);
- **სტატისტიკური გამოთვლები** (ჰისტოგრამების გამოთვლა, ტექსტურების ანალიზი და ა.შ.);

- **გეომეტრიული ოპერაციები** (ცენტრალური პროექცია, არაწრფივი გარდაქმნები და ა.შ.);
- **გარდაქმნები** (ფურიეს, ჰაფის, ადამარის და ა.შ.);
- **სეგმენტაციის მეთოდების გამოყენებით გეომეტრიული ობიექტების გამოყოფა** (სასაზღვრო მრუდების ან არეების).

გეომეტრიული წარმოდგენების დონეზე შესაძლოა საჭირო გახდეს შემდეგი ტიპის ოპერაციები:

- **ბულის ოპერაციები** (გაერთიანებები, გადაკვეთები და დამატებები);
- **წარმოებული სიმრავლეების აგება** (ამოზნექილი გარსების, ვორონოის დიაგრამების, ჩონჩხური სიმრავლეების და ა.შ.);
- **ხილვადობის და მობილურობის გამოთვლები** (ტრაექტორიის დაგეგმვა და ა.შ.);
- **გეომეტრიული თვისებების გამოთვლები** (ბმულობის, მომენტების და ა.შ.).

გამოსახულების დამუშავებისას მასთან აკავშირებენ კოორდინატთა სისტემას. არსებობს სხვადასხვა ტიპის კოორდინატთა სისტემები:

- **რასტრულ კოორდინატთა სისტემა** – მასში გამოიყენება სტრიქონებისა და სვეტების კოორდინატები, ხოლო ანათვლების სათავე $[0,0]$ მოთავსებულია გამოსახულების ზედა მარცხენა კუთხეში.

- მართკუთხა კოორდინატთა სისტემა – სათავით [0,0] გამოსახულების ქვედა მარცხენა კუთხეში.
- მართკუთხა კოორდინატთა სისტემა – სათავით [0,0] გამოსახულების ცენტრში.

კოორდინატთა სისტემა გამოიყენება თითოეული პიქსელის ადგილმდებარეობის მისათითებლად, რაც აუცილებელია გამოსახულების კომპიუტერული დამუშავებისას.

2.5. ციფრული გამოსახულების ფორმატები

ციფრულ გამოსახულებებს ფართო გამოყენება აქვს ტელეკომუნიკაციურ სისტემებში, მონაცემთა ბაზებში და მანქანურ ხელვაში. ამის გამო სხვადასხვა ტიპის აპარატურულ და პროგრამულ საშუალებებს შორის მონაცემთა გაცვლის უზრუნველსაყოფად გამოსახულებებისათვის შემუშავდა სტანდარტული ფორმატები. მათი რაოდენობა რამდენიმე ათეულს აღწევს. ჩვენ მოკლედ განვიხილავთ მხოლოდ რამდენიმე მნიშვნელოვანს.

გამოსახულება შეიძლება წარმოვიდგინოთ როგორც ბაიტნაკადი, რომელიც გამოსახულების პიქსელებს წარმოადგენს რიგრიგობით, სტრიქონ-სტრიქონ. ბაიტების ასეთ რიგს ხშირად **რასტრულ რიგს** უწოდებენ (raster order). ბაიტნაკადში შეიძლება იმყოფებოდნენ სტრიქონების ერთმანეთისგან გამომყოფი მარკერები (ჭდეები). საწყისი გამოსახულების პიქსელური მონაცემები არ შეიცავს პარამეტრებს გამოსახულების ტიპის, განზომილების, დროის და მიღების ხერხის შესახებ. ასეთი სახის ინფორმაციას ხშირად წერენ ხელით ვიდეოკასეტის ეტიკეტზე ან ვთქვათ ქაღალდის ფურცელზე. შედარებით ახლახან შემუშავდა სტანდარტული ფორმატები, რომლებშიც გათვალისწინებულია მარკირებისათვის ან პიქსელური მონაცე-

მების დეკოდირებისთვის საჭირო არაგრაფიკული ინფორმაციის სათაურის შენახვა. ზოგიერთი კომპანიის მიერ, რომლებიც უშვებენ გრაფიკულ მოწყობილობებს და აპარატურას გამოსახულებების დამუშავებისთვის, შემუშავებულ იქნა რამდენიმე სტანდარტული ფორმატი, მაგრამ მხოლოდ მათი ნაწილისთვის არსებობს ღია დოკუმენტაცია.

გამოსახულების **ფაილის სათაური** (file header) შეიცავს ყველა იმ მახასიათებელს, რომელიც სრულიად საკმარისია იმისათვის, რომ მოხდეს მისი დამუშავება სხვადასხვა პროგრამების საშუალებით. სათაური შეიძლება შეიცავდეს მონაცემებს გამოსახულების ტიპისა და განზომილების, შექმნის დროის, სახელის და ა.შ. შესახებ.

ზოგიერთი სტანდარტული ფორმატი შექმნილია გამოსახულებათა მხოლოდ გარკვეული ტიპისთვის, მაგალითად, ბინარული და შავ-თეთრი. ფორმატები შეიძლება ერთმანეთისგან განსხვავდებოდეს გამოსახულების განზომილების დასაშვები ზღვრული მნიშვნელობებით. არსებობს ფორმატები, რომლებიც განკუთვნილია კადრების მიმდევრობასთან სამუშაოდ. სულ ახლახანს შემუშავებული მულტიმედია-ფორმატები შესაძლებლობას იძლევა, რომ გამოსახულების მონაცემებთან ერთად შეინახოს ასევე ტექსტური, აუდიო და გრაფიკული ინფორმაციაც. ასევე ზოგიერთ ფაილში გათვალისწინებულია გამოსახულების მონაცემების შეკუმშვის პროცედურა. ამასთან, უშუალოდ გამოსახულების ფაილში ყველა პიქსელის მნიშვნელობა არ იწერება. გამომდინარე იმ მოთხოვნიდან, თუ რა ხარისხის შეკუმშვაა საჭირო და რა მეთოდის გამოყენებით, შესაძლებელია გამოსახულების ფაილის ზომის შემცირება საწყისი გამოსახულების 30 პროცენტამდე. მონაცემთა შეკუმშვის მეთოდები იყოფა **შეკუმშვის მეთოდებად დანაკარგის გარეშე** (lossless compression) და **დანაკარგებით** (lossy compression). პირ-

ველ შემთხვევაში საწყისი გამოსახულება შეიძლება ზუსტად იქნეს აღდგენილი გამოსახულების ფაილის მონაცემების საფუძველზე. მეორე შემთხვევაში კი, საწყისი გამოსახულების აბსოლუტურად ზუსტი აღდგენა შეუძლებელია. ამასთან დანაკარგის შემჩნევა ხან შესაძლებელია და ხან არა.

ყველაზე უფრო ხშირად გამოყენებული ფორმატებიდან აღსანიშნავია შემდეგი:

- **PBM** (Portable Bit Map, გადასატანი რასტრული ფორმატი) ტიპის ფორმატების ოჯახში შედის ფორმატები: PBM, PGM და PPM. ისინი წარმოადგენენ გამოსახულებათა ფაილების უმარტივეს ფორმატებს. ამ ფორმატებში გამოსახულების სათაური და პიქსელური მონაცემები ინახება ASCII კოდში.
- **GIF** (Graphic Interchange Format, გრაფიკული მონაცემების ურთიერთგაცვლის ფორმატი). ეს ფორმატი გამოიყენება დიდი რაოდენობის მონაცემების შესანახად მსოფლიო ინტერნეტ ქსელში და არსებულ მონაცემთა ბაზებში. ის მარტივი სტრუქტურისაა, თუმცა ფერის კოდირებისთვის მასში გამოიყენება მხოლოდ 8 ბიტის სიტყვები. ამის გამო იგი არ არის განკუთვნილი იმ შემთხვევისათვის, როცა გამოსახულება შეიცავს ფერთა დიდ რაოდენობას ანუ 256 განსხვავებულ ფერზე მეტს.
- **TIFF** (Tag Image File Format, გამოსახულებათა ფაილების ტეგური ფორმატი). ის რთული აგებულებისაა, თუმცა საკმაოდ მოქნილია. ის გამოიყენება თითქმის ყველა ტიპის კომპიუტერებში და სკანერებისთვის წარმოადგენს ძირითად ფორმატს. TIFF საშუალებას იძლევა, რომ ერთ ფაილში შევინახოთ რამდენიმე გამოსახულება ფერების გარჩევისუნარიანობით 1-დან 24 ბიტამდე/პიქსელზე.

მასში გათვალისწინებულია მონაცემთა შეკუმშვა დანაკარგებით და დანაკარგების გარეშე.

- **JPEG (JFIF / JFI / JPG)** (Joint Photographic Experts Group). ის შედარებით ახალშექმნილი ფორმატია, რომლის შემუშავებისას ავტორები მიზნად ისახავდნენ შეექმნათ უძრავი მაღალხარისხოვანი ფერადი გამოსახულებების შეკუმშვის ხერხი. ამ ფორმატში გამოყენებულ კოდირების მეთოდს შეუძლია მონაცემთა ნაკადური დამუშავება და შესაძლებლობას იძლევა, რომ მის საფუძველზე შემუშავდეს რეალურ დროში გამოსახულებათა შეკუმშვისა და გახსნის აპარატურული მოწყობილობები. ამ ფორმატში გამოსახულების განზომილებამ შეიძლება მიაღწიოს ($64K \times 64K$) პიქსელამდე ფერების გარჩევისუნარიანობით 24 ბიტი/პიქსელზე. თითოეულ ფაილში შესაძლებელია მხოლოდ ერთი გამოსახულების შენახვა, ხოლო გამოსახულების სათაურში კი მისი მინიატურული ვარიანტი $64K$ -მდე არაშეკუმშული ბაიტის სახით. ამ ფორმატის მნიშვნელოვან უპირატესობას წარმოადგენს მისი დამოუკიდებლობა ფერების წარმოდგენის სისტემისაგან. მონაცემთა შეკუმშვის მაღალი ხარისხის მისაღწევად გამოიყენება მოქნილი, მაგრამ ამავე დროს რთული ალგორითმი დანაკარგებით. ხშირად მისი საშუალებით შესაძლებელია მაღალი ხარისხის გამოსახულების შეკუმშვა ფარდობით 20:1 ხარისხის შეუმჩნეველი დანაკარგის გარეშე.
- **PostScript (BDF / PDL / EPS)**. ამ ფორმატების ოჯახი საშუალებას იძლევა შევინახოთ გამოსახულებათა მონაცემები ASCII – ის ცხრილის საბუჯდი სიმბოლოების გამოყენებით. ეს ფორმატები გამოიყენება გამოსახულებათა ეკრანული ვიზუალიზაციისათვის. ენები PDL

(გვერდების აღწერის ენა) და EPS (ინკაფსულირებული PostScript) გამოიყენება გრაფიკისა და გამოსახულებების შესანახად მათი სხვა დოკუმენტებში შემდგომი ჩასმის მიზნით. გამოსახულების პიქსელების კოდირება ხდება ASCII – ის 7 ბიტისანი კოდებით და აქედან გამომდინარე შესაძლებელია მათი რედაქტირება და ნახვა ტექსტური რედაქტორის საშუალებით. ფორმატი PostScript განკუთვნილია ნახევარტონური და ფერადი გამოსახულებების შესანახად გარჩევისუნარიანობით 75-3000 წერტილი/დუიმი-მზე.

- **MPEG (MPG / MPEG-1 / MPEG-2).** ეს ფორმატები წარმოადგენენ ვიდეო-აუდიო მონაცემების, ტექსტისა და გრაფიკის შეკუმშვისათვის განკუთვნილ ნაკადური ტიპის მეთოდებს. აბრევიატურა MPEG იშიფრება ასეთნაირად – Motion Picture Experts Group (მოძრავი გამოსახულებების დამუშავების ექსპერტული ჯგუფი). ფორმატი MPEG-1 განკუთვნილია მულტიმედია კომპიუტერებისთვის და უზრუნველყოფს შეკუმშული აუდიომონაცემების გადაცემას სიჩქარით 0,25 მგბიტი/წმ და შეკუმშული ვიდეომონაცემების გადაცემას სიჩქარით 1,25 მგბიტი/წმ. MPEG-2 ფორმატით კი გადაცემის სიჩქარე არის 15 მგბიტი/წმ, რაც უზრუნველყოფს ამ მონაცემებისთვის სატელევიზიო ხარისხს, განსხვავებით MPEG-1 ფორმატისაგან.

III თავი. გამოსახულების დისკრეტიზაცია და კვანტირება

3.1. სივრცითი გაზომვები გამოსახულებაზე

როგორც უკვე ზემოთ აღვნიშნეთ, ციფრული გამოსახულების ყოველ პიქსელში ჩვენ გვაქვს უწყვეტი გამოსახულების რაღაც ელემენტარულ არეში (მიდამოში) ინტენსიობის გაზომვის შედეგად მიღებული სიდიდე, ანათვალი. თუ ახლა მოვახდენთ გამოსახულების სიბრტყის რომელიმე პიქსელის უკუპროექტირებას დაკვირვებული სცენის ობიექტის ზედაპირზე, მაშინ სცენის შესაბამისი ელემენტის ზომა (პროექციის ზომა) შესაძლებლობას გვაძლევს შევაფასოთ გადამწოდის ნომინალური გარჩევისუნარიანობა (nominal resolution).

განსაზღვრება 3.1. მკვ გადამწოდის ნომინალური გარჩევისუნარიანობა ტოლია სცენის იმ ზედაპირული ელემენტის გვერდის განზომილებისა, რომელსაც გამოსახულების სიბრტყეში ცალკე პიქსელი შეესაბამება.

მაგალითად, თუ კვადრატული ფურცელი გვერდის ზომით 25სმ. წარმოდგენილია (500 × 500) პიქსელი განზომილების ციფრული გამოსახულებით, მაშინ გადამწოდის ნომინალური გარჩევისუნარიანობა უდრის 0,05სმ. აღსანიშნავია, რომ ნომინალური გარჩევისუნარიანობის ცნების გამოყენება არ არის სასურველი იმ შემთხვევაში, როცა სცენის ობიექტებს შორის დაშორება სიღრმის (დამკვირვებლიდან სიშორის) თვალსაზრისით დიდია. ამის მიზეზი ის არის, რომ ამ დროს ნომინალური

გარჩევისუნარიანობა იცვლება ობიექტების ზედაპირების სიშორისა და ორიენტაციისგან დამოკიდებულებით.

გადამწოდის მიერ სცენის დაკვირვებული სივრცის მოცულობის მახასიათებელს წარმოადგენს გამოსახულების გადამწოდის **ხედვის ველი** (field of view).

განსაზღვრება 3.2. *გამოსახულების გადამწოდის ხედვის ველი განისაზღვრება დაკვირვებული სცენის იმ ზომით, რომლის აღქმაც მას შეუძლია (მაგალითად, 20სმ×20სმ). იმის გამო, რომ ეს ზომა შეიძლება იცვლებოდეს სცენის სიღრმის ცვლილებასთან ერთად, უფრო აზრიანია ცნების - ხედვის კუთხური ველის (angular field of view) – გამოყენება (მაგალითად, $30^{\circ} \times 27^{\circ}$).*

გადამწოდის **გარჩევისუნარიანობა** ახასიათებს მის სიზუსტეს სიცრცითი გაზომვების ან მახასიათებელი ნიშან-თვისებების აღმოჩენისას.

განსაზღვრება 3.3. *ტერმინი გარჩევისუნარიანობა ეხება გადამწოდის სიზუსტეს გაზომვების შესრულების დროს.*

ფორმალური თვალსაზრისით გადამწოდის გარჩევისუნარიანობა შეიძლება განისაზღვროს რამდენიმე ხერხით. მაგალითად, რეალური სამყაროს ტერმინებში ნომინალური გარჩევისუნარიანობის საშუალებით, ვთქვათ ასეთნაირად – “მოცემული სკანერის გარჩევისუნარიანობა დედამიწის ზედაპირზე ტოლია ერთი მეტრისა” ან “ერთ მილიმეტრზე მონაკვეთთა იმ წყვილების რაოდენობით, რომელთა გარჩევაც შეიძლება მიღებულ გამოსახულებაზე”. მაგრამ სხვა შინაარსი დევს ტერმინ გარჩევისუნარიანობაში, როცა ვლავარაკობთ იმ პიქსელების რაოდენ-

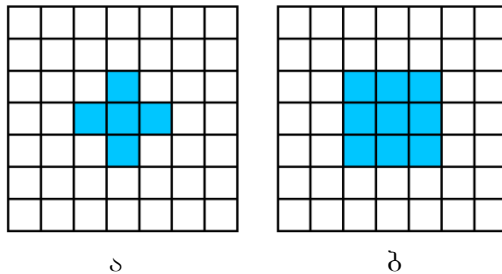
ნობაზე, რომლებზედაც კამერას ხელი მიუწვევება სცენაზე დაკვირვებისას, ანუ კამერის ხედვის ველში მოხვედრილი პიქსელების რაოდენობაზე. მაგალითად, თუ გადამწოდის საშუალებით მივიღეთ გამოსახულება გარჩევისუნარიანობით (640×480) პიქსელი, მაშინ ხედვის ველიც ამდენ ნაწილად დაიყო. სხვა სიტყვებით, ის გვიჩვენებს, თუ რამდენ ნაწილად დაიყო გამოსახულების გადამწოდის ხედვის ველი. ამას დიდი მნიშვნელობა აქვს როგორც ზუსტი გაზომვების ჩასატარებლად, ასევე სცენის მოცემულ არეზე დაკვირვებისთვის. ადამიანს შეუძლია გაარჩიოს ნაცნობი ადამიანის სახე გამოსახულებაზე გარჩევისუნარიანობით (64×64), შესაძლოა ასევე (32×32), მაგრამ უკვე გარჩევისუნარიანობა (16×16) ამისთვის საკმარისი აღარ არის. სცენისათვის გარჩევისუნარიანობის შერჩევას დიდი მნიშვნელობა ენიჭება. ძალიან დაბალი გარჩევისუნარიანობა ცუდ შედეგს იძლევა ობიექტების ამოცნობისას ანუ გამოსახულებაზე გაზომვების ჩატარებისას, და პირიქით, ძალიან მაღალს შეუძლია გააძნელოს ალგორითმების შესრულება და ამით გამოიწვიოს მესხიერების დიდი დანახარჯი.

3.2. გამოსახულების დისკრეტიზაცია

გამოსახულებათა დამუშავების ციფრულ სისტემებში უწყვეტი გამოსახულება შედის რიცხვების მასივის სახით, რომელიც მიიღება უწყვეტი გამოსახულების დისკრეტიზაციისა და კვანტირების შედეგად. დისკრეტიზაციის ქვეშ იგულისხმება უწყვეტი გამოსახულების აღმწერი ფუნქციის შეცვლა მისი ანათვლებით, ანუ მნიშვნელობებით არგუმენტის მნიშვნელობათა სასრული სიმრავლიდან. კვანტირება კი გულისხმობს იმას, რომ ყოველ წერტილში ფუნქციის მნიშვნელობები იცვლება მათთან უახლოესი მნიშვნელობებით, რომლებიც აირჩევა მნიშვნელო-

ბათა სასრული სიმრავლიდან (კვანტირების დონეებიდან). რეალურად ეს იმას ნიშნავს, რომ სინათლის უწყვეტი სპექტრი იყოფა სასრული რაოდენობის კვანტირების დონეებად. ანათელების მდებარეობა და კვანტირების დონეები შეიძლება წინასწარ იყოს მოცემული ან არჩეულ იქნეს ფუნქციის ხასიათიდან გამომდინარე. მთავარი მოთხოვნა, რომელიც წაყენება დისკრეტიზაციისა და კვანტირების პროცესს, არის ის, რომ მიღებული ციფრული გამოსახულება კარგად აღსაქმელი უნდა იყოს დამკვირვებლის მიერ.

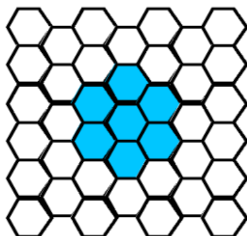
ჩვეულებრივად ანათელები აიღება ერთმანეთისგან თანაბრად დაშორებული ინტერვალებით ისეთნაირად, რომ მათ შექმნან რეგულარული ბადე. ამისათვის უწყვეტი გამოსახულების სიბრტყეზე დებენ დისკრეტიზაციის ბადეს (ხშირად გამოიყენება კვადრატული ბადე) და იღებენ სიკაშკაშის ფუნქციის ან ფერის ანათვალს ბადის ყოველი უჯრედის ცენტრში ან მთელი უჯრედის გასაშუალებულ მნიშვნელობას (სურ. 3.1).



სურ. 3.1

თუმცა ზოგჯერ ანათელებს შორის არათანაბარ ინტერვალებსაც ირჩევენ. დისკრეტიზაციისთვის ასევე გამოიყენება ექვსკუთხა (ჰექსაგონალური) ბადე (სურ. 3.2)

ამავე დროს მნიშვნელოვანია იმის გაგება, თუ როგორ ხდება უწყვეტი გამოსახულების დისკრეტიზაცია ანუ ანათვლების აღება და როგორ განსაზღვრავს ის სხვადასხვა სახის მეზობლობას, რომელიც თავის მხრივ მნიშვნელოვანია გამოსახულების დამუშავებისას.



სურ. 3.2

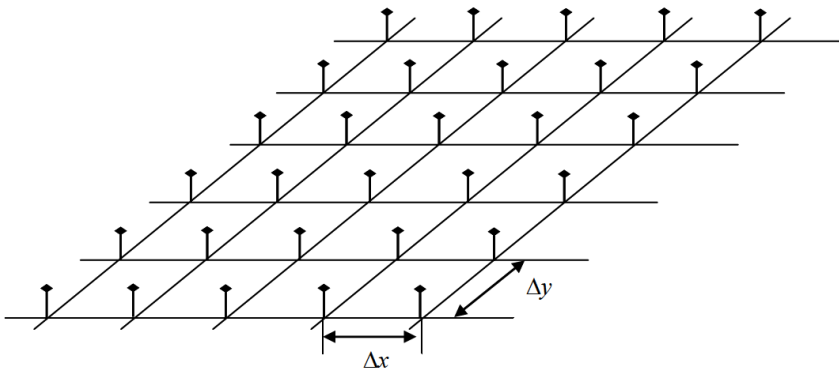
საზოგადოდ, ყველაზე უფრო ხშირად ანათვლების მართკუთხა ბადით აღებისას გამოიყენება ორი ტიპის მეზობლობა, 4-ბმული და 8-ბმული, ხოლო ექვსკუთხა ბადის შემთხვევაში კი 6-ბმული. ამიტომ ციფრული გამოსახულების დამუშავებისას მოცემული პიქსელის მიდამო, ანუ მეზობლობა განისაზღვრება იმისდა მიხედვით, თუ რამდენ ბმულიან მეზობლობას ავირჩევთ.

დისკრეტიზაციის რეალურ სისტემებში დისკრეტიზაციის ბადეს აქვს სასრული ზომები და ანათვლები აიღება რეალური გამოსახულების გარკვეული ფიზიკური მახასიათებლების, მაგალითად, სიკაშკაშის გაზომვის შედეგად. იმის გამო, რომ გამზომი ხელსაწყოებისთვის დამახასიათებელია ცდომილების გარკვეული სიდიდე, დისკრეტიზაციის შედეგად მიღებული გამოსახულება არ წარმოადგენს იდეალურ ციფრულ გამოსახულებას. თუ ანათვლების წერტილები ქმნიან $n \times n$ ელემენტებისგან შედგენილ კვადრატულ ბადეს, მაშინ უწყვეტი გამოსა-

ხულებიდან ასეთი გზით მიღებული გამოსახულება წარმოადგენს $n \times n$ განზომილების დისკრეტულ გამოსახულებას, რომელიც განისაზღვრება ანათვლების წერტილებში მისი მნიშვნელობების მატრიცით.

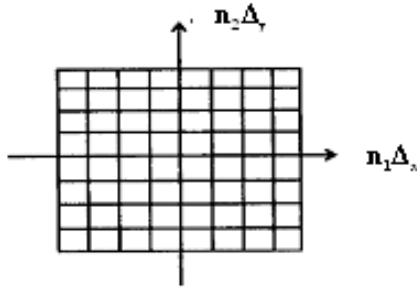
ვთქვათ, უწყვეტი გამოსახულება შემოსაზღვრულ მართკუთხა $-L_1 \leq x \leq L_1, -L_2 \leq y \leq L_2$ არეში წარმოდგენილია ორი სივრცითი x და y ცვლადის $f(x, y)$ უწყვეტი ფუნქციით, რომელიც რეალურ გამოსახულებაზე აღწერს სიკაშკაშის, ოპტიკური სიმკვრივის ან სხვა რომელიმე პარამეტრის განაწილებას.

აღვნიშნოთ Δ_x -ით დისკრეტიზაციის ბიჯი სივრცითი x ცვლადისთვის და Δ_y -ით სივრცითი y ცვლადისთვის. შედეგად ვღებულობთ მართკუთხა ბადეს, რომლის კვანძებში განვალაგოთ წერტილოვანი ვიდეოგადაამწოდები (სურ. 3.3).



სურ. 3.3

ასეთ შემთხვევაში $f(x, y)$ გამოსახულება მოცემული აღმოჩნდება $2D$ განზომილების მართკუთხა გისოსზე (სურ. 3.4),



სურ. 3.4

რაც ასეთნაირად ჩაიწერება

$$f(n_1 \Delta_x, n_2 \Delta_y) = f(x, y)|_{x=n_1 \Delta_x, y=n_2 \Delta_y}.$$

ამ ფორმულის მოკლედ ჩაწერისათვის შემოვიტანოთ აღნიშვნა:

$$f(n_1 \Delta_x, n_2 \Delta_y) \equiv f(n_1, n_2).$$

ფუნქცია $f(n_1, n_2)$ წარმოადგენს ორი დისკრეტული ცვლადის ფუნქციას და მას ეწოდება **ორგანზომილებიანი მიმდევრობა**. ამგვარად, უწყვეტი $f(x, y)$ გამოსახულება, რომელიც წარმოადგენს ორი სივრცითი ცვლადის უწყვეტ ფუნქციას, დისკრეტიზაციის შედეგად გარდაიქმნება ამოკრეფითი მნიშვნელობების (ანათვლების) ცხრილად. ამ ცხრილის განზომილება (სტრიქონებისა და სვეტების რაოდენობა) განისაზღვრება საწყისი შემოსაზღვრული მართკუთხა არის $(-L_1 \leq x \leq L_1, -L_2 \leq y \leq L_2)$ გეომეტრიული ზომებით და დისკრეტიზაციის Δ_x და Δ_y ბიჯის არჩევით შემდეგი ფორმულების საფუძველზე

$$N_1 = \lceil 2L_1/\Delta_x \rceil, \quad N_2 = \lceil 2L_2/\Delta_y \rceil,$$

სადაც [...] -ით აღნიშნულია რიცხვის მთელი ნაწილი.

იმ შემთხვევაში, თუ უწყვეტი გამოსახულების განსაზღვრის არე წარმოადგენს კვადრატს, ანუ $L_1 = L_2 = L$, ხოლო დისკრეტიზაციის x და y ბიჯები ღერძებზე ერთმანეთის ტოლია, ანუ $\Delta_x = \Delta_y = \Delta$, მაშინ

$$N_1 = N_2 = N,$$

და ცხრილის განზომილება ტოლია N^2 .

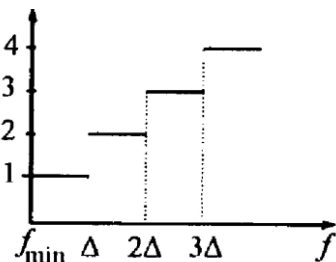
უწყვეტი გამოსახულების დისკრეტიზაციის შედეგად მიღებული ცხრილის ელემენტი წარმოადგენს სწორედ პიქსელს ანუ ანათვალს. განვიხილოთ ახლა ამ ცხრილის $f(n_1, n_2)$ ელემენტი (პიქსელი), რომელიც წარმოადგენს ნამდვილ რიცხვს.

კომპიუტერი ოპერირებს მხოლოდ დისკრეტულ სახეში წარმოდგენილ რიცხვებთან. ამიტომ მის მეხსიერებაში უწყვეტი f სიდიდის ჩასაწერად საჭიროა ამ უკანასკნელზე მოვახდინოთ ანალოგურ-ციფრული გარდაქმნა ბიჯით Δf . ამ ოპერაციას (უწყვეტი სიდიდის დისკრეტიზაციას დონის მიხედვით) უწოდებენ **კვანტირებას** (სურ. 3.5).

თუ სიკაშკაშის ფუნქციის მნიშვნელობები მოთავსებულია ინტერვალში $[f_{min}, f_{min} + c]$, მაშინ კვანტირების დონეების რიცხვი უდრის

$$Q = \lceil c/\Delta f \rceil.$$

ამგვარად, კვანტირებისას გაზომვის შედეგად მიღებული ანათელებების ნამდვილი მნიშვნელობები გარდაიქმნება მთელ რიცხვებად ანუ ბიტების კრებულად.



სურ. 3.5

ზემოთქმულის საფუძველზე შეიძლება დავასკვნათ, რომ კომპიუტერის მესხიერებაში შენახული გამოსახულების პიქსელები წარმოადგენს საწყისი უწყვეტი გამოსახულების არგუმენტების და ღონეების მიხედვით დისკრეტიზაციის შედეგს. ცხადია იმისთვის, რომ დისკრეტიზაციის შედეგად მიღებულ ციფრულ გამოსახულებაში შენარჩუნებულ იქნეს უწყვეტი გამოსახულებისთვის დამახასიათებელი ძირითადი ინფორმაცია, ანუ უმნიშვნელო დისკრეტიზაციის ცდომილების მისაღებად დისკრეტიზაციის ბიჯები Δ_x და Δ_y უნდა იყოს საკმაოდ მცირე სიდიდის. ამავე დროს, უნდა გავითვალისწინოთ ისიც, რომ რაც უფრო მცირეა დისკრეტიზაციის და კვანტირების ბიჯები, მით დიდია კომპიუტერში გამოსახულების წარმომდგენი რიცხვითი მასივის მოცულობა და, შესაბამისად, მით უკეთესია დისკრეტიზებული გამოსახულების ხარისხი.

იმისთვის, რომ მივიღოთ მაღალი ხარისხის კვანტირებული გამოსახულება, არ არის აუცილებელი კვანტირების ღონეების რაოდენობის წინასწარ შერჩევა. საზოგადოდ ეს

რაოდენობა შეიძლება განსაზღვრულ იქნეს გამომდინარე თვით გამოსახულების ხასიათიდან. კერძოდ, საჭიროა გამოსახულების ყოველი წერტილის გამოკვლევა კონტურის (გამოსახულებაზე სიკაშკაშის მკვეთრი ცვლილების ადგილების) არსებობის თვალსაზრისით. იმ უბნებისთვის, სადაც სიკაშკაშე მდოვრედ იცვლება, საჭიროა კვანტირების დონეების მაღალი რიცხვი, ანუ უნდა მოხდეს უფრო ზუსტი კვანტირება. წინააღმდეგ შემთხვევაში გამოსახულებაზე შეიძლება მივიღოთ მცდარი კონტურები, ხოლო იმ უბნებისთვის, სადაც ადგილი აქვს სიკაშკაშეების მკვეთრ ცვლილებებს, საჭიროა უხეში კვანტირება. კვანტირების ყველაზე უფრო მარტივ შემთხვევას წარმოადგენს თანაბარბიჯიანი კვანტირების გამოყენება. დისკრეტიზაციის დროს კი პირიქითაა, ანათვლები უნდა იყოს განლაგებული ხშირად, ანუ ერთმანეთთან ახლოს კონტურების სიახლოვეს, რათა შენარჩუნებულ იქნეს მათი მკაფიოობა, ხოლო იქ სადაც სიკაშკაშეები მდოვრედ იცვლება – იშვიათად.

აღსანიშნავია, რომ გამოსახულების დისკრეტიზაციის, ჩაწერის და აღდგენის პროცესში გამოსახულების ხარისხი ფუჭდება. ამის მიზეზი შესაძლოა იყოს: ოპტიკური სისტემის დიფრაქციული დამახინჯებები და აბერაცია, ფოტომასალების არაწრფივი დამახინჯებები, ატმოსფერის ტურბულენტობით გამოწვეული დამახინჯებები, გეომეტრიული დამახინჯებები და ა.შ. ცხადია, ასეთ გამოსახულებებზე გაძნელებულია ნიშანთვისებებისა და ერთგვაროვანი ბმული არეების (სეგმენტების) გამოყოფა, რაც ასე აუცილებელია სცენათა ანალიზის ამოცანის გადასაწყვეტად.

IV თავი. ბინარული გამოსახულებების ანალიზი

4.1. პიქსელი და მისი გარემო

ზოგიერთ გამოყენებით ამოცანებში, მაგალითად დოკუმენტების დამუშავებისას და სამრეწველო მანქანური ხედვის ზოგიერთ ამოცანებში, ბინარული (ორ დონიანი) გამოსახულებები გამოიყენება როგორც შესასვლელი მონაცემები გამოსახულებათა დამუშავების ალგორითმებისთვის. ნახევარტონური ან ფერადი გამოსახულებიდან ბინარული გამოსახულების ფორმირებისათვის შეიძლება განხორციელდეს ოპერაცია, რომელიც გამოსახულებიდან აირჩევს პიქსელთა გარკვეულ ქვესიმრავლეს როგორც წინა პლანის პიქსელებს (foreground). დანარჩენი პიქსელები კი განიხილება როგორც ფონური პიქსელები (background). პიქსელების შერჩევის ოპერაცია შეიძლება იყოს მარტივი, მაგალითად ზღურბლური ოპერატორი, რომელიც პიქსელების მნიშვნელობას ირჩევს სიკაშკაშის მოცემული დიაპაზონიდან ან წარმოადგენს კლასიფიკაციის რომელიმე ალგორითმს.

ბინარული B გამოსახულების პიქსელები იღებს მნიშვნელობებს 0 ან 1. დაშვებულია, რომ მნიშვნელობა 1 შეესაბამება წინა პლანის პიქსელებს, ხოლო 0 – ფონურ პიქსელებს. პიქსელის მნიშვნელობა i სტრიქონისა და j სვეტის გადაკვეთაში აღვნიშნოთ $B[i, j]$ -თი. გამოსახულება განზომილებით $N \times M$ შედგება N სტრიქონისგან, რომელიც გადანომრილია 0-დან $N - 1$ -ის ჩათვლით, და M სვეტისგან, გადანომრილი 0-დან $M - 1$ -ის ჩათვლით. ამგვარად, ბინარული B გამოსახულების ელემენტი $B[0,0]$ შეესაბამება გამოსახულების მარცხენა ზედა

პიქსელის მნიშვნელობას, ხოლო ელემენტი $B[N-1, M-1]$ კი – მარჯვენა ქვედა პიქსელის მნიშვნელობას.

გამოსახულების დამუშავების ალგორითმებს საქმე აქვს არა მარტო პიქსელების მნიშვნელობებთან, არამედ მათი მეზობელი პიქსელების მნიშვნელობებთანაც. როგორც უკვე ზემოთ აღვნიშნეთ, **პიქსელის მიდამოს**, ანუ **მეზობლობის** განხილვისას გამოიყენება მეზობლობის ორი განსაზღვრება: 4-ბმული და 8-ბმული მიდამო. $I[i, j]$ ციფრული გამოსახულების $[i, j]$, $i = \overline{0, N-1}$, $j = \overline{0, M-1}$ პიქსელის 4-ბმული მიდამო – $N_4[i, j]$, შედგება შემდეგი პიქსელებისაგან: $[i-1, j]$, $[i+1, j]$, $[i, j-1]$, $[i, j+1]$, ხოლო 8-ბმული მიდამო – $N_8[i, j]$ კი შეიცავს 4-ბმული მიდამოს ყველა პიქსელს და ასევე დიაგონალურად განლაგებულ მეზობელ პიქსელებსაც: $[i-1, j-1]$, $[i-1, j+1]$, $[i+1, j-1]$, $[i+1, j+1]$. გრაფიკულად $[i, j]$ პიქსელის ეს ორი მიდამო წარმოდგენილია 4.1 სურათზე, სადაც (ა) 4-ბმული მიდამოა, (ბ) 8-ბმული მიდამო.

	ზ	
მც	*	მჯ
	ქ	

ა

მცზ	ზ	მჯზ
მც	*	მჯ
მცქ	ქ	მჯქ

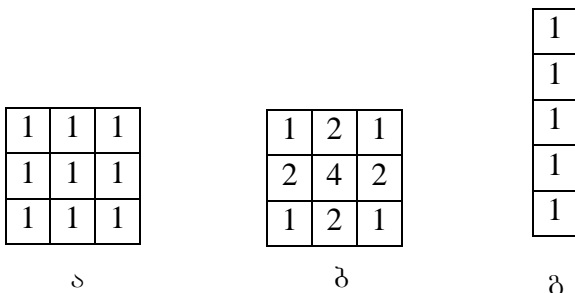
ბ

სურ. 4.1

4.2. ტრაფარეტების გამოყენება გამოსახულებების დამუშავებისას

გამოსახულებათა დამუშავების ამოცანებში ტრაფარეტის ცნება მნიშვნელოვან როლს ასრულებს. ტრაფარეტი ეს არის პიქ-

სელთა კოორდინატების სიმრავლე, რომელთაც მინიჭებული აქვს წონებად წოდებული მნიშვნელობები. 4.2 სურათზე ნაჩვენებია სამი ერთმანეთისგან განსხვავებული ტრაფარეტი.



სურ. 4.2

პირველი ორი – (ა) და (ბ) წარმოადგენს კვადრატული ტიპის ტრაფარეტებს. პირველის წონების მნიშვნელობები ტოლია 1-ის, ხოლო მეორესი კი ერთმანეთისგან განსხვავებულეებია. მესამეს კი აქვს მართკუთხედის ფორმა და ერთნაირი წონები. ტრაფარეტს აქვს განსაზღვრული კოორდინატთა სათავე. ჩვეულებრივად ის ემთხვევა ერთ-ერთი პიქსელის კოორდინატებს. თუ ტრაფარეტი სიმეტრიულია, მაშინ კოორდინატთა სათავედ აიღება მისი ცენტრალური პიქსელი (მაგალითად, (ა) და (ბ) შემთხვევა). ხოლო არასიმეტრიული ტრაფარეტის შემთხვევაში კი – ნებისმიერი პიქსელი. მაგალითად, (გ) შემთხვევაში სათავედ შეიძლება არჩეულ იქნეს ზედა პიქსელი. 4.3 სურათზე ნაჩვენებია წონითი კოეფიციენტებიანი ტრაფარეტის გამოყენების შედეგები ნახევარტონურ გამოსახულებაზე, სადაც (ა) საწყისი ნახევარტონური გამოსახულებაა, (ბ) ტრაფარეტი განზომილებით (3 × 3), (გ) ტრაფარეტის გამოყენების შედეგი და (დ) შედეგის ნორმირება ტრაფარეტის წონების ჯამზე გაყოფით.

40	40	80	80	80
40	40	80	80	80
40	40	80	80	80
40	40	80	80	80
40	40	80	80	80

ა

1	2	1
2	4	2
1	2	1

ბ

640	800	1120	1280	1280
640	800	1120	1280	1280
640	800	1120	1280	1280
640	800	1120	1280	1280
640	800	1120	1280	1280

გ

40	50	70	80	80
40	50	70	80	80
40	50	70	80	80
40	50	70	80	80
40	50	70	80	80

დ

სურ. 4.3

შესასვლელ გამოსახულებაზე ტრაფარეტის მოქმედების შედეგად ფორმირდება გამოსასვლელი გამოსახულება. ამ გამოსახულებათა ზომები ერთნაირია. ტრაფარეტის გამოყენების მექანიზმი შემდეგში მდგომარეობს. ჯერ ტრაფარეტი უნდა შეუთავსდეს შესასვლელ გამოსახულებას ისეთნაირად, რომ ტრაფარეტის ცენტრი უნდა ემთხვეოდეს დამუშავებაში მყოფ მიმდინარე პიქსელს. ამის შემდეგ ტრაფარეტის ყოველი პიქსელის წონა მრავლდება შესასვლელი გამოსახულების შესაბამისი პიქსელის მნიშვნელობაზე. შემდეგ ეს ნამრავლები ჯამდება და მიღებული მნიშვნელობა თავსდება შესასვლელი გამოსახულების დამუშავებაში მყოფი მიმდინარე პიქსელის ადგილას.

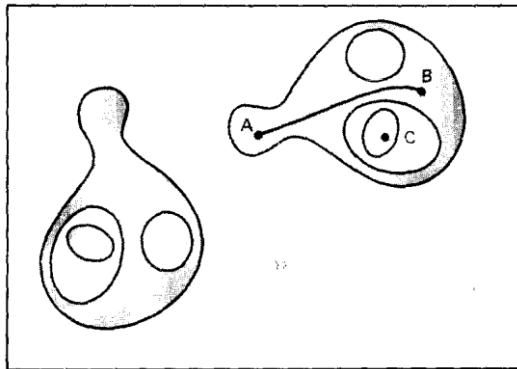
იმ შემთხვევაში, როცა ტრაფარეტის ცენტრი ემთხვევა სასახლვრო სვეტებში ან სტრიქონებში მოთავსებულ პიქსელებს, ტრაფარეტის ზოგიერთი პიქსელი გამოდის გამოსახულების საზღვრებს გარეთ. მაგრამ იმის გამო, რომ შესასვლელი და გამოსასვლელი გამოსახულებების ზომები ერთნაირი უნდა იყოს, ამიტომ შესაძლებელია, რომ შესასვლელ გამოსახულებას პერიმეტრის გასწვრივ დაემატოს რამდენიმე წარმოსახვითი სტრიქონი და სვეტი, რომლებშიც პიქსელების მნიშვნელობები ტოლი იქნება 0-ის ან რაიმე მუდმივასი. პრაქტიკაში ხშირად პიქსელების მნიშვნელობებად იღებენ მნიშვნელობებს უახლოესი სტრიქონიდან (ან სვეტიდან). 4.3 სურათზე (ბ) ნიღბის გამოყენების შედეგად მიღებული (გ) გამოსახულება წარმოადგენს შესასვლელი (ა) გამოსახულების დაგლუვებულ გამოსახულებას.

4.3. ბმული კომპონენტები და მათი მარკირება

ვთქვათ, მოცემული გვაქვს ბინარული გამოსახულება B და $B[i, j] = B[i', j'] = v$, სადაც $v = 0$ ან $v = 1$. მაშინ,

განსაზღვრება 4.1. *პიქსელს $[i, j]$ ეწოდება ბმული პიქსელთან $[i', j']$ მნიშვნელობით v , თუ არსებობს პიქსელების ისეთი მიმდევრობა $[i, j] = [i_0, j_0], [i_1, j_1], \dots, [i_k, j_k] = [i', j']$, რომ $B[i_k, j_k] = v$, $k = \overline{0, l}$ და $[i_k, j_k]$ წარმოადგენს $[i_{k-1}, j_{k-1}]$ -ის მეზობელს ნებისმიერი $k = \overline{1, l}$ -თვის. ხოლო პიქსელების მიმდევრობა $[i_0, j_0], [i_1, j_1], \dots, [i_k, j_k]$ ჰქმნის ბმულ გზას $[i, j]$ პიქსელიდან $[i', j']$ პიქსელისკენ.*

ბმული კომპონენტი (connected component) მნიშვნელობით ν – ეს არის ისეთი C პიქსელების მაქსიმალური სიმრავლე, რომელთა მნიშვნელობები ტოლია ν -სი და ამ პიქსელების ყოველი წყვილი ბმულია მნიშვნელობით ν . სხვა სიტყვებით, ბმული კომპონენტი ეს არის ბმული წერტილების ისეთი მაქსიმალური სიმრავლე, რომლის ნებისმიერ ორ წერტილს შორის არსებობს ბმული გზა. მაგალითად, 4.4 სურათზე მოცემულ ბინარულ გამოსახულებაზე A წერტილი ბმულია B წერტილთან, რადგანაც შეგვიძლია ვიპოვოთ ამ წერტილების შემაერთებელი და მთლიანად ამ არეში მოთავსებული უწყვეტი წირი. მაგრამ A წერტილი არ არის ბმული C -თან, რადგანაც მათი შემაერთებელი ასეთი წირის პოვნა შეუძლებელია.



სურ. 4.4

ბმულობა წარმოადგენს გამოსახულებაზე გამოყოფილი რაიმე ქვესიმრავლის (არის) ყველაზე მეტად ცნობილ ტოპოლოგიურ თვისებას. 4.5 სურათზე ნახვენებია (ა) ბინარული გამოსახულება, რომელიც შეიცავს ხუთ ბმულ კომპონენტს, რომლებშიც პიქსელების მნიშვნელობა არის ერთის ტოლი.

ეს კომპონენტები აკმაყოფილებს როგორც 4-ბმულობის, ასევე 8-ბმულობის პირობებს, ანუ ბმულები არიან ორივე პირობის მიხედვით.

1	1	0	1	1	1	0	1
1	1	0	1	0	1	0	1
1	1	1	1	0	0	0	1
0	0	0	0	0	0	0	1
1	1	1	1	0	1	0	1
0	0	0	1	0	1	0	1
1	1	0	1	0	0	0	1
1	1	0	1	0	1	1	1

ა

1	1	0	1	1	1	0	2
1	1	0	1	0	1	0	2
1	1	1	1	0	0	0	2
0	0	0	0	0	0	0	2
3	3	3	3	0	4	0	2
0	0	0	3	0	4	0	2
5	5	0	3	0	0	0	2
5	5	0	3	0	2	2	2

ბ

სურ. 4.5

განსაზღვრება 4.2. *B* ბინარული გამოსახულების ბმული კომპონენტების მარკირების ქვეშ იგულისხმება ისეთი მარკირებული *LB* გამოსახულების ფორმირება, რომელშიც ყოველ პიქსელს მინიჭებული აქვს იმ ბმული კომპონენტის ჭდე, რომელსაც მოცემული პიქსელი მიეკუთვნება.

ჭდე წარმოადგენს გარკვეულ იდენტიფიკატორს, რომელიც გამოყენებულია როგორც ობიექტის უნიკალური სახელი. ხანდახან, თუმცა იშვიათად, გამოიყენებენ სიმბოლურ ჭდეებსაც, მაგრამ საერთოდ ბმული კომპონენტების მარკირებისთვის უფრო მოხერხებულია ჭდეებად დადებითი მთელი რიცხვების გამოყენება. 4.5,ბ სურათზე ნაჩვენებია (ა) ბინარული გამოსახულების ბმული კომპონენტების მარკირების შედეგად მიღებული მარკირებული გამოსახულება.

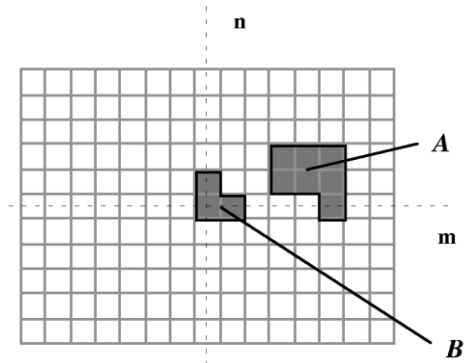
ბმული კომპონენტების მარკირებისთვის გამოიყენება რამდენიმე ერთმანეთისგან განსხვავებული ალგორითმი. ზოგიერთ ალგორითმში ნაგარაუდევია, რომ გამოსახულება მთლიანად არის შენახული მეხსიერებაში და მაშინ თითოეული კომპონენტის მარკირებისთვის გამოიყენება უბრალო რეკურსიული დამუშავება, რომლის დროსაც შესაძლებელია მთელი გამოსახულების პიქსელების გადათვალიერება. მაგრამ არსებობს ალგორითმები, რომლებიც განკუთვნილია ისეთი დიდი განზომილების გამოსახულებების დამუშავებისათვის, რომლებიც მთლიანად ვერ თავსდება მეხსიერებაში და მათი დამუშავება ხდება ნაწილ-ნაწილ ორ-ორ სტრიქონად.

4.4. ბინარული გამოსახულებების მორფოლოგია

ტერმინ მორფოლოგიის ქვეშ იგულისხმება ობიექტების ზომის, სტრუქტურისა და ფორმის თვისებების აღწერა. კომპიუტერულ ხედვაში ამ ტერმინის ქვეშ გულისხმობენ გამოსახულებაზე არსებული არეების (ობიექტების) ფორმის თვისებების აღწერას. მათემატიკური მორფოლოგია არის გეომეტრიული სტრუქტურების დამუშავებისა და ანალიზის თეორია და მეთოდოლოგია. ციფრულ ბინარულ და ნახევარტონურ გამოსახულებებში ის ჩვეულებრივად გამოიყენება როგორც გამოსახულებაზე არსებული არეების ფორმის თვისებების შესახებ გეომეტრიული ინფორმაციის გამოყოფის ინსტრუმენტი მათი შემდგომი წარმოდგენისა და აღწერის მიზნით. დასაწყისში მათემატიკური მორფოლოგიის ოპერაციები განსაზღვრული იყო როგორც ოპერაციები სიმრავლეებზე, მაგრამ ისინი ასევე სასარგებლო აღმოჩნდა ორგანზომილებიანი წერტილების სიმრავლეების დამუშავების ამოცანებშიც. ჩვენ მიერ განხილული იქნება ბინარული მორფოლოგიის ოპერაციები და ნაჩვენები იქნება ამ

ოპერაციების გამოყენების შესაძლებლობა ბმული კომპონენტების მარკირების შედეგად მიღებული არეების დამუშავებისათვის.

ჩვენ ზემოთ უკვე შემოვიტანეთ უწყვეტი და ციფრული გამოსახულებების განსაზღვრებები. მაგრამ აღმოჩნდა, რომ გამოსახულება (როგორც უწყვეტი, ასევე დისკრეტული) შეიძლება განისაზღვროს სხვანაირადაც. კერძოდ, როგორც კოორდინატების სიმრავლე. ეს უკანასკნელი შეესაბამება გამოსახულებაზე არსებული ობიექტების შესაბამის წერტილებს (პიქსელებს). გამოსახულების ასეთი წარმოდგენა ილუსტრირებულია 4.6 სურათზე, სადაც მოცემულია ბინარული გამოსახულება, რომელიც შედგება ორი ობიექტისგან, ანუ A და B სიმრავლეებისგან, და ასეთი წარმოდგენისთვის აუცილებელი კოორდინატთა (m, n) სისტემისგან. მაგალითისთვის, ობიექტი B მოცემულ გამოსახულებაზე შედგება წერტილებისგან – $[0,0]$, $[1,0]$, $[0,1]$.

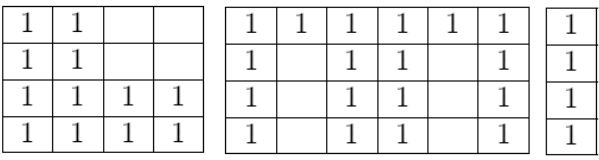
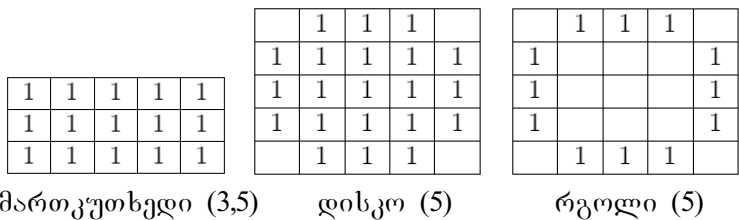


სურ. 4.6

ბინარული მორფოლოგიის ოპერაციებისთვის შესასვლელ მონაცემებს წარმოადგენს ბინარული გამოსახულება B და ე.წ. სტრუქტურული ელემენტი (structuring element) S , რომელიც ჩვე-

ულებრივად წარმოადგენს კონკრეტული კონფიგურაციის მქონე ბინარული დისკრეტული წერტილების სიმრავლეს, გამიზნულს განსაზღვრული ოპერაციების ჩატარებისთვის. სტრუქტურული ელემენტი შეიძლება ჩავთვალოთ როგორც რაიმე ფორმის მქონე არის აღწერა. ამ არეს შეიძლება ჰქონდეს ნებისმიერი ზომა და ბინარულ სახეში წარმოდგენილი ფორმა. ყოველი სტრუქტურული ელემენტი შეიძლება წარმოვიდგინოთ როგორც ზონდი, რომელიც გამოიყენება ბინარული გამოსახულებიდან მონაცემების ასაღებად. ასევე სტრუქტურული ელემენტისთვის უნდა განისაზღვროს ის წერტილი (პიქსელი), რომელიც აირჩევა როგორც მისი *საწყისი* – *ათვლის წერტილი*, რათა მოვახდინოთ მისი პოზიციონირება გამოსახულების რომელიმე პიქსელის მიმართ. ეს ნიშნავს შემდეგს: სტრუქტურული ელემენტი გამოსახულების x წერტილში გვიჩვენებს, რომ მისი ათვლის წერტილი ემთხვევა x წერტილს. სიმეტრიული სტრუქტურული ელემენტის შემთხვევაში ათვლის წერტილად აიღება ცენტრალური წერტილი, თუმცა შეიძლება ნებისმიერი სხვა წერტილის აღებაც. გამოსახულების ელემენტარული სიმეტრიული სტრუქტურული ელემენტი განისაზღვრება როგორც წერტილი და მისი მეზობლები განსახილველ მიდამოში. მაგალითად, ელემენტარული სამკუთხედები ხშირად განიხილება ჰექსაგონალურ ბადეში, ხოლო (2×2) განზომილების კვადრატები კი – კვადრატულ ბადეში. ამის მაგალითს $2D$ გამოსახულების შემთხვევაში წარმოადგენს (3×3) განზომილების ცენტრირებული ფანჯარა განსაზღვრული 8-ბმული ელემენტის ბადეში. არსებობს ყველაზე უფრო ხშირად გამოყენებული ასეთი სტრუქტურული ელემენტების კრებული, მაგალითად, მოცემული ზომების მქონე მართკუთხედი – $[მართკუთხედი(l, w)]$ ან მოცემული დიამეტრის მქონე დისკო – $[დისკო(d)]$. გამოსახულებათა დამუშავების პაკეტებში არსებობს ასეთი ელემენტებ-

ის მთელი ბიბლიოთეკა. რამდენიმე სახის სტრუქტურული ელემენტი ნაჩვენებია 4.7 სურათზე.



სურ. 4.7

ათვის წერტილის ბინარულ გამოსახულებაზე მიბმით სტრუქტურული ელემენტი შეიძლება გადატანილ იქნეს გამოსახულების ნებისმიერ პიქსელში და გამოყენებული ბინარული გამოსახულების არისთვის რომელიმე ფორმის მქონე ელემენტის დასამატებლად, ან იმის შესამოწმებლად, თავსდება თუ არა ელემენტის ფორმა გამოსახულების მოცემულ არეში. მაგალითად, რომელიმე დეტალზე არსებული ნახვრეტის ზომის კონტროლისათვის დისკოს ფორმის სტრუქტურული ელემენტის გამოყენებით შესაძლებელია შევამოწმოთ, რომ პატარა ზომის დისკო მთლიანად თავსდება მოცემულ არეში, ხოლო დიდი ზომის დისკო კი – არა.

4.4.1. ბინარული მორფოლოგიის ოპერაციები

ჩვენ განვიხილავთ ბინარული მორფოლოგიის ძირითად ოპერაციებს. ასეთებია: გაფართოება (dilation), ეროზია (erosion), ჩა-

კეტვა (closing) და **გაღება** (opening). ოპერაციების ამ სახელებიდან ჩანს მათი არსი, კერძოდ, გაფართოების ოპერაცია აფართოებს გამოსახულებაზე არსებულ არეს, ხოლო ეროზია კი პირიქით – არეს ამცირებს. ჩაკეტვის ოპერაცია შესაძლებლობას იძლევა, რომ მოვახდინოთ არის შიდა ხვრელების ჩაკეტვა და თავიდან მოვიშოროთ არის საზღვრის გასწვრივ არსებული შეზნექილობები (უბეები), ხოლო გაღების ოპერაციის საშუალებით ხდება არის საზღვართან შევრილების სახით არსებული პატარა ფრაგმენტების თავიდან მოშორება.

დახურვისა და გაღების ოპერაციების გამოყენება სასურველია გამოსახულებათა დამუშავების ისეთ ამოცანებში, სადაც წინასწარი დამუშავების ოპერაციების (მაგალითად, ზღურბლური ბინარიზაციის ან წინასწარი დამუშავების რაიმე სხვა ოპერაციის) ჩატარების შემდეგ ხდება ბინარული გამოსახულების ფორმირება პატარა ხვრელების მქონე ბმული სახეების სახით, ან როცა ბმული კომპონენტების წყვილები, რომლებიც გამოსახულებაზე ერთმანეთისგან განცალკავებული უნდა იყოს, აღმოჩნდება შეერთებული წინა პლანის პიქსელების ვიწრო ზოლით.

განვიხილოთ ახლა მორფოლოგიური ოპერაციების მათემატიკური განსაზღვრებები.

განსაზღვრება 4.3. *პიქსელთა X სიმრავლის t ვექტორით გადატანის (translation) X_t ოპერაცია მოიცემა შემდეგნაირად*

$$X_t = \{x + t | x \in X\}.$$

როგორც ვხედავთ, ბინარულ გამოსახულებაზე 1-ის ტოლი სიკაშკაშის პიქსელების სიმრავლის გადატანას თან სდევს სიმრავლის ყველა პიქსელის წანაცვლება მოცემული მანძილ-

ით. გადატანის t ვექტორი მოიცემა მოწესრიგებული (δ_i, δ_j) წყვილის საშუალებით, სადაც, δ_i გადატანის ვექტორის კომპონენტია სტრიქონის მიმართულებით, ხოლო δ_j – სვეტის მიმართულებით.

განსაზღვრება 4.4. ბინარული B გამოსახულების S სტრუქტურული ელემენტით გაფართოება აღინიშნება როგორც $B \oplus S$ და მოიცემა გამოსახულებით

$$B \oplus S = \cup_{s \in S} (B + s).$$

როგორც ვხედავთ გაფართოება წარმოადგენს პიქსელის მიდამოზე გამოყენებულ გაერთიანების ოპერაციას. S სტრუქტურული ელემენტი გამოიყენება ბინარული გამოსახულების ყველა პიქსელის მიმართ. ყოველთვის, როცა ხდება სტრუქტურული ელემენტის კოორდინატთა სათავის და 1-ის ტოლი ბინარული პიქსელის შეთავსება, მთელი სტრუქტურული ელემენტის მიმართ გამოიყენება გადატანის და შემდეგ კი ლოგიკური შეკრების ოპერაციები გამოსახულების შესაბამისი პიქსელების მიმართ. ლოგიკური შეკრების შედეგები კი გვაძლევს ბინარულ გამოსახველელ გამოსახულებას (სურ. 4.8), სადაც (ა) საწყისი B ბინარული გამოსახულებაა, (ბ) (3×3) განზომილების S მართკუთხა სტრუქტურული ელემენტი, (გ) გაფართოების შედეგად მიღებული ბინარული გამოსახულება.

განვიხილოთ ახლა თუ რას წარმოადგენს გაფართოების ოპერაცია მათემატიკური თვალსაზრისით B ბინარული გამოსახულების პირველი 1-ის ტოლი სიკაშკაშის პიქსელის მაგალითზე. ამ პიქსელის კოორდინატებია $[1,0]$, ანუ ის მდებარეობს B გამოსახულების პირველი სტრიქონისა და ნულოვანი სვეტის გადაკვეთის ადგილას. ამ პიქსელის მიმართ გადატანის

$S_{(1,0)}$ ოპერაციის შემდეგ S სტრუქტურული ელემენტის კოორდინატა სათავე დაემთხვევა $[1,0]$ პიქსელს.

1	1	1	1	1	1	1	1	
			1	1	1	1	1	
			1	1	1	1	1	
		1	1	1	1	1	1	
		1	1	1	1	1	1	
		1	1					

ა

1	1	1
1	1	1
1	1	1

ბ

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1	1	1	1
	1	1	1	1			

გ

				1	1		
				1	1		
				1	1		

დ

	1	1	1	1	1	1	
		1	1	1	1	1	
		1	1	1	1	1	
		1	1	1	1	1	
		1	1	1	1	1	
		1	1	1	1	1	
		1	1				

ე

		1	1	1	1		
		1	1	1	1		
		1	1	1	1		
		1	1	1	1		
		1	1	1	1		

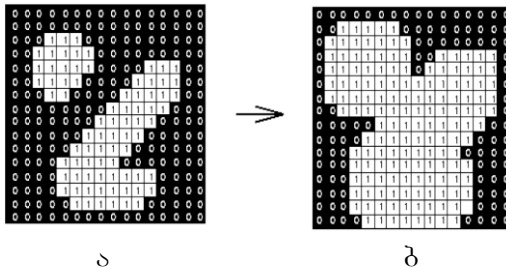
ვ

სურ. 4.8

ლოგიკური შეკრების ოპერაციის შედეგად გამოსასვლელ გამოსახულებაზე (რომელიც დასაწყისში ნულებითაა შევსებული) 1-ის ტოლი სიკაშკაშის პიქსელები მოთავსდება კოორდინატებში $[0,0]$, $[0,1]$, $[1,0]$, $[1,1]$, $[2,0]$ და $[2,1]$, და ასევე ვირტუალურ კოორდინატებში $[0,-1]$, $[1,-1]$ და $[2,-1]$, რომ-

ლებიც მხედველობაში არ მიიღება. B ბინარული გამოსახულების შემდეგი 1 -ის ტოლი სიკაშკაშის პიქსელისთვის კოორდინატებით $[1,1]$ სრულდება გადატანა $S_{(1,1)}$. ლოგიკური ოპერაციის ჩატარების შედეგად გამოსახულელ გამოსახულებაზე მივიღებთ პიქსელებს კოორდინატებით $[0,0], [0,1], [0,2], [1,0], [1,1], [1,2], [2,0], [2,1], [2,2]$. ასე გაგრძელდება მანამ, სანამ არ მოხდება სტრუქტურული ელემენტის ასლების (გადატანის ოპერაციის შედეგად მიღებულების) B საწყისი გამოსახულების თითოეულ 1 -ის ტოლი სიკაშკაშის პიქსელთან ლოგიკური შეკრება. შედეგად მიიღება 4.8.გ სურათზე წარმოდგენილი გამოსახულელი გაფართოებული გამოსახულება.

4.9,ა სურათზე წარმოდგენილია წყვეტის შემცველი ბინარული გამოსახულება, რომელზედაც წყვეტის შესავსებად გამოყენებულია გაფართოება სტრუქტურული ელემენტით (3×3) (სურ. 4.9,ბ).



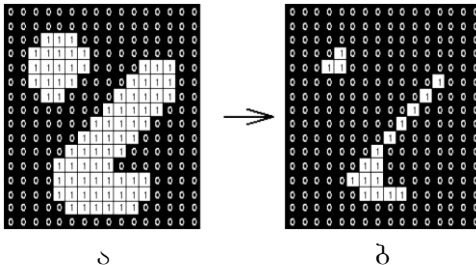
სურ. 4.9

განსაზღვრება 4.5. ბინარული B გამოსახულების S სტრუქტურული ელემენტით *ეროზია* აღინიშნება როგორც $B \ominus S$ და მოიცემა გამოსახულებით

$$B \ominus S = \{b | b + s \in B \ \forall s \in S\}.$$

ეროზიის ოპერაციის შესრულებისას სტრუქტურული ელემენტი, ისევე როგორც გაფართოებისას, ელემენტ-ელემენტ “გაიფლის” შესასვლელი ბინარული გამოსახულების ყველა პიქსელს. თუ გამოსახულებაზე გადაადგილებისას რომელიმე პოზიციაში მყოფი სტრუქტურული ელემენტის ყოველი 1-ის ტოლი სიკაშკაშის პიქსელი ემთხვევა ბინარული გამოსახულების 1-ის ტოლი სიკაშკაშის პიქსელს, მაშინ ხდება სტრუქტურული ელემენტის ცენტრალური პიქსელის ლოგიკური აჯამვა გამოსასვლელი გამოსახულების შესაბამის პიქსელთან. 4.8,დ სურათზე ნაჩვენებია ბინარული გამოსახულების (სურ. 4.8,ა) სტრუქტურული ელემენტით (სურ. 4.8,ბ) ეროზიის ოპერაციის გამოყენების შედეგი.

4.10 სურათზე წარმოდგენილია ბინარული გამოსახულება ბიდან (სურ. 4.10,ა) არასასურველი დეტალების მოცილების მიზნით ეროზიის ოპერაციის გამოყენება სტრუქტურული ელემენტით (3 × 3) (სურ. 4.10,ბ).



სურ. 4.10

გაფართოება და ეროზია წარმოადგენს მათემატიკური მორფოლოგიის ყველაზე უფრო მარტივ ოპერაციებს. მათი

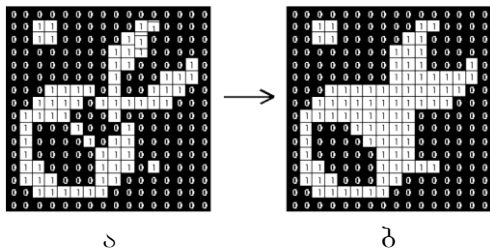
კომპოზიცია კი წარმოშობს კიდევ ორ, ხშირად გამოყენებულ ოპერაციებს: ჩაკეტვას და გაღებას.

განსაზღვრება 4.6. ბინარული B გამოსახულების S სტრუქტურული ელემენტით **ჩაკეტვა** აღინიშნება როგორც $B \cdot S$ და მოიცემა გამოსახულებით

$$B \cdot S = (B \oplus S) \ominus S.$$

როგორც ვხედავთ ბინარული გამოსახულების ჩაკეტვა ნიშნავს ერთი და იგივე მასტრუქტურირებელი ელემენტის გამოყენებით ჯერ ეროზიის ოპერაციის ჩატარებას და შემდეგ გაფართოების. 4.8,ა სურათზე წარმოდგენილი ბინარული გამოსახულების სტრუქტურული ელემენტით (სურ. 4.8,ბ) ჩაკეტვის ოპერაციის შედეგი წარმოდგენილია 4.8,ე სურათზე.

სურათზე 4.11 წარმოდგენილია ბინარულ გამოსახულებაზე (სურ. 4.11,ა) ჩაკეტვის ოპერაციის გამოყენების შედეგი სტრუქტურული ელემენტით (3×3) (სურ. 4.11,ბ).



სურ. 4.11

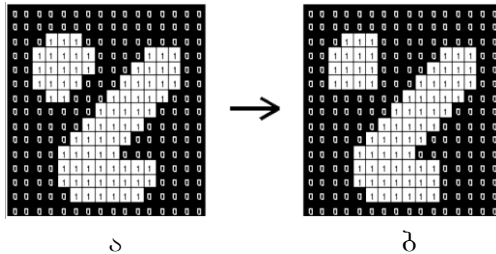
განსაზღვრება 4.7. ბინარული B გამოსახულების S სტრუქტურული ელემენტით **გაღება** აღინიშნება როგორც $B \circ S$ და მოიცემა გამოსახულებით

$$B \circ S = (B \ominus S) \oplus S.$$

გაღების ოპერაციაში კი პირიქითაა. ჯერ ერთი და იგივე სტრუქტურული ელემენტით ხდება ბინარული გამოსახულების გაფართოება, ხოლო შემდეგ კი – ეროზია. ბინარულ გამოსახულებაზე (სურ. 4.8,ა) სტრუქტურული ელემენტით (სურ. 4.8,ბ) გაღების ოპერატორის მოქმედების შედეგი მოცემულია 4.8,ვ სურათზე.

ასევე გაღების ოპერატორის ბინარულ გამოსახულებაზე მოქმედების შედეგი (სურ. 4.12,ა) სტრუქტურული ელემენტით (3×3) , წარმოდგენილია სურათზე 4.12,ბ.

ბინარული მორფოლოგიის მეთოდები შეიძლება გამოყენებულ იქნეს მანქანური ხედვის ინსპექტირების სამრეწველო სისტემებში, მაგალითად, საათების მექანიზმში არსებულ კბილანებზე დეფექტების გამოსავლენად, სამედიცინო გამოსახულებების დამუშავებისას და ა.შ.



სურ. 4.12

ბინარული მორფოლოგია ასევე შეიძლება გამოყენებულ იქნეს ობიექტის პრიმიტიული (ელემენტარული) მახასიათებელი ნიშან-თვისებების გამოსაყოფად ობიექტის ამოცნობის პროცესში მათი შემდგომი გამოყენების მიზნით (სურ. 4.13). მაგალითად, ბრტყელი ორგანზომილებიანი ობიექტის (სურ. 4.13,ა) ფორმის

ამოცნობისთვის პრიმიტიული ნიშან-თვისებების როლში შეიძლება გამოყენებულ იქნეს მისი კუთხეები. თუ მოვახდენთ მკვეთრი კუთხეების მქონე ობიექტის გამოსახულების გაღებას დისკოს ტიპის სტრუქტურული ელემენტით, მაშინ მივიღებთ “კუთხეებმოჭრილი” ობიექტის გამოსახულებას (სურ. 4.13,ბ). თუ ახლა საწყის გამოსახულებას გამოვაკლებთ ამ უკანასკნელს, მივიღებთ გამოსახულებას, რომელზედაც ობიექტის მხოლოდ კუთხეები იქნება მოცემული (სურ. 4.13,გ).

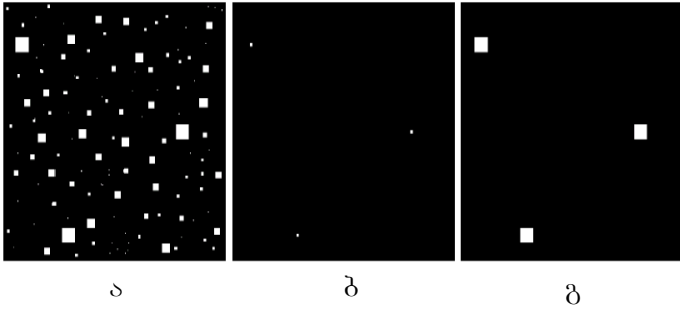


სურ. 4.13

ეს გამოსახულება შეიძლება გამოყენებულ იქნეს ობიექტის ფორმის ამოცნობისათვის. საზოგადოდ კი, ობიექტთა ფორმის ამოცნობის სისტემაში ნიშან-თვისებათა მორფოლოგიური გამოყოფის გამოყენება სასარგებლოა, რადგანაც იგი პრიმიტივების სწრაფი აღმოჩენის შესაძლებლობას იძლევა.

ბინარული მორფოლოგიის ოპერაციების გამოყენებით გამოსახულებაზე მცირე ზომის არეების (ობიექტების) მოშორების მაგალითი წარმოდგენილია 4.14 სურათზე. როგორც ვხედავთ, საწყისი ბინარული გამოსახულება (სურ. 4.14,ა) შეიცავს სხვადასხვა ზომის კვადრატებს, მათ შორის ძალიან მცირე ზომისაც, რომელთა მოსაშორებლად გამოყენებულია კვადრატული ფორმის სტრუქტურული ელემენტი. ჯერ ხდება საწყისი გამოსახულების ეროზია (სურ. 4.14,ბ), ხოლო მიღებულ გამოსა-

ხულებზე შემდეგ იყენებენ გაფართოების ოპერატორს. შედეგად მიიღება ბინარული გამოსახულება, რომელზედაც უკვე მცირე ზომის ობიექტები აღარ არი (სურ. 4.14,გ).



სურ. 4.14

4.5. ბინარულ გამოსახულებაზე არეების მომიჯნაობის აღწერა გრაფების საშუალებით

გამოსახულებათა ანალიზისას მნიშვნელოვანია არა მარტო ცალკეული არეების თვისებების განსაზღვრა, არამედ მათ შორის სივრცითი ურთიერთმიმართებებისა და ურთიერთკავშირების ცოდნაც. ეს უკანასკნელი გულისხმობს გამოსახულებაზე ერთმანეთის **მომიჯნავე არეების (region adjacency)** აღმოჩენას. ორი არე ერთმანეთის მომიჯნავედ ჩაითვლება იმ შემთხვევაში, თუ ერთი არის პიქსელი წარმოადგენს მეორე არის პიქსელის მეზობელს. როგორც ვიცით, ბინარულ გამოსახულებაზე სულ გვაქვს ორი ტიპის არე: წინა პლანის არეები და ფონური არეები. ამიტომ, ზემოთქმულის თანახმად, ასეთ გამოსახულებაზე წინა პლანის ყველა არე ფონური არეების მომიჯნავეა, მაგრამ არა ერთმანეთის. იმ მარტივ შემთხვევაში, როცა ფონი

გამოსახულებაზე წარმოდგენს ერთ ბმულ არეს, მაშინ არეებს შორის მომიჯნაობის თვისებების გამოთვლა ამით სრულდება.

დაეუშვათ, რომ წინა პლანის არეები შეიცავს ხვრელებს. ამასთან, თითოეული ხვრელის არე მიეკუთვნება ფონს. თუ მოვახდენთ წინა პლანის პიქსელებისთვის ბმული კომპონენტების მარკირებას, ანუ მივანიჭებთ მათ შესაბამის ჭდეებს, მაშინ მივიღებთ მარკირებულ გამოსახულებას, რომელზედაც წინა პლანის ყველა არეს ექნება გარკვეული რიცხვითი ჭდეები, ხოლო ფონური არეების ყველა პიქსელი მარკირებული იქნება ნულის ტოლი სიდიდის ჭდით. ანალოგიური პროცედურით შეიძლება ჩატარდეს ბმული კომპონენტების მარკირება ფონური პიქსელებისთვისაც. შედეგად ყველა ფონურ არეს ექნება შესაბამისი ჭდე. შეიძლება აღმოჩნდეს, რომ არეებიდან ყველაზე დიდი ფართობის მქონე არე იწყებოდა გამოსახულების მარცხენა ზედა კუთხეში. ამ არეს შეიძლება მიენიჭოს განსაკუთრებული ჭდე, მაგალითად, 0. მაშინ დანარჩენი, ნაკლები ფართის მქონე ფონური არეები შესაბამისი იქნება წინა პლანზე განლაგებული ხვრელების.

ამგვარად, ორი მარკირებული გამოსახულების საშუალებით, რომლებზედაც ჭდეების საშუალებით მონიშნულია (მარკირებულია) როგორც წინა პლანის, ასევე ფონური არეები, შეიძლება განისაზღვროს წინა პლანის თითოეული არისთვის მომიჯნავე ფონური არეები (ან პირიქით). გამოსახულებაზე არეების მომიჯნაობის შესახებ ინფორმაციის წარმოდგენისათვის იყენებენ სტრუქტურას, რომელსაც **არეების მომიჯნაობის გრაფს** (region adjacency graph, RAG) უწოდებენ. ამ სტრუქტურის საშუალებით არეების მომიჯნაობის წარმოდგენა შეიძლება როგორც ბინარული გამოსახულების, ასევე ზოგად შემთხვევაში ნახევარტონური და ფერადი გამოსახულებების სეგმენტაციის პროცესშიც.

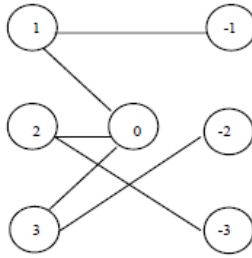
განსაზღვრება 4.8. გრავს, რომლის თითოეულ კვანძს შესაბამეა გამოსახულების რომელიმე არე, ხოლო წიბოებით ერთმანეთთან შეერთებულია მომიჯნავე არეების შესაბამისი კვანძები, არეების მომიჯნაობის გრაფი ეწოდება.

4.17 სურათზე წარმოდგენილია (ა) მარკირებული ბინარული გამოსახულება ჭდეებით მონიშნული წინა პლანის (დადებითი რიცხვები) და ფონური არეებით, რომელთაგან დიდები მარკირებულია ნულოვანი ჭდით, ხოლო ხვრელებიანი არეები კი უარყოფითი რიცხვებით, და მისი შესაბამისი (ბ) არეების მომიჯნაობის გრაფი.

წარმოდგენილ მაგალითში წინა პლანის არეებს, როგორც ეს ჩვეულებრივადაა მიღებული, ჭდეებად მინიჭებული აქვს დადებითი რიცხვები. რაც შეეხება ფონურ არეებს, აქ 0-ანი ჭდე მინიჭებული აქვს გამოსახულების ზედა მარცხენა კუთხეში მოთავსებულ ყველაზე დიდი ფართობის მქონე არეს.

0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	0	2	2	0
0	1	-1	-1	-1	1	0	2	2	0
0	1	1	1	1	1	0	2	2	0
0	0	0	0	0	0	0	2	2	0
0	3	3	3	0	2	2	2	2	0
0	3	-2	3	0	2	-3	-3	2	0
0	3	-2	3	0	2	-3	-3	2	0
0	3	3	3	0	2	2	2	2	0
0	0	0	0	0	0	0	0	0	0

ა



ბ

სურ. 4.17

შემდეგ კი მისი მნიშვნელობა მცირდება უარყოფითი მნიშვნელობებისკენ გამოსახულებაზე არსებული სხვადასხვა ფორმის მქონე ხერხელების შესაბამისი არეებისთვის, რომლებიც წინა პლანის არეებს მიეკუთვნება.

მომიჯნაობის გრაფის ასაგებად იყენებენ ალგორითმს, რომელიც შემდეგნაირად მოქმედებს: გამოსახულების დამუშავება ხდება ორსტრიქონიან პორციებად, რომლის დროსაც გამოსახულებაზე როგორც ჰორიზონტალური, ასევე ვერტიკალური მომიჯნაობის მიმართულებებით, ხოლო 8-ბმულიან შემთხვევაში დიაგონალური მიმართულებითაც, იძებნება განსხვავებული ჭდეების მქონე წერტილები (პიქსელები). ალგორითმი მანამდე არამომიჯნავე არეებს შორის მომიჯნაობის აღმოჩენის შემთხვევაში გრაფს ახალ წიბოებს ამატებს.

აღნიშნული ალგორითმის ეფექტურობაზე გავლენას ახდენს ორი ფაქტორი:

- **მონაცემთა სტრუქტურის ზომა.**

მარკირების შედეგად გამოსახულებაზე შეიძლება გექონდეს ათასობით ჭდე. ასეთ შემთხვევაში შესაძლოა შეუქმდებელი ან არარაციონალური აღმოჩნდეს მონაცემ-

თა მთელი გრაფული სტრუქტურის მთლიანად მესხიერებაში შენახვა.

- **გამოსახულების დამუშავების დროითი დანახარჯი.**

გამოსახულების პიქსელ-პიქსელ დამუშავებისას მომიჯნაობის ერთი და იგივე მიმართება (ანუ არეების ერთი და იგივე წყვილებისთვის) მრავალჯერ შეგვხვდება. შესაბამისად, ყოველი ასეთი შეხვედრისას ალგორითმს მოუხდება გრაფზე ერთი და იგივე წიბოს დამატება, რაც გამოსახულების დამუშავების დროს გაზრდის. ამიტომ სასურველია, რომ გრაფზე წიბოს დამატება მომიჯნავე არეების ყოველი წყვილისთვის მოხდეს ერთხელ.

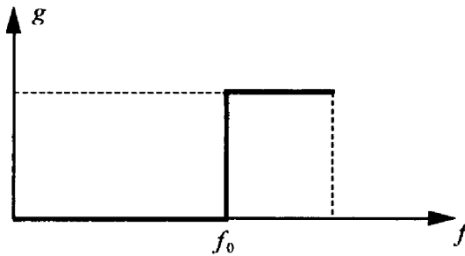
4.6. ზღურბლის გამოყენება ნახევარტონური გამოსახულებების ბინარიზაციისათვის

ნახევარტონური გამოსახულებიდან ბინარული გამოსახულების მისაღებად გამოიყენება გამოსახულების **ზღურბლური დამუშავების** ანუ **ზღურბლური ბინარიზაციის მეთოდი**. ასეთ გარდაქმნას იყენებენ იმისათვის, რომ გამოსახულებაზე ერთმანეთისაგან მკვეთრად გამიჯნონ წინა პლანის (foreground) პიქსელები, რომლებიც ჩვენთვის საინტერესო ობიექტებს შეესაბამებიან, იმ დანარჩენი პიქსელებისგან, რომლებიც ფონს (background) მიეკუთვნება, და აგრეთვე გამოსახულებას მოაცილონ მასზე არსებული არასაჭირო ინფორმაცია. ვიცით რა გამოსახულებაზე ინტენსივობის მნიშვნელობების განაწილება, ზღურბლის მოძებნა ასეთ შემთხვევაში პრობლემას არ წარმოადგენს. ზღურბლის მნიშვნელობად შეიძლება არჩეულ იქნეს ინტენსივობის ის მნიშვნელობა, რომელიც ყველაზე კარგად მოახდენს ამ დაყოფას. თუ $f(i, j)$ წარმოადგენს უწყვეტ შესასვლელ გამოსახულებას, ხოლო $g(i, j)$ ზღურბლური გარდაქმნის

შედეგად მიღებულ უწყვეტ გამოსახვლელ გამოსახულებას, მაშინ აღნიშნული გარდაქმნა ფორმალურად ასე ჩაიწერება:

$$g(i, j) = \begin{cases} 1, & \text{როცა } f(i, j) \geq f_0 \\ 0, & \text{როცა } f(i, j) < f_0 \end{cases}$$

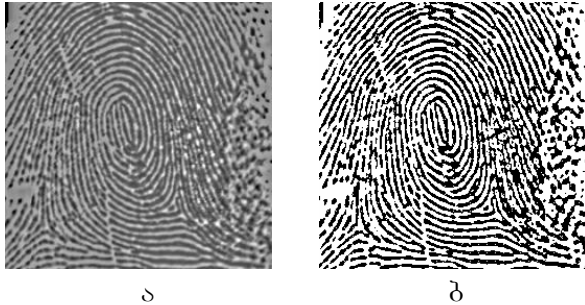
სადაც, f_0 ზღურბლის მნიშვნელობაა ინტენსივობის მიხედვით. გრაფიკულად ეს ასე შეიძლება წარმოვადგინოთ (სურ. 4.18).



სურ. 4.18

თითის ანაბეჭდის გამოსახულების (სურ. 4.19,ა) ზღურბლური ბინარიზაციის (სურ. 4.19,ბ) მაგალითი წარმოდგენილია 4.19 სურათზე.

მარტივი გამოსახულების შემთხვევაში, ანუ როცა გვაქვს ინტენსივობის ერთი მნიშვნელობის მქონე ობიექტები და ინტენსივობის მეორე მნიშვნელობის მქონე ფონი (მაგალითად, ნაბეჭდი ტექსტი, ნახაზი და ა.შ.), მაშინ ზღურბლის მნიშვნელობად აირჩევა ინტენსივობის ერთი $f_0 = t$ მნიშვნელობა. შედეგად, ყველა პიქსელი, რომელშიც ინტენსივობის მნიშვნელობა მეტი ან ტოლი იქნება t , ხდება წინა პლანის პიქსელი, ხოლო დანარჩენი კი ფონური. ასეთ ოპერაციას **ზედა ზღურბლური ბინარიზაცია** (threshold above) ეწოდება.

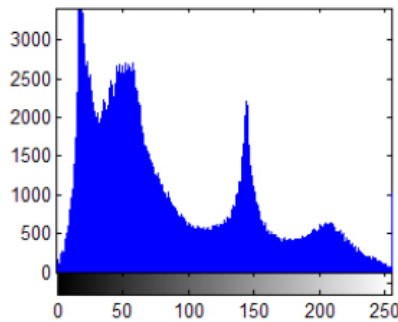


სურ. 4.19

მაგრამ არსებობს ზღურბლური ბინარიზაციის სხვა ნაირსახეობებიც. მაგალითად, **ქვედა ზღურბლური ბინარიზაცია (threshold below)**, როცა ზემოაღწერილისგან განსხვავებით ხდება პირიქით: წინა პლანის პიქსელებად ხდებიან ის პიქსელები, რომელთა ინტენსივობის მნიშვნელობები ნაკლებია ან ტოლი ზღურბლის t მნიშვნელობაზე. იმ შემთხვევაში, თუ ზღურბლად გამოიყენება ორი მნიშვნელობა, ქვედა და ზედა, მაშინ საქმე გვაქვს **ზღურბლურ ბინარიზაციასთან დიაპაზონის მიხედვით (threshold inside)**. ასეთ შემთხვევაში წინა პლანის პიქსელებად აირჩევა ის პიქსელები, რომელთა მნიშვნელობები მოთავსებულია ზღურბლურ მნიშვნელობებს შორის. ხოლო, თუ ხდება პირიქით, ანუ წინა პლანის პიქსელებად აირჩევა ის პიქსელები, რომელთა ინტენსივობის მნიშვნელობები მოთავსებულია ზღურბლური მნიშვნელობების გარეთ, მაშინ საქმე გვაქვს **გარე ზღურბლურ ბინარიზაციასთან (threshold outside)**. ნახევარტონური გამოსახულების ბინარულით შეცვლა ძალიან მოსახერხებელია, რადგანაც ამ დროს: ა) გამოსახულების აღქმა უფრო ადვილი ხდება და ბ) მკვეთრად მცირდება გამოსახულების შესანახად საჭირო მეხსიერების მოცულობა, რადგანაც, თუ ნახევარტონური გამოსახულების თითოეული პიქსელის შე-

ნახვა ძირითადად მოითხოვს მესხიერების 8 ბიტს, ბინარული გამოსახულების შემთხვევაში ეს 1 ბიტის ტოლია.

ამგვარად, ნახევარტონური გამოსახულების ბინარიზაციისათვის ზღურბლის არჩევა აუცილებელ ოპერაციას წარმოადგენს. გამოსახულებათა დამუშავების ბევრ ინტერაქტიულ პაკეტში მომხმარებლისთვის გათვალისწინებულია ზღურბლის შერჩევის შესაძლებლობა. ცხადია, გამოსახულებათა ავტომატური ანალიზის შემთხვევაში ზღურბლის შერჩევა ავტომატურად უნდა ხდებოდეს. ნახევარტონურ გამოსახულებაზე ზღურბლის შერჩევის მარტივ საშუალებას წარმოადგენს ინტენსივობის მნიშვნელობების (ნახევარტონების დონეების) განაწილების **ჰისტოგრამის** აგება. ჰისტოგრამა წარმოადგენს გრაფიკს (დიაგრამას), რომელიც გვიჩვენებს ინტენსივობის თითოეული მნიშვნელობის შესაბამისი პიქსელების რაოდენობას მოცემულ გამოსახულებაზე (სურ. 4.20).



სურ. 4.20

ის შეიძლება მოცემული იყოს ასეთი გამოსახულების სახით:

$$h(m) = |\{(i,j) | f(i,j) = m\}|,$$

სადაც, m აღნიშნავს ინტენსივობის მნიშვნელობებს უწყვეტ $f(i, j)$ გამოსახულებაზე. მაგალითად, 8-ბიტიან ციფრულ ნახევარტონურ გამოსახულებაზე გვაქვს ნახევარტონების 256 დონე, ანუ $m = 0, 1, 2, \dots, 255$. ჰისტოგრამის ასაგებად საჭიროა ვიმოდროთ გამოსახულებაზე პიქსელ-პიქსელ და დავითვალთ ინტენსივობის თითოეული მნიშვნელობის შესაბამისი პიქსელების საერთო რაოდენობა.

განვიხილოთ 8-ბიტიანი ციფრული ნახევარტონური გამოსახულების ჰისტოგრამის გამოთვლის მაგალითი. ვთქვათ, გვაქვს ასეთი გამოსახულება (სურ. 4.21)

52	55	61	66	70	61	64	73
63	59	55	90	109	85	69	72
62	59	68	113	144	104	66	73
63	58	71	122	154	106	70	69
67	61	68	104	126	88	68	70
79	65	60	70	77	68	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94

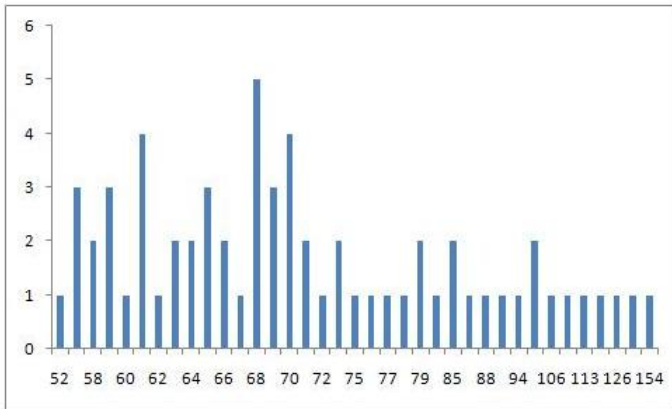
სურ. 4.21

მაშინ მის ჰისტოგრამას ექნება 4.22 სურათზე წარმოდგენილი სახე (კენტონომრიან სვეტებში მოცემულია სიკაშკაშის მნიშვნელობები, ხოლო ლუწონომრიანში კი მათი ჯამური რაოდენობაა გამოსახულებაზე).

52	1	64	2	72	1	85	2	113	1
55	3	65	3	73	2	87	1	122	1
58	2	66	2	75	1	88	1	126	1
59	3	67	1	76	1	90	1	144	1
60	1	68	5	77	1	94	1	154	1
61	4	69	3	78	1	104	2		
62	1	70	4	79	2	106	1		
63	2	71	2	83	1	109	1		

სურ. 4.22

ხოლო გრაფიკულად კი ჰისტოგრამას 4.23 სურათზე მოცემული სახე ექნება.



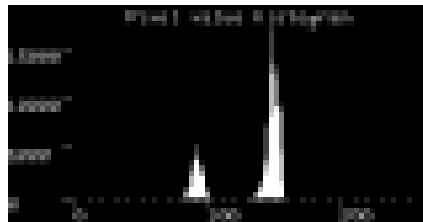
სურ. 4.23

ჰისტოგრამის აგება შეიძლება ასევე ფერადი გამოსახულებისთვის, მხოლოდ ამ დროს ის უნდა აიგოს თითოეული RGB ფერისთვის ცალ-ცალკე.

ავიღოთ მარტივი გამოსახულება, რომელიც წარმოადგენს ფონს მასზე მოთავსებული ერთი ობიექტით (სურ. 4.24,ა). მის ჰისტოგრამას აქვს 4.24,ბ სურათზე წარმოდგენილი სახე.



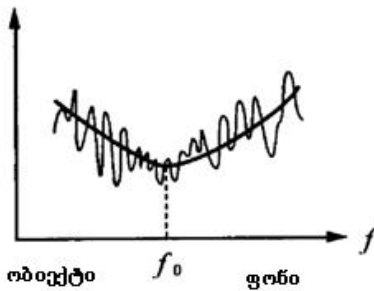
ა



ბ

სურ. 4.24

როგორც ვხედავთ, მიღებულ ჰისტოგრამაზე ინტენსივობის მნიშვნელობების განაწილებას აქვს ორი წვერო (მაქსიმუმი), რომლებიც ერთმანეთისგან საკმაოდ დაშორებულია და ერთმანეთს არ თანაკვეთს. ასეთ ჰისტოგრამას ორმოდიანი ან ბიმოდალური ჰისტოგრამა ეწოდება. აქ ერთი წვერო შეესაბამება ღია პიქსელებს, ხოლო მეორე კი – მუქს. მაგრამ შეიძლება გვქონდეს ისეთი ჰისტოგრამა, რომელზედაც მოდები ერთმანეთს ფარავს. თუ გადაფარვა მცირეა, მაშინ ზღურბლის როლში შეიძლება ავიღოთ მოდებს შორის მოთავსებული დაბლობიდან (ვაკიდან) ნებისმიერი მნიშვნელობა. დიდი გადაფარვის შემთხვევაში კი, ანუ როცა ღია და მუქი პიქსელების განაწილებები ერთმანეთს ძლიერად ფარავს, ზღურბლის არჩევა უკვე რთული ხდება, რადგანაც ამ დროს მოდებს შორის ზღურბლის ასარჩევი დაბლობი აღარ არის. ამას შეიძლება ადგილი ჰქონდეს მაშინ, როცა გამოსახულებაზე გვაქვს ხმაური ან ობიექტის და ფონის ინტენსივობის მნიშვნელობების დიდი გაბნევაა. ასეთ შემთხვევაში ზღურბლის განსაზღვრისათვის იყენებენ სხვადასხვა მეთოდებს. კერძოდ, ახდენენ ჰისტოგრამის მაქსიმუმებს შორის მოთავსებული უბნის აპროქსიმაციას რომელიმე გლუვი ფუნქციის საშუალებით, მაგალითად პარაბოლით და შემდეგ ეძებენ მის მინიმუმს (სურ. 4.25)



სურ. 4.25

V თავი. ფორმის წარმოდგენა და აღწერა

5.1. არის ფორმის წარმოდგენა და აღწერა

ბინარულ გამოსახულებაზე არეების (ობიექტების) აღმოჩენისა და გამოყოფის შემდეგ შესაძლოა საჭირო გახდეს მათი მახასიათებელი ნიშან-თვისებების განსაზღვრა. ეს უკანასკნელები შემდგომში შეიძლება გამოყენებულ იქნეს გამოსახულების ანალიზის უფრო მაღალი დონის პროცედურების ჩასატარებლად, როგორცაა: გამოსახულებაზე ობიექტების ადგილმდებარეობის განსაზღვრა, ობიექტების ამოცნობა, საინსპექციო ამოცანები, მსგავსი ობიექტების მოძებნა გამოსახულებათა დიდ კოლექციებში ან განაწილებულ მონაცემთა ბაზებში. მაგალითად, სამედიცინო გამოსახულების ანალიზისას რომელიმე ორგანოს ფორმა მნიშვნელოვან როლს ასრულებს დაავადების დიაგნოსტიკის საქმეში. ასევე მნიშვნელოვანია ფორმა ასტრონომიებისთვის, რომლებიც მის საფუძველზე ახდენენ გალაქტიკების, ვარსკვლავთა სისტემების, პლანეტების კლასიფიკაციას, რაც აუცილებელი ნაბიჯია სამყაროს სტრუქტურის სხვადასხვა ასპექტების შესწავლის და ახსნის გზაზე. უნდა აღინიშნოს, რომ ბიოლოგიური ფორმის ანალიზი წარმოადგენდა გამოსახულებათა ანალიზის ისტორიულად ერთ-ერთ პირველ გამოყენებით ამოცანას.

საზოგადოდ, არ არსებობს ფორმის უნივერსალური განსაზღვრება. ის შთაბეჭდილება, რომელიც ადამიანში წარმოიშვება ობიექტის ფორმის შესახებ, ყალიბდება ფერის ან ინტენსივობის პატერნების (ტექსტურების) აღქმის შედეგად, საიდანაც ის უკვე აყალიბებს წარმოდგენას ამ ობიექტის გეომეტრიის შესახებ. ჩვენ ტერმინ “ფორმის” ქვეშ ვგულისხმობთ გეომეტრიულ პატერნს (სახეს), რომელიც სიმრავლეთა თეო-

რის მიხედვით შეიძლება წარმოადგენდეს ტოპოლოგიური აზრით კომპაქტურ და ჩაკეტილ წერტილთა რაღაც სიმრავლეს (მრუდებს, ზედაპირებს, მყარ სხეულებს და ა.შ.).

გამოსახულებაზე გამოყოფილი არის (ობიექტის) ფორმა წარმოადგენს გამოსახულების მნიშვნელოვან ვიზუალურ ნიშან-თვისებას, რომელიც გამოიყენება გამოსახულების შინაარსობრივი აღწერისათვის. ზემოთქმულიდან გამომდინარე, დგება ფორმის წარმოდგენისა და აღწერის ამოცანა, რაც თავისთავად რთულ პრობლემას წარმოადგენს.

2D გამოსახულებაზე არის ფორმის წარმოდგენა უპირველეს ყოვლისა მოითხოვს მისთვის მნიშვნელოვანი ნიშან-თვისებების განსაზღვრას იმ ინფორმაციაზე დაყრდნობით, რომელიც დევს:

- **ობიექტის ფორმის გარშემომწერ კონტურსა (საზღვარსა) და მის ნიშან-თვისებებში;**
- **კონტურსა და მისით გარშემოწერილ არეში.**

ასეთი სახის ბევრი ნიშან-თვისებაა შემუშავებული, როგორცაა მაგალითად, **ფორმის მოხაზულობა (shape signature), მოხაზულობის ჰისტოგრამა (signature histogram), ფორმის ინვარიანტები (shape invariants), მომენტები (moments), სიმრუდე (curvature)** და სხვ. ცხადია, თითოეული ეს ნიშან-თვისება ფასდება იმის მიხედვით, თუ რამდენად ზუსტად შეუძლია მათ მოცემულ მონაცემთა ბაზაში მოიძიონ მსგავსი ფორმები. აგალითად, ფორმაზე დამყარებული გამოსახულებების ძიება დღესდღეობით წარმოადგენს ერთ-ერთ აქტუალურ კვლევის საგანს გამოსახულებათა დამუშავებისა და ანალიზის სფეროში. მისი გამოყენების სფეროებია მეტეოროლოგია, მედიცინა, კოსმოსი, წარმოება, გართობის ინდუსტრია, განათლება, მართლ-

წესრიგი და თავდაცვა. ამავე დროს, ფორმის წარმოდგენა და აღწერა უნდა იყოს ისეთი, რომ გაადვილდეს მისი შენახვა, გადაცემა შორ მანძილზე, შედარება სხვა ფორმასთან და ამოცნობა. თავის მხრივ, მონაცემთა ბაზაში მსგავსი ფორმების ეფექტური მოძიება მოითხოვს ფორმის მაღალი სიზუსტით აღწერას. ამასთან, ფორმის აღწერა უნდა იყოს ინვარიანტული გადაადგილების, ბრუნვის და მასშტაბირების მიმართ.

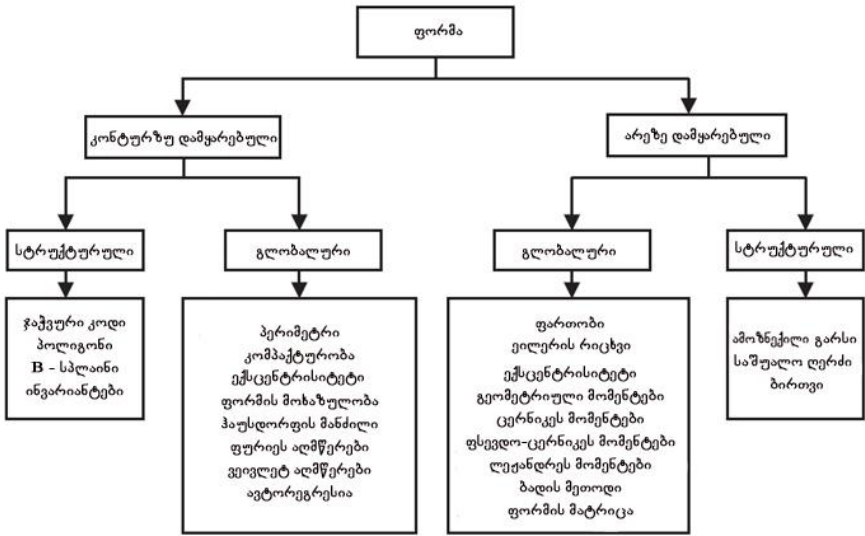
მას შემდეგ, რაც ციფრულ გამოსახულებაზე მოხდა ობიექტის გამოყოფა, (ჩვეულებრივად, სეგმენტაციის და სეგმენტირებული გამოსახულების ბინარიზაციის შედეგად) გამოსახულების ანალიზის ამოცანა მოითხოვს ობიექტის ფორმის აღწერას, რისთვისაც გამოიყენება ე.წ. **ფორმის აღმწერი** (shape descriptor). საზოგადოდ, ფორმის აღმწერი წარმოადგენს რაღაც რიცხვების სიმრავლეს, რომელიც შექმნილია მოცემული ფორმის აღწერისათვის. ამასთან განსხვავებული ფორმებისთვის აღმწერები საკმაოდ განსხვავებულები უნდა იყოს, რათა შეძლოს მათი ერთმანეთისაგან გარჩევა. ასევე სხვადასხვა ფორმებისთვის ისინი უნდა ხასიათდებოდეს ძიების მაღალი სიზუსტით, კომპაქტურობით, დაბალი გამოთვლითი სირთულით, მდგრადი ეფექტური ძიებით. ფორმის აღმწერებს ასევე უნდა შეეძლოს ხმაურის შემცველი, დამახინჯებული და დეფექტიანი ფორმების პოვნა. ფორმის მარტივი აღმწერებია: ფართობი, პერიმეტრი, კომპაქტურობა და ა.შ.

ფორმის აღმწერები ძირითადადში იყოფა ორ კატეგორიად:

- კონტურზე დამყარებული აღმწერები;
- არეზე დამყარებული აღმწერები.

ზემოთქმულიდან გამომდინარე, ფორმის წარმოდგენისა და აღწერის ხერხები იყოფა ორ კლასად: **კონტურზე დამყარე-**

ბულ და არეზე დამყარებულ მეთოდებად. ასეთი დაყოფა განპირობებულია იმით, თუ საიდან ხდება ნიშან-თვისებების გამოყოფა – მხოლოდ კონტურიდან, თუ არიდან. ხოლო თითოეულ ამ კლასში შემავალი მეთოდები კი თავის მხრივ იყოფა სტრუქტურულ და გლობალურ მეთოდებად, რაც განისაზღვრება იმით, ფორმა წარმოდგენილია როგორც ერთი მთლიანი, თუ იგი წარმოდგენილია კონტურის დანაწევრებით მიღებული ცალკეული მონაკვეთების (პრიმიტივების) სახით. ფორმის წარმოდგენისა და აღწერის ზემოწარმოდგენილი კლასიფიკაცია მოცემულია 5.1 სურათზე.



სურ. 5.1

5.2. კონტურზე დამყარებული ფორმის წარმოდგენა და აღწერა

როგორც უკვე აღვნიშნეთ, კონტურზე დამყარებული ფორმის აღწერისას გამოიყენება მხოლოდ კონტურთან დაკავშირებული

ინფორმაცია. აქ განიხილება ორი მიდგომა: **უწყვეტი (გლობალური)** და **დისკრეტული (სტრუქტურული)**. უწყვეტი მიდგომისას კონტური განიხილება მთლიანობაში და არა დანაწევრებული სახით. შესაბამისად, ფორმის აღმწერ ნიშან-თვისებათა ვექტორში სწორედ ეს მთლიანობაა ასახული. ფორმის დისკრეტული აღწერისას კი კონტური გარკვეული კრიტერიუმის საფუძველზე ნაწევრდება ცალკეულ მონაკვეთებად, რომლებსაც **პრიმიტივებს** უწოდებენ, ხოლო თვით ფორმა კი ასეთი მიდგომისას წარმოიდგინება სიმბოლოების მიმდევრობის ან გრაფის (ან ხის) სახით.

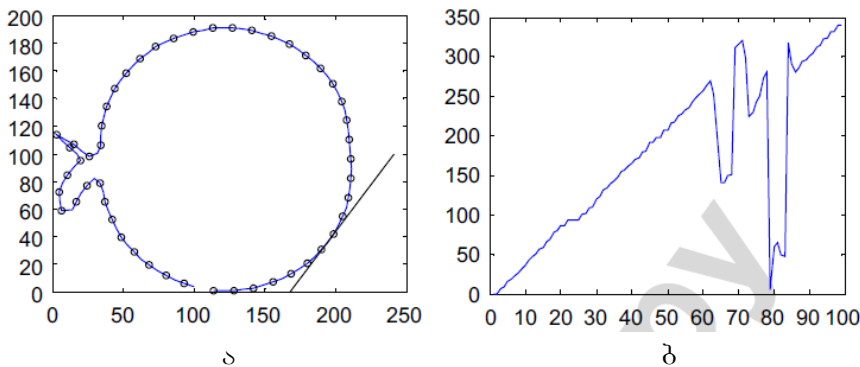
კონტურზე დამყარებული ფორმის აღწერისას საჭიროა დავიცვათ კომპრომისი აღწერის სიზუსტესა და ეფექტურობას შორის. იგულისხმება, რომ ერთი მხრივ ფორმა უნდა იყოს აღწერილი რაც შეიძლება ზუსტად, და ამავე დროს, ეს აღწერა უნდა იყოს რაც შეიძლება კომპაქტური (დაკონური) ფორმების შემდგომი ანალიზისა და მონაცემთა ბაზებში მოძიებისთვის.

ფორმის გლობალური წარმოდგენის მეთოდები ფორმას აღწერს ნიშან-თვისებათა მრავალგანზომილებიანი რიცხვითი ვექტორის საშუალებით კონტურის ინფორმაციაზე დაყრდნობით. ასეთი აღწერისათვის კონტური წარმოიდგინება ფორმის კონტურიდან აღებული სასრული რაოდენობის წერტილთა (ანათვლების) სიმრავლით. წერტილების შერჩევა შეიძლება სხვადასხვა საფუძველზე დაყრდნობით. მაგალითად, სიმძიმის ცენტრიდან დაშორების საფუძველზე. თუ კონტური წარმოდგენილია ერთმანეთისგან თანაბრად დაშორებული k რაოდენობის s_1, s_2, \dots, s_k ანათვალი წერტილების საშუალებით, მაშინ შესაძლებელია ამ კონტურის გარდაქმნა $f(s_1), f(s_2), \dots, f(s_k)$ რიცხვით მიმდევრობაში, სადაც $f(s_i)$, $i = \overline{1, k}$ წარმოადგენს მოცემული კონტურის k წერტილში აღებული რომელიმე ნიშან-თვისების რიცხვით მნიშვნელობას. ამგვარად, ფორმის კონ-

ტური შეიძლება აღიწეროს ნიშან-თვისებათა რიცხვითი ვექტორით

$$\mathbf{X} = (f(s_1), f(s_2), \dots, f(s_k)).$$

აღწერილი გარდაქმნის მაგალითი წარმოდგენილია 5.2, ა, ბ სურათზე, სადაც (ა)-ზე მოცემულია თანაბარანათვლებიანი კონტური, ხოლო (ბ)-ზე კი – რიცხვითი მიმდევრობა, რომლის თითოეული წევრი წარმოადგენს კონტურის თითოეულ ანათვალში აღებული მხების x ღერძის მიმართ დახრის კუთხის მნიშვნელობას. მიღებულ მიმდევრობაში რიცხვითი მნიშვნელობები მოთავსებულია ინტერვალში $(0 - 360)^\circ$, ხოლო კონტურზე მოძრაობისათვის საწყის წერტილად არჩეულია კონტურის უკიდურესად ქვედა წერტილი. ამასთან, კონტურზე მოძრაობა ხდება საათის ისრის ბრუნვის მიმართულებით.



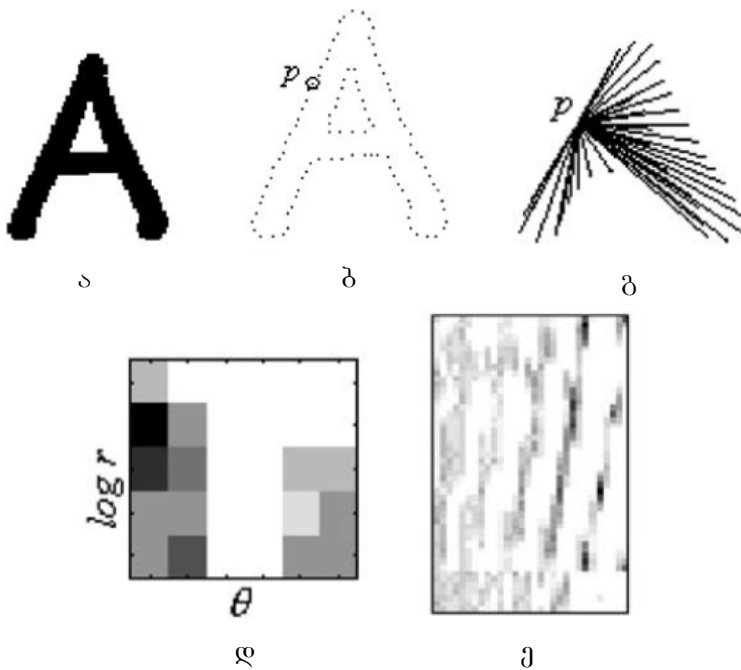
სურ. 5.2

განვიხილოთ მოკლედ ფორმის გლობალური წარმოდგენის ზოგიერთი მეთოდი.

ფორმის კონტექსტი წარმოადგენს ფორმის აღმწერს, რომელიც გამოყოფს ობიექტის ფორმისთვის დამახასიათებელ გლობალურ ნიშან-თვისებას კონტურის გარკვეულ წერტილებში. ფორმის კონტექსტის ანალიზი იწყება N ანათვალის ადებით ფორმაზე კიდეების გამოყოფის შედეგად მიღებული გამოსახულებიდან. განვიხილოთ მისი არსი 5.3,ა სურათზე წარმოდგენილი **A** სიმბოლოს ფორმის მაგალითზე. ჯერ სიმბოლოს გამოსახულების მიმართ გამოვიყენოთ კიდის გამოყოფის ოპერატორი, რაც მოგვცემს მოცემული ფორმის კიდის გამოსახულებას – კონტურს (სურ. 5.3,ბ). მიღებული კონტურის რომელიმე p წერტილში ფორმის კონტექსტის საპოვნელად საჭიროა ვიპოვოთ მოცემული p ანათვალის წერტილიდან ვექტორები კონტურის ყველა სხვა დანარჩენი ანათვლების წერტილებამდე. შედეგად მივიღებთ ვექტორების სიმრავლეს p წერტილში (სურ. 5.3,გ). აღნიშნული ვექტორები გამოხატავს მთლიანი ფორმის ხილულობას მოცემული p წერტილის მიმართ. მიღებული ვექტორების r სიგრძისა და Q ორიენტაციის დაკვანტვით ხდება ჰისტოგრამის რუკის მიღება, რომელიც გამოიყენება p წერტილის წარმოსადგენად (სურ. 5.3,დ). ასე ხდება კონტურის თითოეული ანათვალის წერტილისთვის. მოცემული ფორმის კონტექსტის მისაღებად თითოეული ანათვალის შესაბამისი ჰისტოგრამა გლუვდება, ხოლო შემდეგ კი ხდება მათი გაერთიანება, რის შედეგადაც მიიღება **A** ფორმის კონტექსტი (სურ. 5.3,ე), რომელშიც სტრიქონების რაოდენობა ანათვლების რაოდენობის ტოლია.

ფორმის მოხაზულობა ფორმას წარმოადგენს სასაზღვრო წერტილებიდან (კოორდინატებიდან) წარმოებული ერთგანზომილებიანი ფუნქციის საშუალებით. პრაქტიკაში ფორმის მოხაზულობის რაოდენობრივი აღწერისათვის გამოიყენება რამდე-

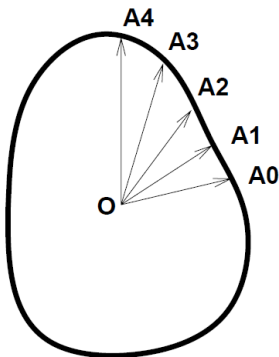
ნიმე აღმწერი, როგორცაა მაგალითად, სიმძიმის ცენტრის პროფილი, მანძილი სიმძიმის ცენტრიდან, სიმრუდე და სხვ.



სურ. 5.3

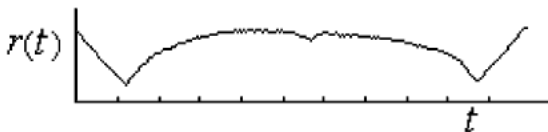
ფორმის აღმწერად სიმძიმის ცენტრიდან მანძილის გამოყენების შემთხვევაში ერთგანზომილებიანი ფუნქციის მნიშვნელობებად აიღება მანძილი ფორმის სიმძიმის ცენტრიდან კონტურის წერტილებამდე. კონტურზე წერტილების არჩევა ხდება ისეთნაირად, რომ მათ შორის რადიალური კუთხეები ან ინტერვალები კონტურის გასწვრივ ერთმანეთის ტოლი იყოს (სურ. 5.4). ამგვარად, რაც უფრო დიდია ასეთი წერტილების რაოდენობა (სიმრავლე) კონტურზე, მით უფრო ზუსტია კონ-

ტურის აღწერაც. მაგრამ ეს თავის მხრივ იწვევს გამოთვლების ფასის ზრდას.

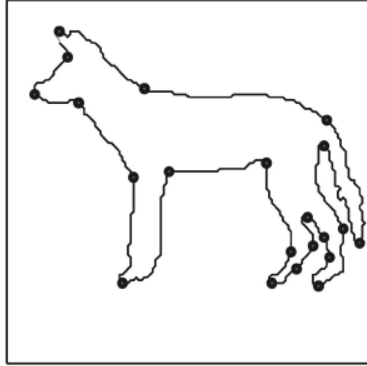


სურ. 5.4

ასეთი ნიშან-თვისების გამოყენების მაგალითი მოცემულია 5.5 სურათზე, სადაც $r(t)$ აღნიშნულია მანძილი ფორმის სიმძიმის ცენტრიდან კონტურის წერტილებამდე. სიმრუდის, როგორც ფორმის აღმწერის გამოყენების შემთხვევა მოცემულია 5.6 სურათზე, სადაც ძალის ფორმა წარმოდგენილია კონტურის იმ წერტილების (ანათვლების) საშუალებით, რომლებშიც სიმრუდეს მაღალი მნიშვნელობები აქვს.



სურ. 5.5



სურ. 5.6

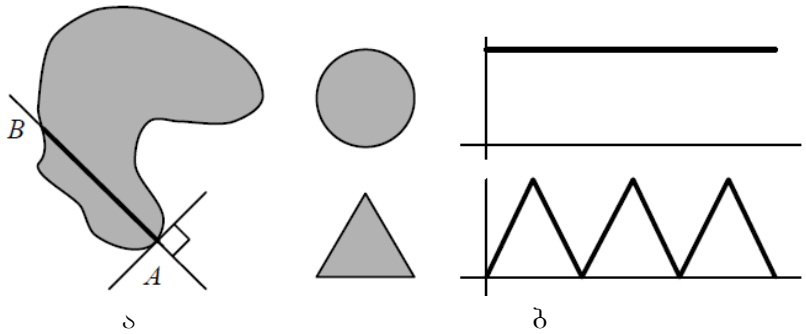
ფორმის კონტურის მოხაზულობის აღმწერ კიდევ ერთ ფუნქციას წარმოადგენს კონტურის ნორმალის მიმართულებით აღებული მანძილების მიმდევრობა. კონტურის ნებისმიერ წერტილში ნორმალის მიმართულებით აღებული მანძილი გამოითვლება როგორც სიგრძის ფუნქცია. 5.7,ა სურათზე წარმოდგენილ შემთხვევაში უმოკლესი მანძილი კონტურის ნებისმიერი A წერტილიდან მის მოპირდაპირედ მთავსებულ B წერტილამდე განისაზღვრება კონტურის A წერტილში მხების მიმართ გავლებული ნორმალის მიმართულებით (უნდა აღინიშნოს, რომ ეს მიმართება წერტილებს შორის სიმეტრიული არ არის). 5.7,ბ სურათზე წარმოდგენილია ფორმის მოხაზულობები წრეწირისა და სამკუთხედისთვის. მოხაზულობა მგრძობიარეა ხმაურის მიმართ. ამ უკანასკნელის შესამცირებლად ახდენენ მოხაზულობის დაგლუვებას ან მოხაზულობის ნიშან-თვისებას იყენებენ უკვე დაგლუვებული კონტურისათვის.

კონტურის აღმწერი წერტილების რაოდენობის (განზომილების) შემცირების (შესაბამისად, გამოთვლების ფასის) ერთ-ერთ გზას წარმოადგენს ფორმის აღმწერად **სასაზღვრო მომენტების** გამოიყენება. თუ ფორმის საზღვარს წარმოვადგ-

ერთ როგორც ფორმის მოხაზულობას $z(i)$, მაშინ r -ული m_r მომენტი და μ_r ცენტრალური მომენტის შეფასება შეიძლება ასე ჩაიწეროს

$$m_r = \frac{1}{N} \sum_{i=1}^N [z(i)]^r \quad \text{და} \quad \mu_r = \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^r,$$

სადაც, N საზღვრის წერტილების რაოდენობაა.



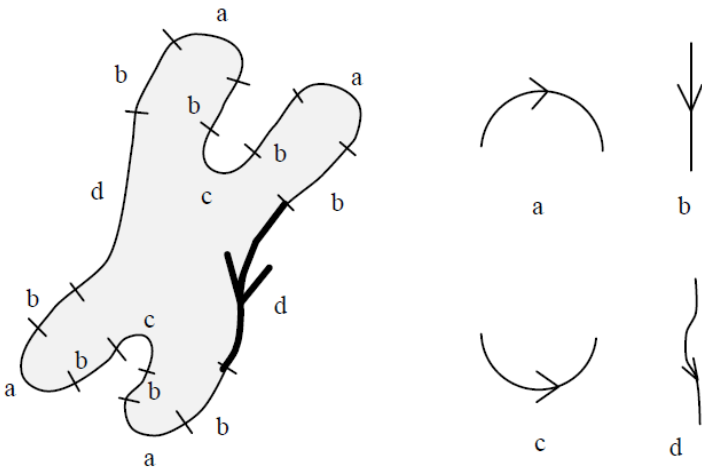
სურ. 5.7

იმისათვის, რომ ფორმის აღწერა იყოს ინვარიანტული გადაადგილების, ბრუნვის და მასშტაბირების მიმართ, იყენებენ ნორმალიზებულ ცენტრალურ მომენტებს:

$$\overline{m}_r = m_r / (\mu_2)^{r/2} \quad \text{და} \quad \overline{\mu}_r = \mu_r / (\mu_2)^{r/2}$$

ფორმის სტრუქტურული აღწერის მეთოდებში ფორმა, როგორც აღვნიშნეთ, დანაწევრებულია მისი კონტურის პრიმიტივებად (მონაკვეთებად), რომლებიც თავის მხრივ შეიძლება წარმოდგენილი იქნეს მეორე რიგის პოლინომებით, როგორი-

ცაა მაგალითად წრიული, ელიფსური ან პარაბოლური მონაკვეთები. 5.8 სურათზე წარმოდგენილია შემთხვევა, როცა ქრომოსომების ამოცნობისთვის გამოყენებულია ფორმის სინტაქსური აღწერა. როგორც სურათიდან ჩანს, ფორმის კონტური დანაწევრებულია მონაკვეთებად, რომლებიც წარმოადგენს დიდი სიმრუდის მქონე ამოზნექილ მონაკვეთებს, დიდი სიმრუდის მქონე ჩაზნექილ მონაკვეთებს, წრფეებს და ა.შ.



სურ. 5.8

ფორმის აღწერის სტრუქტურული მეთოდები ერთმანეთისგან განსხვავდება პრიმიტივების შერჩევითა და მათ ორგანიზაციით. საზღვრის დანაწევრების შედეგი კი მოიცემა ზოგადი სახით წარმოდგენილი სიმბოლოების

$$S = s_1, s_2, \dots, s_n$$

მიმდევრობით, სადაც s_i შეიძლება იყოს მაგალითად, ფრიმანის ჯაჭვური კოდის ელემენტი, მრავალკუთხედის გვერდი, B-

სპლაინი (უბან-უბან პოლინომიალური მრუდი), კვადრატული (მეორე რიგის) რკალი და ა.შ. ასევე s_i შეიძლება წარმოადგენდეს რომელიმე ნიშან-თვისებას, როგორცაა მაგალითად, სიგრძე, საშუალო სიმრუდე, მაქსიმალური სიმრუდე და ა.შ.

ფორმის აღწერები, რომლებიც გამოიყენება სტრუქტურული მიდგომისას, მოცემულია 5.1 სურათზე.

ფორმის სტრუქტურული აღწერის მეთოდების დადებით მხარეს წარმოადგენს უნარი, მართონ ოკლუზიის (დაფარვის) პრობლემა, რასაც ადგილი აქვს დინამიკურ სცენაზე ობიექტების მიერ ერთმანეთის ნაწილობრივი ან მთლიანი გადაფარვის შემთხვევაში. თუმცა ამავე დროს ამ მიდგომას უარყოფითი მხარეებიც გააჩნია. კერძოდ, იმის გამო, რომ არ არსებობს ობიექტისა და ფორმის ფორმალური განსაზღვრება, შესაბამისად უცნობია მოცემული ფორმის აღწერისათვის საჭირო პრიმიტივების რაოდენობა. აქედან გამომდინარე, აღნიშნული მიდგომის გამოყენების წარმატებულობა განპირობებულია მონაცემთა ბაზაში არსებული ფორმის კონტურის ნიშან-თვისებების შესახებ აპრიორული ცოდნით. ასევე უარყოფით მხარეს წარმოადგენს ასეთი მეთოდებისთვის დამახასიათებელი მაღალი გამოთვლითი სირთულე და მგრძობიარობა ხმაურის მიმართ.

5.3. არეზე დამყარებული ფორმის წარმოდგენა და აღწერა

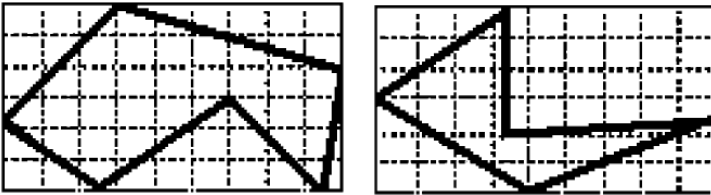
როგორც ზემოთ აღვნიშნეთ, არეზე დამყარებული ფორმის წარმოდგენისა და აღწერის მეთოდები ფორმის წარმოდგენისთვის კონტურთან ერთად იყენებენ არის ყველა პიქსელთან დაკავშირებულ ინფორმაციას. ისევე როგორც კონტურზე დამყარებული მეთოდები, არეზე დამყარებული ფორმის აღწერის მეთოდებიც იყოფა გლობალურ და სტრუქტურულ მეთოდებად,

იმისდა მიხედვით – ახდენენ ისინი ფორმის დანაწევრებას ცალკეულ ქვენაწილებად, თუ არა.

გლობალური მიდგომისას ფორმა განიხილება როგორც ერთი მთლიანი. ასეთ მიდგომაზე დამყარებული მეთოდები ფორმის აღწერისათვის იყენებენ არეში პიქსელების განაწილებას, რომლებიც თავის მხრივ ნაკლებად განიცდის ხმაურისა და ცვლილებების გავლენას. შესაბამისად, ისინი უფრო ადვილად ართმევენ თავს მნიშვნელოვანი დეფექტების მქონე ფორმების აღწერას, განსხვავებით კონტურზე დამყარებული მეთოდებისგან. ფორმის ასეთი წარმოდგენის შედეგია ნიშან-თვისებათა რიცხვითი ვექტორი, რომელსაც იყენებენ ფორმის აღწერისათვის. არეზე დამყარებული ფორმის გლობალური აღმწერები მოცემულია 5.1 სურათზე. განვიხილოთ მოკლედ ზოგიერთი ეს გლობალური მეთოდი.

ფორმის აღწერის პოპულარულ მეთოდებს წარმოადგენს სტატისტიკური **მომენტების** (moments) მეთოდები, რომელთა საშუალებითაც განისაზღვრება არეში პიქსელების განაწილება. დაბალი რიგის მომენტები ანუ მომენტების ინვარიანტები (მათ გეომეტრიულ მომენტებსაც უწოდებენ) ასახავს არეში პიქსელების განაწილებასთან დაკავშირებულ ფიზიკურ არსს. რაც შეეხება მაღალი რიგის მომენტების განსაზღვრას, აქ სიძნელეს წარმოადგენს მათი ფიზიკური ინტერპრეტაციის პრობლემა. **ბადური** (grid) მეთოდები ფორმის აღწერისათვის იყენებს მართკუთხა ბადეს, რომელიც ფორმას ედება ზემოდან. ამის შემდეგ ხდება ბადის სკანირება მარცხნიდან მარჯვნივ და ზემოდან ქვემოთ. შედეგად მიიღება ბინარული მატრიცა, რომელშიც ერთიანები შეესაბამება ბადის იმ უჯრებს, რომლებიც ფორმითაა დაფარული, ხოლო ნულიანები კი იმ უჯრებს, რომლებიც არ არის ფორმით დაფარული. საბოლოოდ ფორმა წარმოიდგინება ბინარული ნიშან-თვისებათა ვექტორით. ასეთი

მეთოდების დადებითი მხარეა წარმოდგენის სიმარტივე, ხოლო უარყოფითი კი – მგრძობიარობა ხმაურის მიმართ. მაგრამ ისინი არ არის ინვარიანტული გადაადგილების, ბრუნვის და მასშტაბირების მიმართ. 5.9 სურათზე მოცემულია ორი სხვადასხვა ფორმის წარმოდგენა მართკუთხა ბადის საშუალებით.



სურ. 5.9

წარმოდგენილ ფორმებზე ბადური აღმწერების გამოყენების შედეგად მიღებულ ბინარულ მატრიცებს შესაბამისად ასეთი სახე აქვს:

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

ბადური მეთოდებისგან განსხვავებით **ფორმის მატრიცის** (shape matrix) მეთოდებში ფორმაზე კვადრატული ბადის ნაცვლად ხდება კონცენტრული წრეწირებისგან შედგენილი პოლარული ბადის და რადიალური წრფეების, ცენტრით სხეულის სიმძიმის ცენტრში, ზედდება (სურ. 5.10). განვიხილოთ პოლარული ბადის აგების პროცედურა 5.10 სურათზე წარმოდგენილი ფორმისთვის, სადაც G აღნიშნულია ფორმის სიმძიმის ცენტრი,

ხოლო GA კი ფორმის მაქსიმალური რადიუსი. შემოვხაზოთ G წერტილიდან ერთმანეთისგან თანაბრად დაშორებული n კონცენტრული წრეწირი. შემდეგ დაწყებული GA -დან საათის მოძრაობის საწინააღმდეგო მიმართულებით G წერტილიდან გავატაროთ რადიალური წრფეები, რომლებიც თითოეულ წრეწირს დაყოფს m რაოდენობის თანაბარ რკალებად. შედეგად მივიღებთ პოლარულ კოორდინატებში წარმოდგენილ ფორმის მატრიცას.



სურ. 5.10

ფორმის აღმწერ ნიშან-თვისებათა ბინარული ვექტორის მნიშვნელობები მიიღება კონცენტრული წრეწირებისა და რადიალური წრფეების გადაკვეთის წერტილებიდან. კერძოდ, წრეწირები შეესაბამება ფორმის მატრიცის სვეტებს, ხოლო რადიალური წრფეები კი – სტრიქონებს. 5.10 სურათზე მოცემული ფორმისთვის პოლარული ბადის აგების შედეგად ვღებულობთ: $n = 5$ სტრიქონს და $m = 12$ სვეტს. შესაბამის ბინარულ მატრიცას კი ასეთი სახე აქვს:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

აღსანიშნავია, რომ ფორმის მატრიცა ინვარიანტულია გადატანის, ბრუნვის და მასშტაბირების მიმართ. მაგრამ, იმის გამო, რომ ფორმის მატრიცა წარმოადგენს ფორმის გაიშვიათებულ ამონაკრებს, ის ადვილად განიცდის ხმაურის გავლენას. ამის გამო ფორმის მატრიცის მეთოდები ნაკლებად საიმედოა.

კონტურზე დამყარებული სტრუქტურული მეთოდების მსგავსად, არეზე დამყარებულ ფორმის წარმოდგენის მიდგომაშიც არსებობს სტრუქტურული მეთოდები. აქაც გამოყოფილი არე ნაწევრდება ქვეარეებად, რომლებიც შემდეგ გამოიყენება ფორმის წარმოდგენისა და აღწერისათვის. ასეთი მეთოდები მოცემულია 5.1 სურათზე.

მოკლედ განვიხილოთ **ამოზნექილი გარსის (convex hull)** მეთოდის არსი.

განსაზღვრება 5.1. *R არეს ეწოდება ამოზნექილი მხოლოდ და მხოლოდ მაშინ, თუ ამ არის ნებისმიერი ორი $x_1, x_2 \in R$ წერტილის შემავრთებელი მონაკვეთი მთლიანად ამ არის შიგნით მდებარეობს.*

R არის **ამოზნექილი გარსი** წარმოადგენს იმ უმცირეს ამოზნექილ H არეს, რომელიც აკმაყოფილებს პირობას: $R \subset H$. 5.11 სურათზე მოცემული ობიექტისთვის მასზე შემოხაზული კონტური წარმოადგენს ამ ობიექტის ამოზნექილ გარსს.

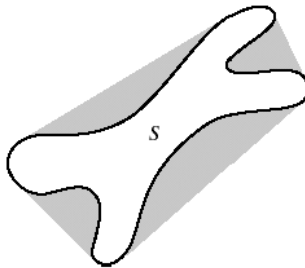


სურ. 5.11

ასეთი მიდგომისას ფორმა წარმოდგინება ამოზნექილი გარსებისგან შედგენილი სერიის საშუალებით. 5.12 სურათზე მოცემული S არისტვის ამოზნექილი გარსი განისაზღვრება როგორც ამ არის შემცველი უმცირესი ამოზნექილი სიმრავლე. ასეთი მიდგომისას ხმაურის გაელენის შესამცირებლად ფორმის დანაწევრებამდე ახდენენ კონტურის დაგლუვებას.

განსაზღვრება 5.2. *სხვაობას $H - R$ ეწოდება R არის ამოზნექილობის დეფიციტი.*

5.12 სურათზე წარმოდგენილია S არე და მისი ამოზნექილი გარსი, ხოლო მუქი ფერით კი აღნიშნულია ამოზნექილობის დეფიციტი.



სურ. 5.12

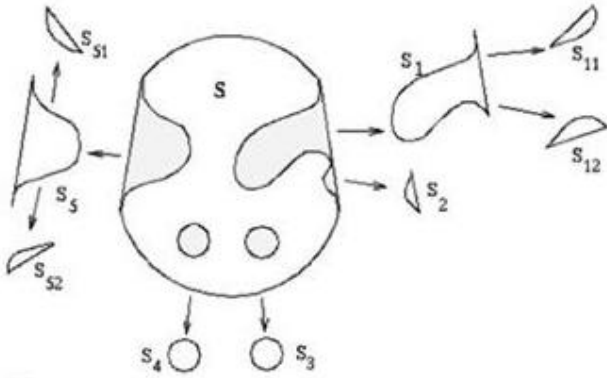
ამოზნექილი გარსის გამოყოფის პროცესი შეიძლება განვიხილოთ როგორც ამოზნექილობის დეფიციტების პოვნის ამოცანა არის საზღვრის გასწვრივ (სურ. 5.12). შესაბამისად, ფორმა შეიძლება წარმოდგენილ იქნეს რეკურსიული პროცედურის შედეგად მიღებული ე.წ. **ჩაზნექილობის ხის** სახით. კერძოდ, პირველად ამოზნექილობის დეფიციტების საშუალებით ხდება მოცემული ობიექტის ამოზნექილი გარსის მიღება. შემდეგ მიღებულ ამოზნექილობის დეფიციტებში ისევ იძებნება ამოზნექილი გარსები და ამოზნექილობის დეფიციტები. შემდეგ მიღებული ამოზნექილობის დეფიციტებით ისევ იძებნება ამოზნექილი გარსები და ამოზნექილობის დეფიციტები და ა.შ. მანამ, სანამ მიღებული ამოზნექილობის დეფიციტები შეიცავს ამოზნექილობებს. ეს პროცედურა წარმოდგენილია 5.13 სურათზე, სადაც ობიექტის ამოზნექილი გარსი და მისი ჩაზნექილობები წარმოდგენილია 5.13,ა სურათზე, ხოლო ამოზნექილი გარსი წარმოდგენილი ჩაზნექილობის ხის სახით მოცემულია 5.13,ბ სურათზე. ამასთან, თითოეული ჩაზნექილობა შეიძლება აღიწეროს მისი ფართობის, ქორდის სიგრძის (ამოჭრილი ჩაზნექილობის შემაერთებელი წირის), მაქსიმალური სიმრუდის და სხვა მახასიათებლების საშუალებით.

5.4. ფორმის მარტივი აღმწერები

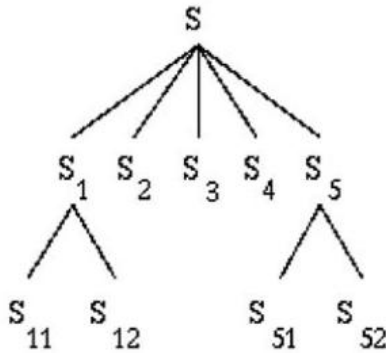
საზოგადოდ, გამოსახულებაზე გამოყოფილი არეების ყველაზე მეტად გავრცელებული მახასიათებელი ნიშან-თვისებები შემდეგია:

- **გეომეტრიული** (მაგალითად, არის ფართობი, სიმძიმის ცენტრი, ექსტრემალური წერტილები);

- ფორმის (მაგალითად, სიმრგვალე, ეილერის რიცხვი, მომენტები);
- ინტენსივობის (მაგალითად, არის საშუალო ინტენსივობა, სხვადასხვა ტექსტურული სტატისტიკები).



ა



ბ

სურ. 5.13

5.1 სურათზე მოცემულია ფორმის აღმწერების ჩამონათვალი, რომლებიც გამოიყენება როგორც კონტურზე, ასევე არეზე დამყარებული ფორმის გლობალური აღწერისას. მათ შორის არის ე.წ. **ფორმის მარტივი აღმწერებიც**, რომლებსაც შეუძლია ერთმანეთისგან განასხვავოს მხოლოდ მკვეთრად (მნიშვნელოვნად) განსხვავებული ფორმები. მართალია, ფორმის მარტივი აღმწერები კომპაქტურებია, მაგრამ ამავე დროს არ გამოირჩევა მაღალი სიზუსტით. სწორედ ამის გამო მათ იყენებენ სხვა აღმწერებთან ერთად კომბინაციაში. განვიხილოთ ზოგიერთი მარტივი ფორმის აღმწერი ბინარული გამოსახულების მაგალითზე.

არის ფართობი:

$$A = \sum_{(i,j) \in R} 1$$

სადაც R -ით აღნიშნულია არის პიქსელთა სიმრავლე. როგორც ვხედავთ, არის ფართობი მასში შემავალი პიქსელების რაოდენობის ტოლია.

არის სიმბიმის ცენტრი (centroid) - (\bar{i}, \bar{j})

$$\bar{i} = \frac{1}{A} \sum_{(i,j) \in R} i$$

$$\bar{j} = \frac{1}{A} \sum_{(i,j) \in R} j$$

ანუ სიმბიმის ცენტრის კოორდინატები (\bar{i}, \bar{j}) ტოლია R არის პიქსელების კოორდინატების საშუალო მნიშვნელობების. უნდა

აღინიშნოს, რომ თუმცა წყვილები $(i, j) \in R$ მთელ რიცხვებს წარმოადგენს, მაგრამ (\bar{i}, \bar{j}) კოორდინატები ზოგადად არ არის მთელი რიცხვები. ამის გამო, საკმარისია, რომ სიმძიმის ცენტრის კოორდინატები გამოთვლილ იქნეს მძიმის შემდეგ ერთი ნიშნის სიზუსტით.

არის პერიმეტრის სიგრძე:

R არის პერიმეტრის სიგრძე P წარმოადგენს არის გლობალურ თვისებას. მარტივად, იმ არის პერიმეტრი, რომელსაც არ აქვს ხვრელები, შეიძლება განსაზღვრული იქნეს მისი სასაზღვრო პიქსელების სიმრავლის საშუალებით. R არის პიქსელს ეწოდება სასაზღვრო, თუ მისი ერთ-ერთი მეზობელი არ ეკუთვნის R -ს. იმ შემთხვევაში, როცა პიქსელის არისადმი ან ფონისადმი მიკუთვნება განისაზღვრება 8-ბმულობით, მაშინ პერიმეტრის პიქსელებისგან მიღებული სიმრავლე იქნება 4-ბმული. და პირიქით, თუ პიქსელის არისადმი ან ფონისადმი მიკუთვნება განისაზღვრება 4-ბმულობით, მაშინ პერიმეტრის პიქსელების სიმრავლე წარმოადგენს 8-ბმულს. აქედან გამომდინარე, R არისთვის ვლებულობთ 4-ბმული P_4 პერიმეტრისა და 8-ბმული P_8 პერიმეტრის შემდეგ განსაზღვრებებს:

$$P_4 = \{(i, j) \in R | N_8(i, j) - R \neq 0\};$$

$$P_8 = \{(i, j) \in R | N_4(i, j) - R \neq 0\}.$$

P პერიმეტრის $|P|$ სიგრძის გამოსათვლელად საჭიროა P პერიმეტრის პიქსელების მოწესრიგება ასეთი სახის მიმდევრობად – $P = \langle (i_0, j_0), \dots, (i_{K-1}, j_{K-1}) \rangle$, რომელშიც პიქსელების ყოველი შემდეგი მომდევნო წყვილი ერთმანეთის მეზობლებია (მათ შორის პირველი და ბოლო პიქსელები). მაშინ პერიმეტრის სიგრძე $|P|$ ასე განისაზღვრება:

$$|P| = |\{k|(i_{k+1}, j_{k+1}) \in N_4(i_k, j_k)\}| + \sqrt{2}|\{k|(i_{k+1}, j_{k+1}) \in N_8(i_k, j_k) - N_4(i_k, j_k)\}|,$$

სადაც $k + 1$ გამოითვლება მოდულით K (K პიქსელთა მიმდევრობის სიგრძის ტოლია). ამგვარად, ორი ვერტიკალურად ან ჰორიზონტალურად მომიჯნავე პიქსელი მოგვცემს პერიმეტრის სიგრძის ზრდას 1-ით, ხოლო დიაგონალურად მომიჯნავე ორი პიქსელი კი – პერიმეტრის სიგრძის ზრდას დაახლოებით 1,4 – ით.

არის A ფართობისა და P პერიმეტრის გამოთვლა საშუალებას გვაძლევს, რომ გამოითვალოს არის კომპაქტურობა (compactness)

კომპაქტურობა (1):

$$C_1 = \frac{|P|^2}{A}.$$

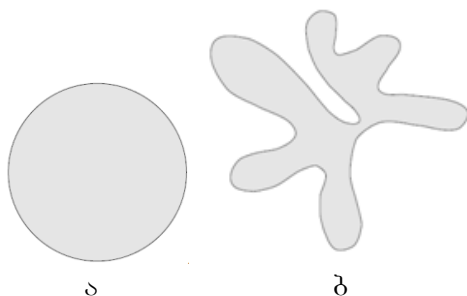
მაგრამ, სამწუხაროდ, ციფრული სახით წარმოდგენილი ფიგურებისთვის სიდიდე $|P|^2/A$ უმცირეს მნიშვნელობას ღებულობს

არა წრეწირისთვის (როგორც ეს უწყვეტი სახით წარმოდგენილი ბრტყელი ფიგურების შემთხვევაშია), არამედ რვაკუთხედებისა და რომბებისთვის. ამის გამო კომპაქტურობის ნიშანთვისების განსაზღვრისთვის შემოტანილი იქნა მეორე ფორმულა

კომპაქტურობა (2):

$$C_2 = \frac{\mu_R}{\sigma_R},$$

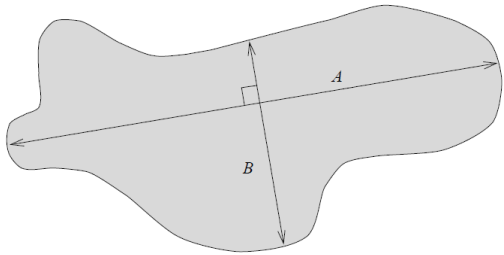
სადაც μ_R და σ_R ტოლია შესაბამისად არის სიმძიმის ცენტრიდან არის სასაზღვრო წერტილებამდე გადახრების სიდიდის საშუალო მნიშვნელობისა და საშუალოკვადრატული გადახრისა შესაბამისად. ევკლიდურ სივრცეში ყველაზე უფრო კომპაქტურ არეს წარმოადგენს წრე. 5.14 სურათზე მოცემულია (ა) კომპაქტური და (ბ) არაკომპაქტური არეების მაგალითები.



სურ. 5.14

ექსცენტრისიტეტი (eccentricity):

ექსცენტრისიტეტი წარმოადგენს ერთმანეთის მიმართ მართობულად განლაგებული მაქსიმალური სიგრძის ქორდების ფარდობას (ან სხვა სიტყვებით, ობიექტის დიდი ღერძის პატარა ღერძთან ფარდობას). 5.15 სურათზე წარმოდგენილ შემთხვევაში ასეთი ქორდებია A და B .



სურ. 5.15

საშუალო რადიალური მანძილი:

$$\mu_R = \frac{1}{K} \sum_{k=0}^{K-1} \|(i_k, j_k) - (\bar{i}, \bar{j})\|,$$

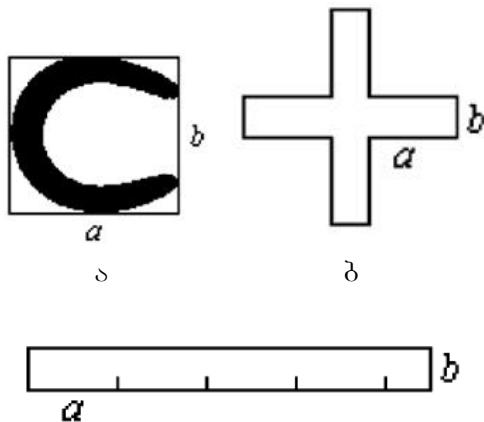
რადიალური მანძილის საშუალოკვადრატული გადახრა:

$$\sigma_R = \frac{1}{K} \sum_{k=0}^{K-1} [|\|(i_k, j_k) - (\bar{i}, \bar{j})\| - \mu_R|]^{1/2},$$

სადაც პიქსელთა სიმრავლე (i_k, j_k) , $k = 0, 1, \dots, K - 1$ წარმოადგენს განსახილველი R არის P პერიმეტრის ელემენტების კრებულს. აღსანიშნავია, რომ კომპაქტურობის C_2 მნიშვნელობა მონოტონურად იზრდება ფიგურის ციფრული სახის წრეხაზთან მიახლოებასთან ერთად და ამჟღავნებს მსგავს ქცევას როგორც ციფრულ, ასევე უწყვეტ სახეში წარმოდგენილი ფიგურებისთვის.

განვიხილოთ კონტურზე დამყარებული ფორმის მარტივი აღმწერების, ექსცენტრისიტეტისა და კომპაქტურობის გამოყენების მაგალითი 5.16 სურათზე წარმოდგენილი ობიექტებისთვის. 5.16,ა სურათზე მოცემული ობიექტის ფორმის ექსცენტრისიტეტი თითქმის 1-ის ტოლია ($a \cong b$). მაგრამ იმის გამო, რომ ობიექტს წაგრძელებული ფორმა არა აქვს, ეს ნიშნისეხება არ გამოდგება როგორც მოცემული ობიექტის ფორმის აღმწერი. კომპაქტურობა ამ შემთხვევაში უფრო შესაფერისი აღმწერი იქნება. რაც შეეხება 5.16,ბ,გ სურათზე წარმოდგენილ ობიექტებს, მიუხედავად იმისა, რომ ისინი განსხვავებულ ფორმებს წარმოადგენს, მათი კომპაქტურობის მნიშვნელობები ერთნაირია ($a = 2b$). ამიტომ კომპაქტურობა, როგორც ფორმის აღმწერი, ასეთ შემთხვევაში არ გამოდგება. შესაბამისად, მათ-

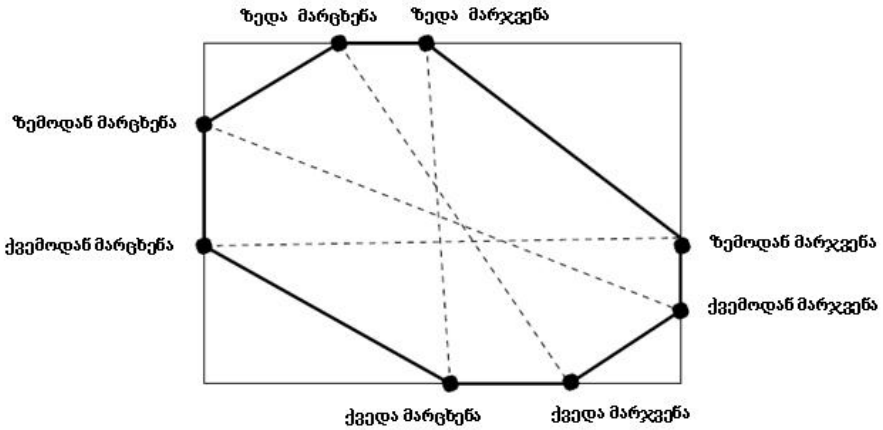
თვის ექსცენტრისიტეტის, როგორც ნიშან-თვისების გამოყენება უფრო შესაფერისი იქნება.



ბ
სურ. 5.16

ჩახაზული სწორკუთხედი და ექსტრემალური წერტილები:

გამოსახულების ანალიზისას ხშირად ძალზე მნიშვნელოვანია და სასურველია ვიცოდეთ რომელიმე არის მიახლოებითი ადგილმდებარეობა გამოსახულებაზე. ამ უკანასკნელთანაა დაკავშირებული აღმწერი მართკუთხედის (bounding box) ცნება. ის წარმოადგენს მართკუთხედს, რომელიც შემოხაზულია არეზე და ეხება მის კიდურა წერტილებს ზემოდან, ქვემოდან, მარცხნიდან და მარჯვნიდან. როგორც 5.17 სურათიდან ჩანს, არეს შეიძლება ჰქონდეს 8 განსხვავებული ექსტრემალური პიქსელი: ზედა მარჯვენა, ზედა მარცხენა, ზემოდან მარცხენა, ქვემოდან მარცხენა, ზემოდან მარჯვენა, ქვემოდან მარჯვენა, ქვედა მარცხენა, ქვედა მარჯვენა.



სურ. 5.17

ექსტრემალური წერტილისთვის დამახასიათებელია ის, რომ მის კოორდინატას აქვს ზღვრული მნიშვნელობა სტრიქონის ან სვეტის მიხედვით. აქ "ზღვრულის" ქვეშ იგულისხმება კოორდინატის მაქსიმალური ან მინიმალური მნიშვნელობა განსახილველი არის ყველა პიქსელისათვის. ექსტრემალური პიქსელები განლაგებულია ჩახაზული მართკუთხედის გვერდებზე. ისინი ქმნის ურთიერთსაწინააღმდეგო მხარეებზე განლაგებული წერტილების წყვილებს: ზედა მარცხნიდან და ქვედა მარჯვნიდან; ზედა მარჯვნიდან და ქვედა მარცხნიდან; მარჯვენა ზემოდან და მარცხენა ქვემოდან; მარჯვენა ქვემოდან და მარცხენა ზემოდან. თითოეული ეს წყვილი განსაზღვრავს მართკუთხედით შემოვლებული ფიგურის ექსტრემალური წერტილების ღერძებს (სურათზე ნაჩვენებია პუნქტირით). გამოსადეგობის თვალსაზრისით ამ ღერძების თვისებებიდან არჩეულია მათი სიგრძე და მიმართულება.

იმის გამო, რომ აღნიშნული ექსტრემალური წერტილები მიღებულია უწყვეტი გამოსახულების ციფრულად გარდაქმნის შედეგად, ამიტომ ევკლიდური მანძილის ფორმულით გამოთვ-

ლილი ღერძის სიგრძე ცოტათი უფრო ნაკლები იქნება, განსხვავებით უწყვეტი შემთხვევისგან. განვიხილოთ, მაგალითად, მანძილის გამოთვლა ორ ჰორიზონტალურად მომიჯნავე პიქსელს შორის გამოსახულებაზე. ამ შემთხვევაში მანძილი მარცხენა პიქსელის მარცხენა კიდესა და მარჯვენა პიქსელის მარჯვენა კიდეს შორის 2-ის ტოლია, ხოლო მანძილი ამ პიქსელების ცენტრებს შორის კი 1-ის. ციფრულ გამოსახულებაზე მანძილის გამოთვლისას შესაძლებელია მხედველობაში მიღებულ იქნეს ეკვიდური მანძილის შესწორება. ეს ნაზრდი დამოკიდებულია ღერძის θ ორიენტაციაზე და ტოლია

$$Q(\theta) = \begin{cases} \frac{1}{|\cos \theta|}: & |\theta| < 45^\circ \\ \frac{1}{|\sin \theta|}: & |\theta| > 45^\circ \end{cases}$$

ამ შესწორების გათვალისწინებით ექსტრემალურ (i_1, j_1) და (i_2, j_2) წერტილებს შორის ექსტრემალური ღერძის სიგრძე ტოლი იქნება

$$D = \sqrt{(i_2 - i_1)^2 + (j_2 - j_1)^2} + Q(\theta).$$

გამოსახულებაზე არის ფორმის აღწერისათვის სშირად იყენებენ მეორე რიგის მომენტებს. კერძოდ:

მეორე რიგის ცენტრალური მომენტი სტრიქონზე:

$$\mu_{ii} = \frac{1}{A} \sum_{(i,j) \in R} (i - \bar{i})^2$$

მეორე რიგის შერეული ცენტრალური მომენტი:

$$\mu_{ij} = \frac{1}{A} \sum_{(i,j) \in R} (i - \bar{i})(j - \bar{j})$$

მეორე რიგის ცენტრალური მომენტი სვეტზე:

$$\mu_{jj} = \frac{1}{A} \sum_{(i,j) \in R} (j - \bar{j})^2$$

მეორე რიგის ცენტრალური მომენტები სტრიქონზე და სვეტზე - μ_{ii} და μ_{jj} შეიძლება გამოყენებულ იქნეს როგორც არის პიქსელების კოორდინატების გადახრის ზომა სიმძიმის ცენტრის (ამიტომაც მათ ცენტრალურ მომენტებს უწოდებენ) კოორდინატებიდან შესაბამისად სტრიქონის და სვეტის მიმართულებით. რაც შეეხება μ_{ij} შერეულ ცენტრალურ მომენტს, ის ითვალისწინებს პიქსელის კოორდინატების გადახრას სიმძიმის ცენტრის კოორდინატების მიმართ ორივე მიმართულებით. აღნიშნულ სიდიდეებს ხშირად იყენებენ როგორც არის ფორმის მარტივ მახასიათებლებს. ისინი ინვარიანტულია გადაადგილების და მასშტაბირების მიმართ და მათი გამოყენება შეიძლება ნებისმიერი ფორმის არეებისთვის.

5.5. ობიექტების ამოცნობა ფორმაზე დაყრდნობით

ობიექტის ფორმის ამოცნობა ერთ-ერთ მნიშვნელოვან როლს ასრულებს გამოსახულებათა ანალიზის ამოცანაში. აქ განიხილება ორი მიდგომა. ერთი, რომლის მიხედვითაც ფორმა განიხილება როგორც ერთი მთლიანი სტრუქტურა და გადაწყვეტი-

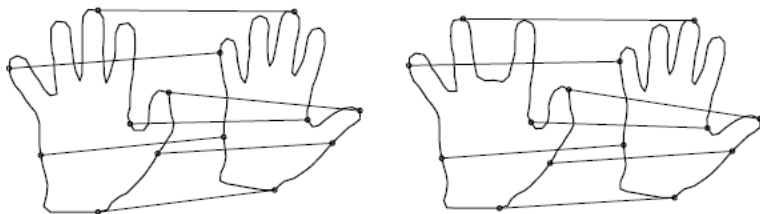
ლების მიღება ხდება ამის გათვალისწინებით. ასეთი მიდგომის მაგალითს წარმოადგენს სტილიზებული ხელნაწერი სიმბოლოები, როგორცაა, მაგალითად, ჩინური სიმბოლოები, სადაც ჯერ ხდება ცალკეული შტრიხებისა და ბლოკების იდენტიფიკაცია, რომლებისგანაც შემდეგ კონსტრუირდება შესაბამისი სიმბოლო. მეორე მიდგომისას კი ხდება ობიექტის სილუეტის კონტურის შესწავლა, კერძოდ, მასზე იძებნება კუთხეები, შეერილები, ამოხნეკილობები, ჩახნეკილობები, მაღალი სიმრუდის მნიშვნელობის მქონე წერტილები და ა.შ. ასეთი მიდგომის მაგალითს წარმოადგენს ადამიანის სახის პროფილის ამოცნობა. მაგრამ არსებობს ამოცანები, რომლებშიც ორივე მიდგომის გამოყენებაა საჭირო, მაგალითად, ტექნიკურ ნახაზებში, სადაც გეხვდება როგორც წირები და სიმბოლოები, რომელთა ამოცნობისთვის მათი სტრუქტურის ანალიზია აუცილებელი, ასევე წრეწირები და მრავალკუთხედები, რომელთა იდენტიფიკაცია მოითხოვს კონტურების ანალიზს.

ფორმაზე დამყარებული ობიექტების ამოცნობა განსხვავდება სხვა ნიშნით (მაგალითად, ობიექტზე დამყარებული ამოცნობა, მოდელზე დამყარებული ამოცნობა) განსაზღვრული ამოცნობისგან. საქმე ისაა, რომ ადამიანის მიერ მოცემული ფორმის აღქმა დამოკიდებულია დაკვირვების წერტილის მდებარეობაზე სივრცეში და ხედვის კუთხეზე. შესაბამისად, მოცემული ფორმა დამკვირვებლისთვის შეიძლება გამოჩნდეს შედარებით უფრო დიდად, ან პატარად, შეიძლება გამოჩნდეს მხოლოდ მისი ზედხედი, ან დახრილად და ა.შ. ამის გამო, ფორმის ინვარიანტულობის პრობლემა აფინური გარდაქმნების (გადატანის, ბრუნვის, მასშტაბირების) მიმართ ფორმაზე დამყარებული ამოცნობის დროს ერთ-ერთი ცენტრალურ პრობლემას წარმოადგენს. გარდა აღნიშნული პრობლემისა, რეალურ სცენაზე ჩვეულებრივ საქმე გვაქვს არა ერთ, არამედ

რამდენიმე ობიექტთან, რომლებიც ამავე დროს შეიძლება ერთმანეთს ნაწილობრივ ფარავდნენ.

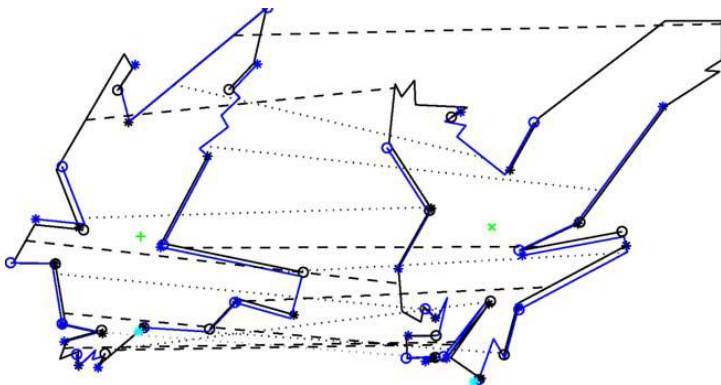
ფორმაზე დამყარებული ობიექტების ამოცნობა მნიშვნელოვანწილად ეფუძნება გამოსახულებაზე არსებული ობიექტების ფორმების რიცხვითი აღმწერების შეთანხმებულობას (matching) მონაცემთა ბაზაში არსებულ სტანდარტულ ფორმებთან. კერძოდ, თითოეული ობიექტის ფორმის აღწერისათვის, გამომდინარე იქედან, თუ რაზეა დამყარებული ფორმის წარმოდგენა და აღწერა, იყენებენ შესაბამის ფორმის აღმწერებს. შედეგად მიიღება ფორმის აღმწერი ნიშან-თვისებათა რიცხვითი ვექტორი. ამის შემდეგ ამოცნობის ამოცანა ემყარება მოცემული ფორმის აღმწერების მნიშვნელობების შეთანხმებულობას ტიპური ფორმების წარმომდგენ სტანდარტულ სიდიდეებთან და შემდგომ კლასიფიკაციას მათ შორის მსგავსების საფუძველზე.

ფორმების შეთანხმებულობისას მნიშვნელოვან ამოცანას წარმოადგენს ფორმებს შორის არსობრივი შესაბამისობების (correspondences) განსაზღვრა. ამოცანა ასეთნაირად ისმება: ვიპოვოთ მოცემული R_1, R_2, \dots, R_n ფორმების ელემენტებს შორის არსობრივი კავშირები (სიახლოვე). ეს გულისხმობს ფორმათა შორის ერთი ფორმის თითოეული ანათვლის წერტილისთვის მეორე ფორმაზე რაიმე ნიშან-თვისების (ან ნიშან-თვისებების) მიხედვით ყველაზე უფრო მსგავსი ანათვლის წერტილის პოვნას. ანათვლების წერტილებს შორის შესაბამისობების განსაზღვრა გვაძლევს შესაბამისობას მთლიანად ფორმებს შორისაც. შესაბამისობები ხელის ორ მტევანზე შერჩეულ წერტილებს (ანათვლებს) შორის ნაჩვენებია 5.18 სურათზე.



სურ. 5.18

იმ შემთხვევაში, თუ ფორმის კონტური დანაწევრებულია ცალკეულ პრიმიტივებად, მაშინ ორ ფორმას შორის შესაბამისობის განსაზღვრისათვის საჭიროა მათი კონტურების პრიმიტივებს შორის მსგავსების განსაზღვრა, რის საფუძველზეც საბოლოოდ განისაზღვრება მსგავსება შეპირისპირებულ ფორმებს შორისაც (სურ. 5.19).



სურ. 5.19

საზოგადოდ, ფორმათა შორის შესაბამისობის დადგენა რთულ ამოცანას წარმოადგენს, რადგანაც იგი მოითხოვს ფორმების სტრუქტურის გაგებას გლობალურ და ლოკალურ დონეებზე და, ასევე, ფორმების შემადგენელი ცალკეული ნაწი-

ლების ფუნქციების ცოდნას. მაგალითად, 5.20 სურათზე მოცემული ფორმებისათვის შესაბამისობის დადგენა რთულ ამოცანას წარმოადგენს გამომდინარე ობიექტების შემადგენელი ნაწილების (სახელურები, ყელები, სადგამები) განსხვავებული რაოდენობისა და ერთი და იგივე ნაწილების განსხვავებული ზომებისა.



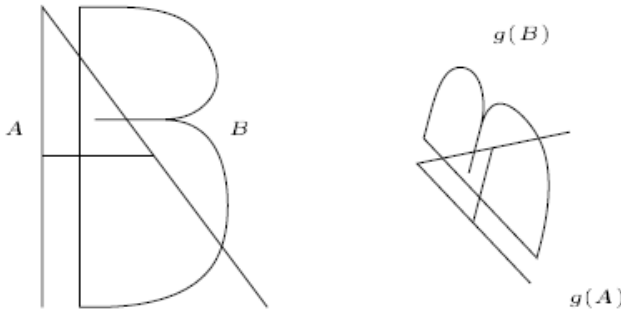
სურ. 5.20

ფორმათა შორის შეთანხმებულობის ანუ მსგავსების განსაზღვრა მოითხოვს შესაბამისი მსგავსების მეტრიკის არსებობას. **ფორმათა მსგავსების ზომა**, ანუ **მანძილის ფუნქცია** ფორმების S ერთობლიობაზე წარმოადგენს ისეთ ფუნქციას $d: S \times S \rightarrow \mathbb{R}$, რომელიც ხასიათდება მეტრიკისთვის დამახასიათებელი თვისებებით. კერძოდ, S ერთობლიობის ყველა A , B , ან C ფორმისთვის სრულდება შემდეგი პირობები:

1. (არაუარყოფითობა) $d(A, B) \geq 0$;
2. (იდენტურობა) $d(A, A) = 0$, ნებისმიერი A ფორმისთვის;
3. (ერთადერთობა) $d(A, B) = 0$, გულისხმობს, რომ $A = B$;
4. (სამკუთხედის აქსიომა) $d(A, B) + d(A, C) \geq d(B, C)$.

მანძილის ფუნქციას, რომელიც აკმაყოფილებს (2), (3) და (4) პირობებს ეწოდება **მეტრიკა**. რაც შეეხება არაუარყოფითობის

პირობას, ის გამომდინარეობს (2) და (4) პირობებიდან. თუ მანძილის ფუნქცია აკმაყოფილებს მხოლოდ (2) და (4) პირობებს, მაშინ მას ფსევდომეტრიკული ფუნქცია ეწოდება. ობიექტების ამოცნობისთვის სასურველია, რომ მსგავსების ზომა ინვარიანტული იყოს აფინური გარდაქმნების მიმართ. 5.21 სურათზე წარმოდგენილია ასეთი გარდაქმნის მაგალითი A და B ფორმებისთვის, სადაც $d(A, B) = d(g(A), g(B))$.



სურ. 5.21

ფორმათა შორის მსგავსების ზომები ძირითადად ეფუძნება ორ წერტილს შორის L_p მანძილის ცნებას. თუ $x, y \in \mathbb{R}^k$, მაშინ L_p მანძილი ასეთნაირად განისაზღვრება

$$L_p(x, y) = \left(\sum_{i=0}^k |x_i - y_i|^p \right)^{1/p}.$$

აღნიშნულ მანძილს მინკოვსკის მანძილსაც უწოდებენ. თუ $p = 2$, მაშინ ვღებულობთ ევკლიდურ L_2 მანძილს; თუ $p = 1$, მაშინ გვაქვს მანჰეტენის L_1 მანძილი; ხოლო თუ $p \rightarrow \infty$, მაშინ ვღებულობთ 'მაქს' მეტრიკას - $\max_i = (|x_i - y_i|)$.

ასევე არსებობს ამოცნობის მეთოდები, რომლებშიც ამოცნობა ხდება მანქანური სწავლების გამოყენებით, მაგრამ მას შემდეგ, რაც მიღებულია შეთანხმებულობის შედეგები. ასეთ ამომცნობ სისტემაში გამოიყენება დისკრიმინანტული კლასიფიკატორი, რომლის სწავლება სწორედ შეთანხმებულობის შედეგების საფუძველზე ხდება. ამ მიდგომისას კლასიფიკატორის აგების ერთ-ერთ ცნობილ მეთოდს წარმოადგენს ბაიესის გადაწყვეტილების თეორიის გამოყენება. აღნიშნული მიდგომის თანახმად თითოეული ფორმის აღმწერის მნიშვნელობების განაწილება ობიექტთა თითოეულ კლასისთვის საშუალებას იძლევა განისაზღვროს მოცემული ობიექტის ამა თუ იმ კლასისადმი მიკუთვნების ალბათობა. კერძოდ, ობიექტი მიეკუთვნება იმ კლასს, რომლის მიმართაც აღმწერს ყველაზე დიდი ალბათობა აქვს.

VI თავი. გამოსახულებების გაუმჯობესება და რესტავრაცია

6.1. შესავალი

ხშირად გადამწოდებიდან მიღებულ გამოსახულებაზე იმ სასურველ ინფორმაციასთან (მონაცემებთან) ერთად, რომლის გამოყოფაც აუცილებელია, გვხვდება აგრეთვე არასასურველი მონაცემებიც, რომლებიც ხელს უშლის გამოსახულების შემდგომი დამუშავებისა და ანალიზის პროცესს. მოვიყვანოთ არასასურველი მონაცემების შემცველი გამოსახულების რამდენიმე მაგალითი:

- გამოსახულების უმეტესი ნაწილი ჩაბნელებულია და მასზე არსებული ობიექტები ძნელად გასარჩევია;
- ფოტოსურათზე გვხვდება რამდენიმე კარგად შესამჩნევი გრძელი ნაკაწრი;
- დოკუმენტის სკანირებისას მისი ტექსტურ ფაილად გარდაქმნის მიზნით ფონზე ჩნდება ხმაურის პიქსელები და ასევე შესაძლებელია ზოგიერთი პიქსელის გამოტოვება ცალკეულ სიმბოლოებში.

ასეთი მონაცემებისგან გამოსახულების გასათავისუფლებლად საჭირო ოპერაციები ხორციელდება ხოლმე ცალკეული პიქსელისა და მისი გარემოს დონეზე. გამოსახულების გასაუმჯობესებლად გამოიყენება გამოსახულებების დამუშავების მეთოდები, კერძოდ ორი პროცედურა: **გამოსახულების ხარისხის გაუმჯობესება და რესტავრაცია (აღდგენა).**

გამოსახულების გაუმჯობესების არსი მდგომარეობს გამოსახულებაზე ისეთი გარდაქმნების (ოპერაციების) ჩატარებაში, რომლებიც გაზრდის მათზე არსებული ობიექტების ან მნიშვნელოვანი დეტალების აღმოჩენის შესაძლებლობას ადამიანის ან მანქანის მიერ, ანუ სხვა სიტყვებით, გააადვილებს გამოსახულების აღქმას მისი ვიზუალური ხარისხის გაუმჯობესების ხარჯზე. ამ ტიპის გარდაქმნებს **გამოსახულების პრეპარირებასაც** უწოდებენ. ასეთი გარდაქმნების მაგალითებია: ხმაურის ჩახშობის, დაგლუვების, კონტრასტის გაძლიერების და კიდევების გამოკვეთის ოპერატორები.

რესტავრაციის ქვეშ გულისხმობენ გამოსახულების რესტავრაციას (აღდგენას), ანუ მასზე ისეთი გარდაქმნების ჩატარებას, რომლის შედეგად დამახინჯებული გამოსახულება მიუახლოვდება იდეალურ გამოსახულებას.

ზემოთქმულიდან გამომდინარე ორივე პროცედურის მიზანია გამოსახულების გარდაქმნა ისეთ ახალ გამოსახულებად, რომელიც უფრო მოხერხებელი იქნება მისი შემდგომი ანალიზისთვის. გამოსახულების სუბიექტური აღქმის გაუმჯობესება შესაძლებელია, თუ გამოსახულების აღმწერ ფუნქციაზე გამოვიყენებთ ადამარის და ფურიეს გარდაქმნებს. გამოსახულების გაუმჯობესების მიზანს არ წარმოადგენს ის, რომ მიღებული გამოსახულება მიუახლოვდეს რაღაც იდეალიზებულ ორიგინალს. ასეთი ამოცანა განიხილება გამოსახულების რესტავრაციისას.

რესტავრაციისას ხდება გამოსახულებაზე იმ დამახინჯებების აღმოფხვრა, რომლებიც წარმოიშობა დამუშავების პროცესის მსვლელობისას აპარატურის ზემოქმედების შედეგად. რესტავრაციის მეთოდებიდან აღსანიშნავია ისეთები, რომლებიც ხელს უწყობს სცენის გამოსახულებაზე ობიექტების გამოყოფის ამოცანის გადაწყვეტას.

განვიხილოთ რამდენიმე მეთოდი, რომელთა მიზანია დამახინჯებების მქონე გამოსახულებებიდან ისეთი გამოსახულებების მიღება, რომლებიც კარგად აღქმადი იქნება დამკვირვებლის მხრიდან და ასევე ადვილად დასამუშავებელი კომპიუტერისთვის.

6.2. ტონების განაწილების ცვლილება გამოსახულებაზე

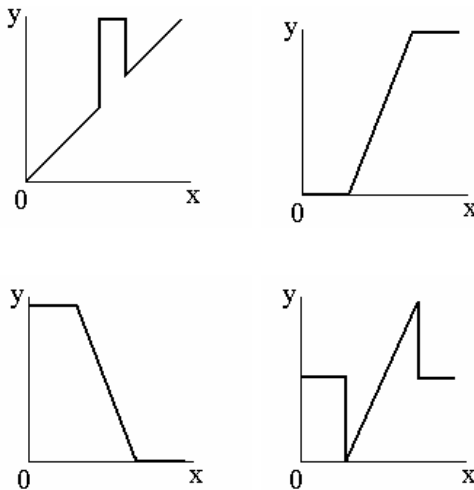
6.2.1. კონტრასტის შეცვლა

ფოტო და სატელევიზიო გამოსახულებების ყველაზე უფრო გავრცელებულ დეფექტს წარმოადგენს სუსტი კონტრასტი, რაც გამოწვეულია აღსადგენი ინტენსივობის მნიშვნელობების დიაპაზონის შეზღუდულობით და ხანდახან შესაბამისი დონეების გადაცემის მახასიათებლის არაწრფივობით. ასეთ დროს გამოსახულებაზე ინტენსივობის დონე იცვლება არა შავიდან თეთრისკენ, არამედ რუხიდან ცოტა უფრო ღია რუხისკენ. ეს ნიშნავს, რომ გამოსახულებაზე ინტენსივობის მნიშვნელობების ცვლილების დიაპაზონი დასაშვებზე უფრო ვიწროა.

ბევრ შემთხვევაში შესაძლებელია კონტრასტის შეცვლა, თუ მოვახდენთ გამოსახულების ყოველი პიქსელის სიკაშკაშის ცვლილებას. დაბალკონტრასტული უწყვეტი გამოსახულებების შემთხვევაში ასეთი კორექცია რთული განსახორციელებელია, რადგან ასეთ დროს ძნელია ნებისმიერი დონის გადაცემის მახასიათებლის რეალიზება მაღალი სიზუსტით. ციფრული გამოსახულების შემთხვევაში კი საჭირო მახასიათებლის მიღება შედარებით ადვილია. თუმცა ამ დროს არაწრფივი ოპერატორის მოძიებისას საჭიროა მხედველობაში მივიღოთ კვანტორების შეცდომები. ასეთი პროცედურის ჩატარების შედეგად შეიძლება აღმოჩნდეს, რომ გამოსახულებული გამოსახულების

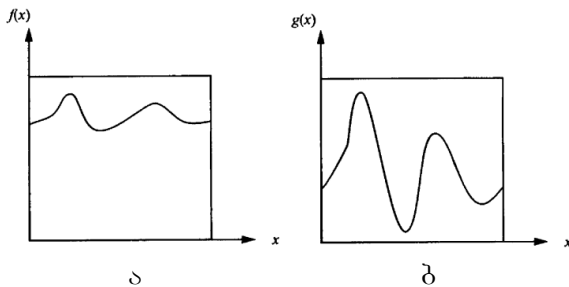
ინტენსივობის დიაპაზონში ზოგიერთი დონე არ იქნეს გამოყენებული. ეს გამოიწვევს იმას, რომ გამოსასვლელ გამოსახულებაზე ინტენსივობის ზოგიერთი ვარდნები გადააჭარბებს შესაბამის ვარდნებს შესასვლელ გამოსახულებაზე, რასაც შედეგად შეიძლება მოჰყვეს მცდარი კონტურების გაჩენა. ამის თავიდან ასაცილებლად გამოსასვლელი გამოსახულება იკვანტება უფრო მეტი რაოდენობის დონეებით, ვიდრე შესასვლელი.

გამოსახულების ხარისხის შესაცვლელად ხშირად იყენებენ პიქსელების ინტენსივობების (ნახევარტონური დონეების) მნიშვნელობების ცვლილებას. ამისათვის გამოიყენება სხვადასხვა გარდაქმნის ფუნქციები (ოპერატორები), რომლებსაც **ტონური განაწილების ფუნქციები** ეწოდება. მათი საშუალებით მოიცემა ინტენსივობის შესასვლელი მნიშვნელობების ახალი მნიშვნელობებით შეცვლის წესი. რამდენიმე ასეთი ტონური განაწილების ფუნქცია წარმოდგენილია 6.1 სურათზე.



სურ. 6.1

ეს ფუნქციები განსაზღვრავს შესასვლელი პიქსელების ინტენსივობების მნიშვნელობების ახალ, გამოსასვლელ მნიშვნელობებად გარდაქმნის წესს. გამოსასვლელზე ინტენსივობის მნიშვნელობების გარდაქმნას **გაჭიმვას** (stretching) ან **ნორმალიზაციას** უწოდებენ, რადგან აღნიშნული გარდაქმნის საშუალებით ხდება ინტენსივობის მნიშვნელობების დინამიკური დიაპაზონის **გაფართოება – გაჭიმვა**. ჰისტოგრამის გარდაქმნის ოპერაცია შეიძლება ჩატარდეს როგორც მთლიან გამოსახულებაზე, ასევე მის რომელიმე არეზე. ასეთი გარდაქმნის მაგალითი მოცემულია 6.2 სურათზე. 6.2,ბ სურათზე წარმოდგენილია ინტენსივობის დინამიკური დიაპაზონის გაფართოების ანუ გაჭიმვის შედეგი 6.2,ა სურათზე წარმოდგენილი შემთხვევისთვის. როგორც ვხედავთ, ინტენსივობის დინამიკური დიაპაზონის გაფართოების შედეგად მოხდა ნახევარტონების ღონეების მნიშვნელობათა დიაპაზონის გაზრდა, რამაც შედეგად გამოსახულების კონტრასტის გაზრდა გამოიწვია. ცხადია, რეალურად ასეთ ოპერაციას პრაქტიკაში იყენებენ ჩაბნელებული გამოსახულებების შემთხვევაში.



სურ. 6.2

გამოსახულებაზე კონტრასტის გაზრდისთვის, ანუ ინტენსივობის მნიშვნელობების დინამიკური დიაპაზონის გაფართოე-

ბისთვის იყენებენ წრფივი კონტრასტირების მეთოდს, რომელიც წარმოადგენს ასეთი სახის გარდაქმნას:

$$g(n_1, n_2) = af(n_1, n_2) + b \quad (6.1)$$

სადაც, a და b მუდმივი პარამეტრებია. აღნიშნული პარამეტრების განსაზღვრა შეიძლება ორი ხერხით.

პირველ შემთხვევაში შესასვლელ გამოსახულებაზე განისაზღვრება არსებული ინტენსივობის f მნიშვნელობები და დიაპაზონი $[f_{min}, f_{max}]$. ამის შემდეგ (6.1) გარდაქმნის თანახმად, გამოსასვლელ გამოსახულებაზე უნდა მივიღოთ ინტენსივობის ახალი g მნიშვნელობები, ანუ $[g_{min}, g_{max}]$ დიაპაზონი. ეს გარდაქმნა ასე შეიძლება ჩაიწეროს:

$$\begin{cases} g_{min} = af_{min} + b \\ g_{max} = af_{max} + b \end{cases}$$

საიდანაც განისაზღვრება a და b პარამეტრების მნიშვნელობები:

$$a = \frac{g_{max} - g_{min}}{f_{max} - f_{min}}, \quad b = \frac{g_{min} f_{max} - g_{max} f_{min}}{f_{max} - f_{min}}$$

მეორე შემთხვევაში, a და b პარამეტრებისთვის იღებენ ისეთ მნიშვნელობებს, რომლებსაც ინტენსივობის მნიშვნელობების მათემატიკური მოლოდინი და დისპერსია დაჰყავთ “სტანდარტულ” სიდიდეებამდე. ამისათვის ჯერ წინასწარ ხდება შესასვლელი გამოსახულების μ_f მათემატიკური ლოდინისა და σ_f^2 დისპერსიის შეფასება, ხოლო შემდეგ a და b პარამეტ-

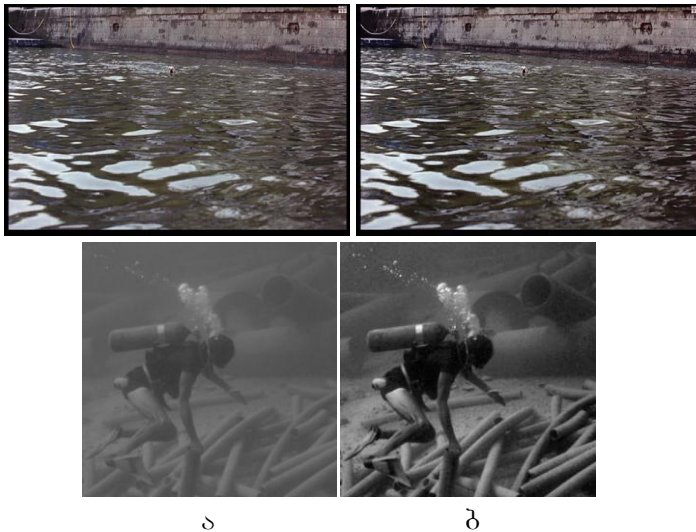
რები ისეთნაირად შეირჩევა, რომ გამოსასვლელი გამოსახულებათვის მივიღოთ “სტანდარტული” μ_g და σ_g^2 :

$$g(n_1, n_2) = \frac{f(n_1, n_2) - \mu_f}{\sigma_f} \cdot \sigma_g + \mu_g = \frac{\sigma_g}{\sigma_f} \cdot f(n_1, n_2) + \mu_g - \mu_f \frac{\sigma_g}{\sigma_f},$$

ანუ შედეგად გვექნება:

$$a = \frac{\sigma_g}{\sigma_f}; \quad b = \mu_g - \mu_f \frac{\sigma_g}{\sigma_f}.$$

გამოსახულებაზე კონტრასტის გაზრდის მაგალითები მოცემულია 6.3 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) კონტრასტის გაზრდის შედეგად მიღებული გამოსახულება.



სურ. 6.3

როგორც ვხედავთ, მოხდა ინტენსივობის მნიშვნელობების დინამიკური დიაპაზონის გაფართოება.

აღსანიშნავია, რომ მოცემულ გამოსახულებებზე, იმისდა მიხედვით, თუ როგორ გამოიყურება ისინი ვიზუალურად, მაგალითად, ვთქვათ საწყისი გამოსახულება ბნელია ან მასზე გვხვდება ცალკეული განსხვავებული სიბნელის მქონე უბნები (არეები), ამ გამოსახულებისთვის შეიძლება გამოყენებული იქნეს ტონური განაწილების ისეთი ფუნქცია, რომელიც ბნელი პიქსელების სიკაშკაშეს გაზრდის, ხოლო ღია პიქსელების სიკაშკაშეს კი შეამცირებს.

ერთ-ერთ ასეთ ტონური განაწილების ფუნქციას აქვს ასეთი სახე – $f(x) = x^{1/\gamma}$. როგორც ვხედავთ, ის არაწრფივი ფუნქციაა და იმისდა მიხედვით, თუ რა მნიშვნელობებს მიიღებს γ , შედეგიც სხვადასხვა იქნება. ასეთი ფუნქციის მიერ განხორციელებულ ინტენსივობის მნიშვნელობების დიაპაზონის გარდაქმნას გამოსახულებაზე **გამა-კორექციას** (Gamma correction) უწოდებენ. თუ $f(x) = x^{0.5}$, მაშინ ასეთი ფუნქცია ინტენსივობის დაბალ მნიშვნელობებს უფრო მეტად გაზრდის, ვიდრე მაღალს. თუ $\gamma = 2.0$, მაშინ ეს გამოიწვევს ინტენსივობის მნიშვნელობების გაზრდას. გარდა აღნიშნული ფუნქციისა, შესაძლებელია გამოყენებული იქნეს უფრო რთული სახის ფუნქციაც. ზოგადად შეიძლება ითქვას, რომ ამ ფუნქციების მოქმედების შედეგად გამოსახველ გამოსახულებაზე პიქსელების ინტენსივობის მნიშვნელობებს შორის განსხვავება უნდა გახდეს უფრო შესამჩნევი, რაც გამოსახულების აღქმას უფრო გააადვილებს. ამგვარად, გამოსახულების გაუმჯობესება შესაძლებელია მის პიქსელებზე გარკვეული ფუნქციის – ოპერატორის გამოყენებით. განსხვავებენ შემდეგი სახის ოპერატორებს:

განსაზღვრება 6.1. *წერტილოვანი ოპერატორი f ეს არის ოპერატორი, რომლის მოქმედების შედეგად გამოსასვლელი პიქსელის მნიშვნელობა განისაზღვრება მხოლოდ ერთი შესასვლელი პიქსელის მნიშვნელობით, ანუ $Out[x, y] = f(Inp[x, y])$.*

აქ ოპერატორის f ფუნქცია შეიძლება დამოკიდებული იყოს გარკვეულ გლობალურ პარამეტრებზე.

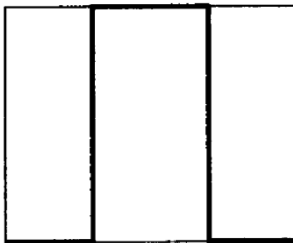
განსაზღვრება 6.2. *კონტრასტის გაზრდის ოპერატორი ეს არის წერტილოვანი ოპერატორი, რომელიც გამოსახულებაზე მნიშვნელოვანი დეტალების გასაძლიერებლად ინტენსივობის შესასვლელი მნიშვნელობებისთვის იყენებს უბან-უბან გლუვ ფუნქციას $f(Inp[x, y])$.*

იმის გამო, რომ წერტილოვანი ოპერატორი მოქმედებს ერთ შესასვლელ პიქსელზე და გამოსასვლელიც ერთი პიქსელია, მათი გამოყენება პიქსელთა მასივზე შესაძლებელია როგორც ნებისმიერი მიმდევრობით, ასევე ერთდროულადაც (პარალელურად). უნდა აღინიშნოს, რომ ძალიან სასარგებლოა გამოსახულების ხარისხის გაუმჯობესება არამონოტონური ტონური განაწილების ფუნქციებით, რადგანაც ასეთი გამოსახულებები კარგად აღიქმება ადამიანის მიერ, მაგალითად, დიზაინსა და ჟურნალისტიკაში. მაგრამ იგივე მიდგომის გამოყენება, მაგალითად, რენტგენოლოგიაში უნდა განხორციელდეს ძალიან ფრთხილად, რათა არ მოხდეს ინტენსივობის იმ არსებითი შკალის შეცვლა, რომელსაც იყენებენ ადამიანი-ექსპერტები და რომელზედაც შეიძლება მომართული იყოს სხვადასხვა გადამწოდები.

6.2.2. პრეპარირება

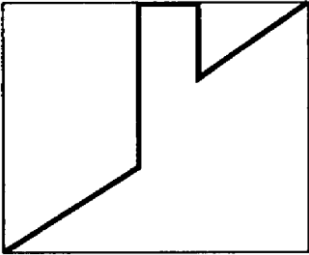
გამოსახულებაზე ინტენსივობის მნიშვნელობების გარდაქმნის საკმაოდ ფართო კლასს **პრეპარირებას** უწოდებენ. პრეპარირების ბევრი პროცედურა გულისხმობს გამოსახულების ელემენტ-ელემენტ გარდაქმნას მისი ვიზუალური ხარისხის გაუმჯობესების მიზნით. ზღურბლური დამუშავება ამ გარდაქმნის კერძო შემთხვევას წარმოადგენს. გამოსახულების ელემენტ-ელემენტ გარდაქმნის ფუნქციების (პრეპარირებისთვის გამოყენებული ფუნქციების) თავისებურებას წარმოადგენს ის, რომ ჯერ ერთი, ძნელია მათი ფიზიკური ინტერპრეტაცია და რეალურად გარდაქმნის ფუნქციის შერჩევა ხდება ემპირიულად, გამომდინარე კონკრეტული ამოცანის გადაწყვეტიდან და, მეორეც, პრეპარირება ხორციელდება დიალოგურ რეჟიმში, შესაბამისად, გარდაქმნის ფუნქცია პარამეტრების მცირე რაოდენობას უნდა შეიცავდეს და მათი არსი ოპერატორისთვის გასაგები უნდა იყოს. განვიხილოთ პრეპარირებისთვის გამოყენებული გარდაქმნის ზოგიერთი ფუნქციის მოქმედება.

გამოსახულების ზღურბლური დამუშავების განზოგადებას წარმოადგენს გარდაქმნა, რომელსაც **ინტენსივობის ჩამოჭრა** ეწოდება (სურ. 6.4).



სურ. 6.4

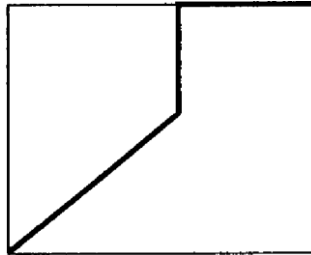
ამ გარდაქმნის არსი შემდეგში მდგომარეობს: შესასვლელი გამოსახულების სიკაშკაშის მნიშვნელობებიდან ხდება მნიშვნელობების გარკვეული ე.წ. “შუშა” დიაპაზონის გამოყოფა. ასეთი დიაპაზონის სიკაშკაშის მნიშვნელობათა შკალაზე გადაადგილებით და მისი სიგანის ცვლილებით შესაძლებელია განისაზღვროს, თუ სიკაშკაშის რა მნიშვნელობებია წარმოდგენილი მოცემულ გამოსახულებაზე (ასევე გამოსახულების რომელ წერტილებში) და რომელი არა; ასევე მოვახდინოთ გამოსახულებაზე ერთმანეთისგან სიკაშკაშეებით განსხვავებული ცალკეული ობიექტების ვიზუალური ანალიზი. გამოსახულებაზე არსებული ის არეები (დეტალები), რომლებიც არ ხვდება სიკაშკაშეების მოცემულ დიაპაზონში, ანუ მიეკუთვნება ფონურ ელემენტებს, დათრგუნული იქნება. სურათზე 6.4. მოცემულ შემთხვევაში ფონი შავია, ანუ აღნიშნული გარდაქმნით ხდება შავი ფონის დათრგუნვა. 6.5 სურათზე ნაჩვენებია შემთხვევა, როცა ადგილი აქვს ინტენსივობის ჩამოჭრის ოპერაციას ფონის შენარჩუნებით.



სურ. 6.5

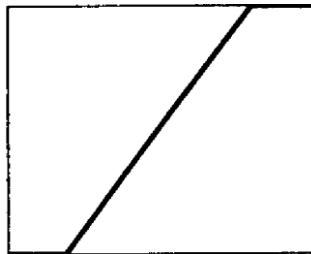
ასეთ შემთხვევაში გამოსახულება ძირითადად შენარჩუნებულია, მაგრამ მასზე ადგილი აქვს სიკაშკაშის მნიშვნელობების მოცემულ ინტერვალში მოხვედრილი ცალკეული უბნების “გამონათებას”. თუ ეს ინტერვალი ეკვრის სიკაშკაშის შკალის

საზღვარს, მაშინ გვაქვს ე.წ. არასრული ზღურბლური გარდაქმნის შემთხვევა (სურ. 6.6).



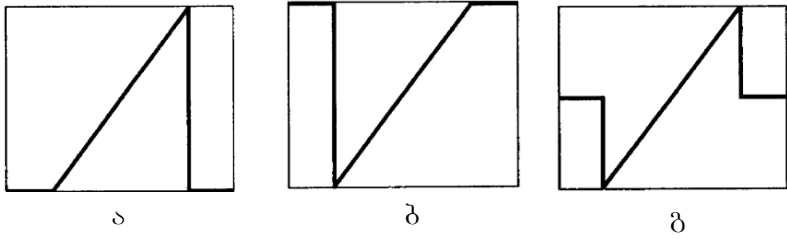
სურ. 6.6

თავისი არსით წრფივი კონტრასტირების გარდაქმნის მსგავსია კონტრასტული მასშტაბირება. ამ ფუნქციის საშუალებით გამოსახულებაზე სიკაშკაშის მნიშვნელობების დიაპაზონის გარკვეული ინტერვალი (უბანი) იჭიმება მთელ შკალაზე (სურ. 6.7). ასეთი ოპერაციის (გარდაქმნის) შედეგად გამოსახულებაზე იმ დეტალების კონტრასტი, რომელთა სიკაშკაშის მნიშვნელობები აღნიშნულ ინტერვალში მოხვდა, იზრდება, ხოლო ის დეტალები, რომელთა სიკაშკაშის მნიშვნელობები ამ ინტერვალის გარეთაა, გარდაიქმნება ერთგვაროვან ფონურ ელემენტებად.



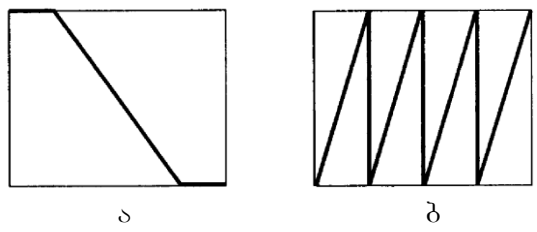
სურ. 6.7

კერძოდ: შავ (სურ. 6.8,ა), თეთრ (სურ. 6.8,ბ) ან რუხ (სურ. 6.8,გ) ფონებად.



სურ. 6.8

კონტრასტული მასშტაბირება შეიძლება ასევე გამოყენებულ იქნეს სიკაშკაშის მნიშვნელობების შებრუნებული მნიშვნელობების მისაღებად, ანუ მოცემული გამოსახულების შებრუნებული გამოსახულების – “ნეგატივის” მისაღებად (სურ. 6.9,ა).



სურ. 6.9

გარდაქმნის კიდევ ერთ ფუნქციას წარმოადგენს კონტრასტული მასშტაბირების ხერხისებური ფუნქცია (სურ. 6.9,ბ). პრაქტიკაში ის საკმაოდ წარმატებულად გამოიყენება ისეთ გამოსახულებებზე, რომლებზედაც გვაქვს რამდენიმე მსხვილი არე სიკაშკაშის მნიშვნელობების მდორე ცვლილებით. ასეთი გარდაქმნის შედეგად პრაქტიკულად არ იცვლება გამოსახულე-

ბის აღქმის ხარისხი, მაგრამ მკვეთრად იზრდება გამოსახულებაზე არსებული წვრილ-წვრილი დეტალების კონტრასტი.

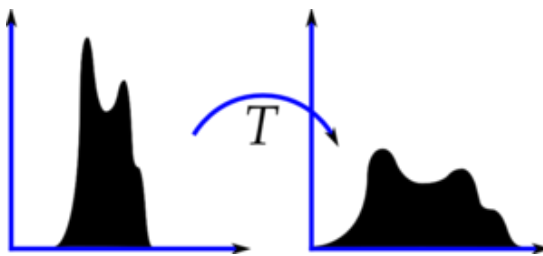
გამოსახულების ელემენტ-ელემენტ გარდაქმნას მიეკუთვნება ასევე გამოსახულების გარდაქმნა **ფსევდოფერებში**. აღნიშნული გარდაქმნისას სიკაშკაშის ყოველ რიცხვით მნიშვნელობას მონიტორის ეკრანზე შეუსაბამებენ განსაზღვრულ ფერს. ცნობილია, რომ ადამიანის თვალი უფრო აღიქვამს ფერების ცვლილებას გამოსახულებაზე, ვიდრე სიკაშკაშის. ამის გამო ასეთი გარდაქმნა ხშირად გამოიყენება პრაქტიკაში, განსაკუთრებით კი სამედიცინო დიაგნოსტიკის სისტემებში. ზოგადად, შესაბამისობის წესი შეიძლება ნებისმიერი იყოს, თუმცა პრაქტიკაში ცდილობენ, რომ გარდაქმნის ფუნქცია თავისი ხასიათით იყოს გლუვი. ასეთი გარდაქმნის შედეგად შესასვლელ გამოსახულებაზე სიკაშკაშის მდოვრე ცვლილებას გამოსასვლელ გამოსახულებაზე ფერის ასევე მდოვრე ცვლილება უნდა შეესაბამებოდეს.

6.2.3. ჰისტოგრამის გარდაქმნა

თუ დაგაკვირდებით წრფივი კვანტირების შედეგად მიღებულ ბუნებრივი სცენის გამოსახულებას, ვნახავთ, რომ მისი სიკაშკაშეების განაწილების ჰისტოგრამაზე უფრო დაბალი დონეების სიკაშკაშეები ჭარბობს, რაც ნიშნავს, რომ გამოსახულების ელემენტების უმრავლესობას ექნება საშუალოზე დაბალი სიკაშკაშის მნიშვნელობები. ასეთ გამოსახულებებზე მუქ უბნებში დეტალების გარჩევა ძნელი ხდება. ასეთი გამოსახულებების გაუმჯობესების ერთ-ერთ მეთოდს წარმოადგენს **ჰისტოგრამის გათანაბრება** (histogram equalization). ამ მიდგომის არსი მდგომარეობს საწყისი გამოსახულების სიკაშკაშეების ისეთ გარდაქმნაში მის ჰისტოგრამაზე დაყრდნობით, რომ ახალი

გამოსახულების ჰისტოგრამას ჰქონდეს სასურველი სახე, ანუ ახალი გამოსახულების ჰისტოგრამაზე მივიღოთ სიკაშკაშეების ისეთი განაწილება, რომელიც გამოსახულებას ვიზუალურად უფრო კარგად აღსაქმელს გახდის. ასეთი გარდაქმნის არსი სქემატურად წარმოდგენილია 6.10 სურათზე.

გამოსახულებაზე ასეთი გარდაქმნის განსახორციელებლად იყენებენ ოპერატორს, რომელსაც ედება შემდეგი ორი შეზღუდვა: ა) გამოსახულელ გამოსახულებაში გამოყენებულ უნდა იქნეს სიკაშკაშის ყველა დასაშვები მნიშვნელობები, და ბ) გამოსახულელ გამოსახულებაში სიკაშკაშის ყოველი მნიშვნელობისთვის უნდა გვქონდეს პიქსელების დაახლოებით ერთნაირი რაოდენობა. პირველი შეზღუდვა იძლევა გამოსახულების კარგი აღქმის შესაძლებლობას, ხოლო მეორე დამატებითა და მისი ვარჯისიანობა ფასდება ექსპერიმენტების შედეგებით. ჰისტოგრამის გათანაბრება შეიძლება შესრულდეს როგორც მთელი გამოსახულებისთვის (global histogram equalization), ასევე მისი ცალკეული არეებისთვის (local histogram equalization).



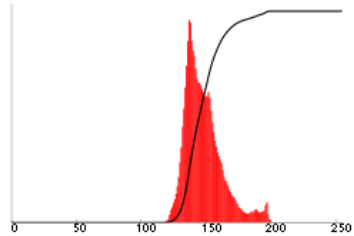
სურ. 6.10

ჰისტოგრამის გათანაბრების ოპერატორის მოქმედების შედეგი ნაჩვენებია 6.11 სურათზე, სადაც მოცემულია როგორც თვით (ა) შესასვლელი გამოსახულება, მისი (ბ) ტონური განა-

წილების ფუნქცია და ჰისტოგრამა, ასევე მის ჰისტოგრამაზე გაჭიმვის ოპერატორის მოქმედების შედეგად მიღებული (გ) გამოსასვლელი გამოსახულება, მისი (დ) ტონური განაწილების ფუნქცია და ჰისტოგრამა.



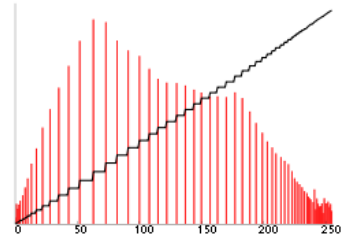
ა



ბ



გ



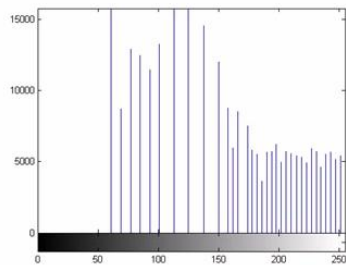
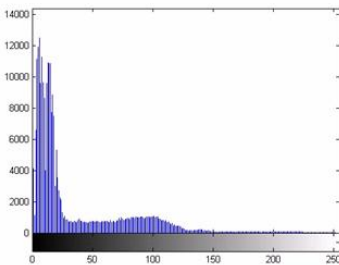
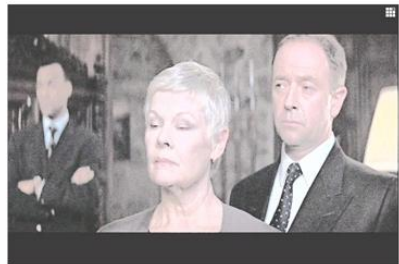
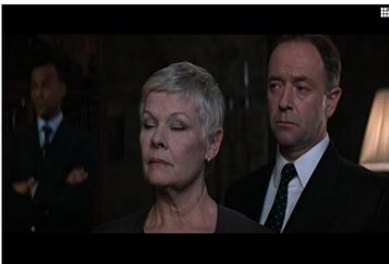
დ

სურ. 6.11

როგორც ვხედავთ, გამოსახულების ტონური განაწილების მრუდის ცვლილება მართლაც იწვევს როგორც მთლიანი გამოსახულების სახის ცვლილებას, ასევე მასზე არსებული ცალკეული დეტალების უფრო გამოკვეთილად წარმოჩენას. ჰისტოგრამის გათანაბრების კიდევ ერთი მაგალითი წარმოდგენილია 6.12 სურათზე, სადაც (ა) საწყისი გამოსახულება და მი-

სი შესაბამისი ჰისტოგრამაა, ხოლო (ბ) ჰისტოგრამის გათანაბრების შედეგად მიღებული გამოსახულება და მისი შესაბამისი ჰისტოგრამაა. სურათიდან ჩანს, რომ საწყისი გამოსახულების ჰისტოგრამას გათანაბრებამდე ჰქონდა პიკის სახე, ხოლო გათანაბრების შემდეგ მოხდა მისი გაჭიმვა, რამაც მკვეთრად შეცვალა საწყის გამოსახულებაზე სიკაშკაშეების განაწილება და შესაბამისად გამოსახულება ვიზუალურად უფრო კარგად აღსაქმელი გახდა.

მაგალითად, დიდ ერთგვაროვან არეებზე (ცა, მინდორი და ა.შ.) ჰისტოგრამის გათანაბრების გამოყენების შედეგად აღნიშნული შეზღუდვის გამო მიღებულ გამოსახულებაზე გვექნება სიკაშკაშის უფრო მეტი განსხვავებული მნიშვნელობები და, შესაბამისად, გამოსახულება უფრო კონტრასტული იქნება.



ა

ბ

სურ. 6.12

(ა) და (ბ) შეზღუდვები იმას ნიშნავს, რომ მიღებულ გამოსასვლელ გამოსახულებაზე გვექნება სიკაშკაშის ყველა დასაშვები $z = z_1, z = z_2, \dots, z = z_n$ მნიშვნელობები; ამასთან თითოეული z_k გამოყენებული იქნება დაახლოებით $q = (N \times M)/n$ -ჯერ. აქ N და M გამოსახულების სტრიქონებისა და სვეტების რაოდენობაა. იმისათვის, რომ განისაზღვროს გამოსახულების დინამიკური დიაპაზონის გამჭვირავი ტონური გარდაქმნის f ფუნქცია, საჭიროა ვიცოდეთ მხოლოდ შესასვლელი გამოსახულების $H_{imp}[i]$ ჰისტოგრამა (i -სიკაშკაშის მნიშვნელობაა გამოსახულებაზე). ამ ჰისტოგრამის თითოეული ჯიბე $[i]$ ინახავს შესასვლელ გამოსახულებაზე z_i სიკაშკაშის მნიშვნელობის მქონე პიქსელების რაოდენობას. ამის შემდეგ ხდება სიკაშკაშის პირველი ზღურბლური t_1 მნიშვნელობის განსაზღვრა. ამისათვის შესასვლელი გამოსახულების ჰისტოგრამაზე ხდება i -ური ინდექსის გაზრდა მანამ, სანამ ნაკლები სიკაშკაშის მნიშვნელობების მქონე პიქსელების ჯამური რაოდენობა არ გადააჭარბებს q_1 ზღურბლს. შედეგად მივიღებთ იმას, რომ შესასვლელი გამოსახულების ყველა იმ პიქსელს, რომელთა სიკაშკაშის მნიშვნელობა $z_k < t_1 - 1$ -ზე, გარდაქმნის შედეგად გამოსასვლელ გამოსახულებაზე მიენიჭება z_1 მნიშვნელობა. თვით t_1 ზღურბლის მნიშვნელობა კი განისაზღვრება შემდეგი უტოლობიდან

$$\sum_{i=1}^{t_1-1} H_{imp}[i] \leq q_1 < \sum_{i=1}^{t_1} H_{imp}[i].$$

უტოლობა გვიჩვენებს, რომ t_1 წარმოადგენს სიკაშკაშის იმ უმცირეს მნიშვნელობას, რომლისთვისაც ჰისტოგრამაზე არსებობს სიკაშკაშის ნაკლები მნიშვნელობების მქონე არაუმეტეს

q_1 პიქსელისა. k -ური ზღურბლური მნიშვნელობა t_k გამოითვლება იტერაციულად შემდეგი უტოლობის საფუძველზე

$$\sum_{i=1}^{t_k-1} H_{inp} [i] \leq (q_1 + q_2 + \dots + q_k) < \sum_{i=1}^{t_k} H_{inp} [i].$$

პრაქტიკაში კი გარდაქმნის f ფუნქცია მოიცემა ე.წ. **გარდაქმნის ცხრილის** (look-up table) სახით. მისი აგება ადვილია, თუ გამოვიყენებთ ზემომოყვანილ ფორმულებს. t_k ზღურბლების გამოთვლილი მნიშვნელობები თავსდება $T[i]$ მასივში მანამ, სანამ სრულდება უტოლობა. ამგვარი მოქმედებების შედეგად მიიღება გარდაქმნის ფუნქცია

$$z_{out} = f(z_{inp}) = T[z_{inp}].$$

უნდა აღინიშნოს, რომ ხშირად გამოსასვლელი გამოსახულების სიკაშკაშის დასაშვებ მნიშვნელობათა დიაპაზონი არის ხოლმე უფრო ფართო, ვიდრე შესასვლელის. ასეთ შემთხვევაში შეუძლებელია გარდაქმნის ისეთი f ფუნქციის აგება, რომელიც შესასვლელი გამოსახულების ნახევარტონების დონეებს ასახავს გამოსასვლელი გამოსახულების დონეების მთელ დიაპაზონზე.

გარდა კონტრასტის გაზრდისა ჰისტოგრამის გათანაბრებას იყენებენ იმ შემთხვევაშიც, როცა საჭიროა ორი ან მეტი გამოსახულების ერთმანეთთან შედარება რომელიმე სპეციფიკური მახასიათებლის, მაგალითად ტექსტურული მახასიათებლის მიხედვით. ამისათვის ჯერ მათი ჰისტოგრამები ნორმალიზაციის საშუალებით დაიყვანება “სტანდარტულ” სახეზე (ეს განსაკუთრებით სასარგებლოა, როცა გამოსახულებები მიღებულია სხვადასხვა გარემოდან), ხოლო შემდეგ კი $b = f(a)$

(*a*-ინტენსივობის მნიშვნელობაა შესასვლელ გამოსახულებაზე, *b*-კი გამოსასვლელ გამოსახულებაზე) ფუნქციის გამოყენებით ხდება ჰისტოგრამის გარდაქმნა ისეთ ჰისტოგრამად, რომელიც მუდმივი იქნება ინტენსივობის ნებისმიერი მნიშვნელობისთვის. გამოსახულების სტატისტიკური აღწერის შემთხვევაში ჰისტოგრამის “სტანდარტულ” სახეზე დაყვანა ნიშნავს ინტენსივობის სიმკვრივის ფუნქციის დაყვანას “სტანდარტულ” სახეზე. ასეთი გარდაქმნის რეალიზაციისთვის გამოსახულების თითოეულ ელემენტზე ხორციელდება არაწრფივი გარდაქმნა, რომელიც აიგება შესასვლელი გამოსახულების ინტენსივობის მნიშვნელობათა განაწილების ჰისტოგრამის საშუალებით. ეს შეესაბამება ისეთი გამოსასვლელი გამოსახულების მიღებას, რომელზედაც ინტენსივობის მნიშვნელობები მოცემული დინამიკური დიაპაზონის $[g_{min}, g_{max}]$ ფარგლებში თანაბარაღბათური იქნება (სურ. 6.13,ა), ანუ

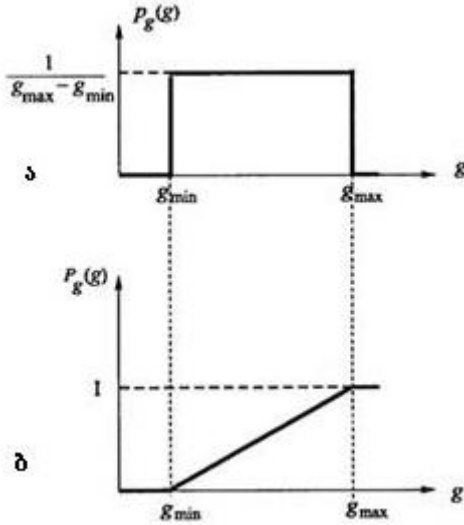
$$p_g(g) = \frac{1}{g_{max} - g_{min}} , \text{ სადაც, } g_{min} \leq g \leq g_{max} .$$

აღნიშნულ ინტერვალში განაწილების ფუნქცია წრფივია (სურ. 6.13,ბ):

$$P_g(g) = \frac{g - g_{min}}{g_{max} - g_{min}} .$$

აქედან,

$$P_f(f) = \frac{g(f) - g_{min}}{g_{max} - g_{min}} .$$



სურ. 6.13.

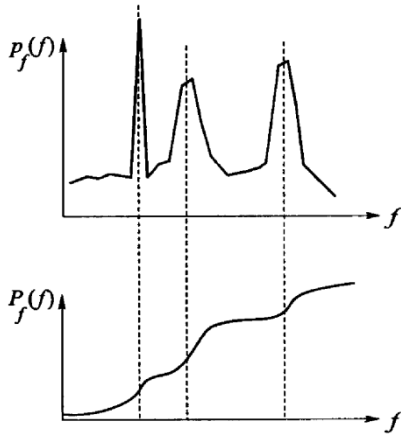
შედეგად შესასვლელი გამოსახულების ჰისტოგრამის გარდაქმნის ფუნქცია ასეთი იქნება:

$$g(f) = (g_{max} - g_{min})P_f(f) + g_{min}.$$

თუმცა, პრაქტიკულად, ასეთი შედეგი ნებისმიერი გამოსახულებისთვის შეიძლება მიახლოებითი იყოს. რეალურად მიღებული ჰისტოგრამები ძალიან არათანაბარია და ხასიათდება ამოხნეკილობებისა და ჩავარდნების დიდი რაოდენობით. ჰისტოგრამაზე ინტენსივობის იმ მნიშვნელობებს, რომლებიც ყველაზე უფრო მაღალაღბათურია, $p_f(f)$ ჰისტოგრამაზე შეესაბამება წვეროები (მაქსიმუმები), ხოლო ინტენსივობის განაწილების ფუნქცია $P_f(f)$ ამ ადგილებში მკვეთრად ზრდადი იქნება (სურ. 6.14)

“შესაფერისი” გარდაქმნის $b = f(a) = g(f)$ ფუნქციის აგების შემთხვევაში, დამოკიდებულება ინტენსივობის შესასვლელ და გამოსასვლელ სიმკვრივის ფუნქციებსა და გარდაქმნის $b = f(a)$ ფუნქციას შორის მოიცემა ასეთნაირად:

$$p_b(b) = db = p_a(a)da \Rightarrow df = \frac{p_a(a)da}{p_b(b)} \quad (6.2)$$



სურ. 6.14

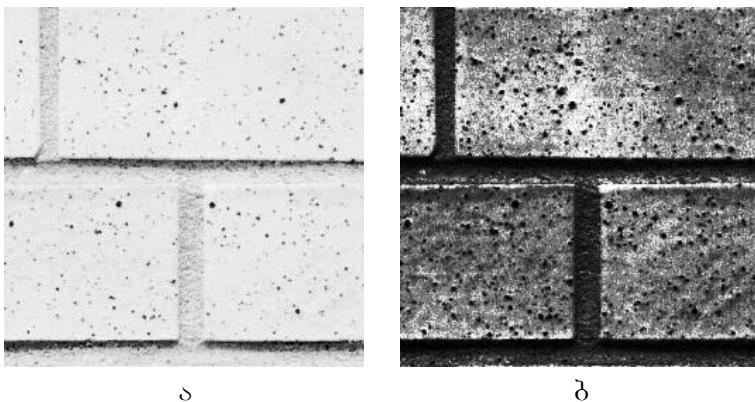
(6.2) გამოსახულებიდან ჩანს, რომ შესაფერისობა $b = f(a)$ ფუნქციისთვის ნიშნავს მის დიფერენცირებადობას და იმას, რომ $df/da \geq 0$. მაშინ, ზემოთქმულის შესაბამისად, ჰისტოგრამის გათანაბრების მოთხოვნა მდგომარეობს იმაში, რომ $p_b(b) = \text{constant}$. ეს უკანასკნელი კი ნიშნავს:

$$f(a) = (2^B - 1) \cdot P(a),$$

სადაც, $P(a)$ ინტენსივობის განაწილების ფუნქციაა. ამგვარად, კვანტირებული განაწილების ფუნქცია, ნორმირებული 0-დან

($2^B - 1$)-მდე, წარმოადგენს გარდაქმნის ცხრილს, აუცილებელს ჰისტოგრამის გათანაბრებისთვის.

ზემოაღწერილი კონტრასტის გაჭიმვისა და ჰისტოგრამის გაფართოების (გარდაქმნის) ოპერატორების გამოსახულებაზე მოქმედების შედეგები წარმოდგენილია 6.15 სურათზე, სადაც (ა) საწყისი, ხოლო (ბ) ჰისტოგრამის გათანაბრების შედეგად მიღებული გამოსახულებებია.



სურ. 6.15

ამავე დროს, უნდა აღინიშნოს ის ნაკლოვანებებიც, რაც თან ახლავს გარდაქმნის ფუნქციის გამოყენებას ჰისტოგრამის გაფართოებისას. კერძოდ, ადგილი აქვს, ერთი მხრივ, ინტენსივობის მადალაღბათური ანათვლების დიაპაზონის გაფართოებას, რაც დონეების მიხედვით ინტენსივობის კვანტირებისას იწვევს შეცდომებს და შესაბამისად გამოსახულებაზე არასასურველ ეფექტებს (მაგალითად “გაუპარსაობის” ეფექტი პროგრეტზე) და, მეორე მხრივ, დაბალაღბათური ანათვლების დიაპაზონის შევიწროებას, რაც გამოიწვევს “არატიპიური” ინტენსივობის მქონე დეტალებისთვის კონტრასტულობის დაკარ-

გვას. ასეთი ეფექტებისგან თავის ასარიდებლად გარდაქმნის ფუნქციას აგებენ არა რეალური, არამედ დაგლუვებული ჰისტოგრამის საშუალებით. ამის გამო ჰისტოგრამის გარდაქმნა მიახლოებითი ხდება.

6.3. მცირე ზომის არეების მოშორება გამოსახულებაზე

გამოსახულებაზე ხშირად გვხვდება მცირე ზომის არეები, რომლებიც წარმოშობილია ხმაურის მიზეზით ან არსობრივი თვალსაზრისით წარმოადგენს არასაინტერესო (უმნიშვნელო) დეტალს, რომლის მხედველობაში მიღება გამოსახულების აღწერისას საჭირო არ არის. სასურველია, რომ გამოსახულება გათავისუფლდეს ასეთი არეებისგან ცალკეულ პიქსელებზე ზემოქმედების ან ბმული კომპონენტებიდან შესაბამისი კომპონენტების მოშორების გზით.

გამოსახულების ხმაურისგან დამახინჯების მიზეზები შეიძლება სხვადასხვა წარმომავლობის იყოს. მაგალითად, ვიდეოგადამწოდის ხმაური, ხმაური გამოწვეული ოპტიკური სისტემის მიერ გამოსახულების ფორმირებისას, ფოტომასალის მარცვლოვნება, ხმაური გამოსახულების გადამცემ არხში და ა.შ. მათი გავლენის მინიმიზებისთვის შეიძლება გამოყენებულ იქნეს სტატისტიკური ფილტრაციის კლასიკური მეთოდები. მეორე მიდგომა გულისხმობს სივრცითი დამუშავების ევრისტიკული მეთოდების გამოყენებას. ხმაური, რომელიც განპირობებულია ვიდეოგადამწოდების და გადამცემის არხების მიერ, გამოსახულებაზე ვლინდება როგორც განცალკევებული ელემენტების იზოლირებული ცვლილებები, რომელთა შორის სივრცით კორელაციას ადგილი არ აქვს. ასეთ დროს პიქსელების დამახინჯებული მნიშვნელობები საგრძნობლად განსხვავდება თავიანთი მეზობლებისგან. მარტივი მიდგომა, რომელიც

ამ დროს შეიძლება გამოყენებულ იქნეს, მდგომარეობს ზღურბლური მეთოდის გამოყენებაში, რომლის დროსაც ხდება გამოსახულების ყოველ პიქსელში სიკაშკაშის მნიშვნელობის გაზომვა და თუ ის აჭარბებს მისი უშუალო მეზობლებისგან შემდგარი ჯგუფის საშუალო სიკაშკაშეს ზღურბლის გარკვეული სიდიდით, მაშინ ამ პიქსელში სიკაშკაშე იცვლება ჯგუფის საშუალო სიკაშკაშის სიდიდით. ასევე ხმაურის ჩახშობა შესაძლებელია დაბალსიხშირული სივრცული ფილტრაციის გამოყენებით იმ მიზეზით, რომ ხმაური სივრცულად დეკორელირებულია და მისი სპექტრი შეიცავს უფრო მაღალ სივრცით სიხშირეებს, ვიდრე ჩვეულებრივი გამოსახულების სპექტრი.

6.3.1. “მარილი და წიწაკა” ტიპის ხმაურისგან გაწმენდა

გამოსახულების ზღურბლური ბინარიზაციის დროს ხშირად წარმოიშობა ხმაური - ე.წ. “მარილი და წიწაკა”. “მარილი” გამოსახულებაზე შეესაბამება მუქი ფერის არეების იმ პიქსელებს, რომლებმაც გამოსახულების ბინარიზაციისას რაღაც მიზეზების გამო შეძლეს გადაელახათ ღია ფერის პიქსელებისთვის დადგენილი ზღურბლი, ხოლო “წიწაკა” კი პირიქით – შეესაბამება ღია ფერის არეების იმ პიქსელებს, რომელთა მნიშვნელობები ბინარიზაციის ზღურბლზე ნაკლები აღმოჩნდა. ასეთ დროს ბინარულ გამოსახულებაზე არსებულ ღია ფერის არეებში გვაქვს 1-ის ტოლი სიკაშკაშის მუქი პიქსელები ან, პირიქით, 1-ის ტოლი სიკაშკაშის ღია პიქსელები მუქი ფერის არეების შიგნით. სხვადასხვა რაოდენობის “მარილი და წიწაკა” ტიპის ხმაურის შემცველი გამოსახულება წარმოდგენილია 6.16 სურათზე, სადაც თეთრი წერტილები აღნიშნავს “მარილს”, ხოლო შავი წერტილები კი – “წიწაკას”. 6.16 სურათზე მოცემულია: (ა) საწყისი გამოსახუ-

ლება, (ბ) გამოსახულება, რომელზედაც “მარილი და წიწაკა” წერტილების რაოდენობა პიქსელთა საერთო რაოდენობის 1%-ს შეადგენს, (გ) “მარილი და წიწაკა” წერტილების რაოდენობა პიქსელთა საერთო რაოდენობის 10%-ს შეადგენს, და (დ) “მარილი და წიწაკა” წერტილების რაოდენობა პიქსელთა საერთო რაოდენობის 20%-ს შეადგენს.



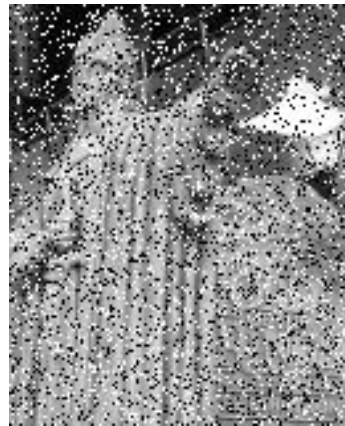
ა



ბ



გ



დ

სურ. 6.16

გამოსახულების მასზე არსებული ასეთი 1-ის ტოლი სიკაშკაშის ანომალური პიქსელებისგან გასაწმენდად, რომლებიც პრაქტიკულად განლაგებულია ერთგვაროვანი არეების შიგნით, საჭიროა მხედველობაში მივიღოთ მათი ურთიერთკავშირები უახლოეს მიდამოში განლაგებულ მცირე რაოდენობის პიქსელებთან (მაგალითად, უახლოესი სტრიქონებისა და სვეტების პიქსელებთან). ბინარული გამოსახულების შემთხვევაში პიქსელების იზოლირებული მნიშვნელობები 1 და 0 იცვლება საწინააღმდეგო მნიშვნელობებით. ფაქტობრივად ამით ხდება დისკრეტიზაციის დროს გამოსახულებაზე წარმოშობილი ხმაურის მოშორება, ანუ გამოსახულების გამარტივება მასზე არსებული მცირე არეების მოშორებით, რომლებიც გამოსახულებაზე შეიძლება წარმოადგენდეს რაიმე მცირე, უმნიშვნელო დეტალს ან დეფექტს ქაღალდის გამოსახულებაზე. ერთ-ერთი მეთოდი, რომელსაც იყენებენ ასეთი გამოსახულების დაგლუვებისთვის, არის **ლოგიკური გასაშუალოების** ანუ **ლოგიკური დაგლუვების** მეთოდი. ის ეყრდნობა დაშვებას, რომ გამოსახულების პიქსელები, რომლებიც მოთავსებულია გამასაშუალოებელი ფანჯრის საზღვრებში, განიხილება როგორც ბულის ანუ ლოგიკური ცვლადები. აქედან გამომდინარე, მოცემულ წერტილში ინტენსივობის დაგლუვებული ფუნქციის სიდიდე შეიძლება განისაზღვროს ამ ცვლადებით განსაზღვრული ბულის ფუნქციით. ვთქვათ, გვინდა ავაგოთ შემდეგი მახასიათებლების მქონე დაგლუვების ლოგიკური ოპერატორი:

- პიქსელში მნიშვნელობა “0” იცვლება მნიშვნელობით “1” მხოლოდ და მხოლოდ მაშინ, თუ ყველა მეზობელ პიქსელში მნიშვნელობა არის “1”;

- პიქსელში მნიშვნელობა “1” იცვლება მნიშვნელობით “0” მხოლოდ და მხოლოდ მაშინ, თუ ყველა მეზობელ პიქსელში მნიშვნელობა არის “0”.

ამ მოთხოვნების დასაკმაყოფილებლად საჭიროა შესაბამისი ლოგიკური ოპერატორის აგება. მაგალითად, ვთქვათ, გვაქვს დაგლუვების ლოგიკური ოპერატორი, რომელიც აკმაყოფილებს ზემომოყვანილ პირობებს და იყენებს (3×3) განზომილების ფანჯარას (სურ. 6.17), რომლის ელემენტებია

a	b	c
d	e	f
g	h	i

სურ. 6.17

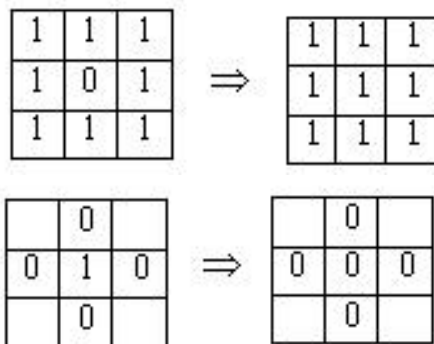
მაშინ ფანჯრის ცენტრალური პიქსელისთვის ინტენსივობის ახალი მნიშვნელობა e' გამოითვლება ფორმულით

$$e' = \bar{e}(abcdfghi)Ve(aVbVcVdVfVgVhVi),$$

რომლის მიხედვითაც $e' = 1$ მხოლოდ და მხოლოდ მაშინ, თუ $e = 0$ და ყველა მეზობელ პიქსელში კი გვაქვს 1; ან თუ $e = 1$ და მეზობელი პიქსელების მნიშვნელობებიდან ერთი მაინც უდრის 1. მაგრამ ზოგიერთ შემთხვევაში ეს იზოლირებული პიქსელები შეიძლება არ წარმოადგენდეს ხმაურის შედეგს, არამედ შეესაბამებოდეს უფრო მსხვილი ობიექტების ფონზე

არსებულ მცირე კონტრასტულ დეტალებს (მაგალითად, ქვის ლოდი დიდ მინდორში, პატარა მინდორი ტყის მასივში, დიდი პერანგზე და ა.შ.). გამოსახულების აღწერისას შეიძლება საჭირო გახდეს ასეთი მცირე დეტალების იზნორირება.

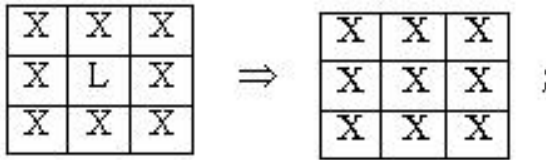
“მარილი და წიწაკა” ტიპის ხმაურისგან ბინარული გამოსახულების გასაწმენდად ასევე იყენებენ ტრაფარეტებს, რომელთა მოქმედება ნახევნების 6.18 სურათზე.



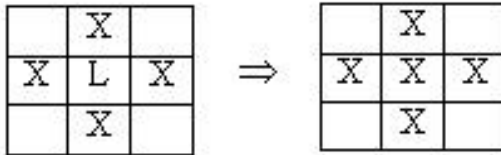
სურ. 6.18

კერძოდ, ოპერაციები, რომლებიც სრულდება შესასვლელ გამოსახულებაზე, წარმოდგენილია ტრაფარეტების საშუალებით. როგორც სურათიდან ჩანს, თუ შესასვლელი გამოსახულების მიდამო ემთხვევა მარცხნივ მყოფ ტრაფარეტს, მაშინ ეს მიდამო იცვლება მარჯვენა ტრაფარეტით მოცემული მიდამოთი. აღწერილ მეთოდში გამოიყენება მხოლოდ ორი ტრაფარეტი.

თუ შესასვლელი გამოსახულება წარმოადგენს მარკირებულ გამოსახულებას, რომლის ფორმირება მოხდა ზღურბლების ნაკრების გამოყენებით ან კლასიფიკაციის შედეგად, მაშინ მისი ხმაურისაგან გასაწმენდად შეიძლება გამოყენებულ იქნეს განზოგადებული სახის ტრაფარეტი (სურ. 6.19).



ა



ბ

სურ. 6.19

სურათიდან ჩანს, რომ ნებისმიერ L ჭდის მქონე პიქსელს, რომელიც იზოლირებულია ერთნაირი ჭდეებით მარკირებული თავისი 8-ბმული მეზობლებით (სურ. 6.19,ა), ენიჭება ჭდე X . აქ L ჭდე შეიძლება წარმოადგენდეს გამოსახულებაზე გამოყენებულ k ჭდეებს შორის ნებისმიერს. უნდა აღინიშნოს, რომ 4-ბმული მეზობლობის გამოყენების შემთხვევაში (სურ. 6.19,ბ) დიაგონალური ელემენტები მხედველობაში არ მიიღება.

გარდა ამისა გამოიყენება სხვა მიდგომაც, კერძოდ, სასაზღვრო პიქსელების მნიშვნელობები იცვლება ფონური ელემენტების პიქსელების მნიშვნელობებით.

6.3.2. მცირე კომპონენტების მოშორება

ჩვენ მიერ განხილული იყო ბინარულ გამოსახულებაზე ბმული კომპონენტების გამოყოფის ხერხები და ასევე ის ნიშანთვისებები, რომელთა გამოთვლაც შესაძლებელია პიქსელთა

იმ სიმრავლეზე, რომლებიც ერთი ბმულ კომპონენტს ქმნის. საზოგადოდ, გამოსახულების აღწერა შეიძლება შეიცავდეს კომპონენტების სიმრავლეს, რომელთაგან თითოეული შეესაბამება გამოსახულების ფონიდან გამოყოფილ რომელიმე არეს და ნიშან-თვისებათა სიმრავლეს, გამოთვლილს თითოეული ამ კომპონენტისთვის. გამოსახულების “მარილი და წიწკა” ტიპის ხმაურისგან გაწმენდის ალგორითმმა გამოთვლილი ნიშან-თვისებების გათვალისწინებით შესაძლებელია ამ აღწერიდან მოახდინოს ნებისმიერი კომპონენტის წაშლა. მაგალითად, შეიძლება წაიშალოს მცირე რაოდენობის პიქსელების შემცველი კომპონენტები ან ძალიან წვრილი კომპონენტები. გამოსახულების ამგვარმა დამუშავებამ შეიძლება ხელი შეუწყოს არის საზღვართან ზოგიერთი ხმაურის შემცველი არეების წაშლას.

6.4. გამოსახულების დაგლუვება

ზოგადად, რეალური გამოსახულება შეიძლება წარმოვიდგინოთ როგორც შედგენილი ორი გამოსახულებისგან: ძირითადი - იდეალური გამოსახულებისგან და რაიმე შემთხვევითი ხმაურისგან წარმოშობილი გამოსახულებისგან. მაგალითად, ასეთი გამოსახულების მარტივი მოდელის ასაგებად შეიძლება დავუშვათ, რომ გამოსახულებაზე რაიმე ერთგვაროვანი ობიექტის შესაბამისი არის პიქსელების მნიშვნელობები ტოლია $g_r + N(0, \sigma)$, სადაც g_r - ინტენსივობის მნიშვნელობაა იდეალურ გამოსახულებაზე, ხოლო $N(0, \sigma)$ - გაუსის (ნორმალური) განაწილების მქონე ხმაურია გამოსახულებაზე.

ხმაურისგან გამოსახულების გასაწმენდად ხშირად იყენებენ გასაშუალოების ოპერაციას, კერძოდ, მოცემულ პიქსელში ინტენსივობის მნიშვნელობა იცვლება მის უშუალო

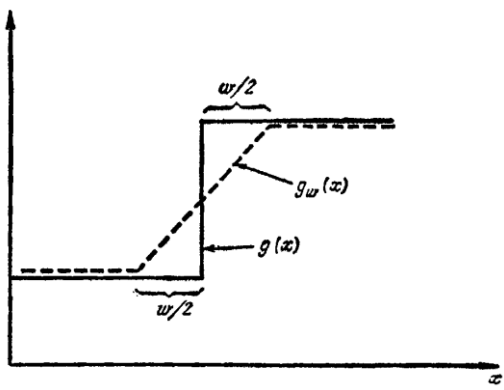
მიდამოში ინტენსივობის ფუნქციის საშუალო მნიშვნელობით. ამასთან, აღნიშნული ოპერატორის მოქმედების შედეგად გამოსახულების აღმწერი ინტენსივობის ფუნქცია ისე უნდა დაგლუვდეს, რომ მოხდეს საშუალოდ ყველაზე უფრო ძლიერი ხმაურის ჩახშობა, მაგრამ ისე, რომ გამოსახულებაზე არ დაიკარგოს (შესამჩნევი დარჩეს) ინტენსივობის მნიშვნელობების ვარდნები. ზოგადად, ეს შეიძლება ასე ჩაიწეროს:

$g_w(i, j) = g(i, j)$ -ის რაიმე მიდამოს საშუალო მნიშვნელობას.

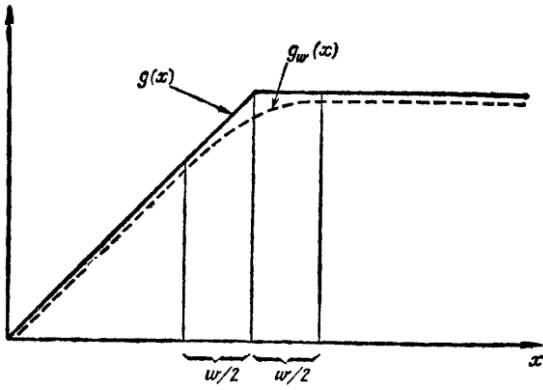
სადაც, $g_w(i, j)$ - დაგლუვებული გამოსახვლელი, ხოლო $g(i, j)$ შესასვლელი გამოსახულებებია. ინტენსივობის ფუნქციის გასაშუალოების პროცესს **ფუნქციის რეგულარიზაციას** უწოდებენ. გასაშუალოებას ის უარყოფითი ეფექტი მოსდევს თან, რომ გამოსახულებაზე დეტალები უფრო დეფოკუსირებული ანუ ბუნდოვანი ხდება.

ინტენსივობის ერთგანზომილებიანი უწყვეტი ფუნქციის შემთხვევაში რეგულარიზაციის ორი მაგალითი წარმოდგენილია 6.20, ა, ბ სურათზე, სადაც $g(x)$ ინტენსივობის ფუნქციაა, ხოლო $g_w(x)$ კი მისი დაგლუვებული ვარიანტია. სურათებიდან ჩანს, რომ რეგულარიზაცია მართლაც ახდენს ფუნქციების გარკვეული აზრით, დაგლუვებას. კერძოდ, თუ $g(x)$ წყვეტილია, მაშინ ფუნქცია $g_w(x)$ უწყვეტია და თუ $g(x)$ უწყვეტი ფუნქციაა, მაშინ $g_w(x)$ ფუნქციას აქვს უწყვეტი წარმოებულობა. როგორც ვხედავთ, $g_w(x)$ ფუნქციის “გადასასვლელი არის” სიგანე დამოკიდებულია ფანჯრის ზომაზე. კერძოდ, ფანჯრის w ზომის ზრდასთან ერთად გამოსახულება უფრო ბუნდოვანი (დეფოკუსირებული) ხდება და, პირიქით, რაც მცირეა ფანჯრის w ზომა, მით უფრო ზუსტად ახდენს $g_w(x)$ ფუნქცია $g(x)$

ფუნქციის აპროქსიმაციას. რეგულარიზებულ $g_w(x)$ ფუნქციას ასევე ხშირად $g(x)$ ფუნქციის მცოცავ ან მიმდინარე საშუალოს უწოდებენ.



ა



ბ

სურ. 6.20

თუ $I[i, j]$ ინტენსივობის დისკრეტული ფუნქციაა, მაშინ მისი დაგლუვებული ფუნქცია $I_w[i, j]$ განისაზღვრება ფორმულით

$$I_w[i, j] = \frac{1}{A_w} \sum_{w(i, j)} \sum I[m, n],$$

სადაც $w(i, j)$ არის გამოსახულების $[i, j]$ ელემენტის A_w ფართის მქონე ფანჯარა (მიდამო). კონკრეტულ შემთხვევაში ფანჯარა შეიძლება იყოს მარტივი გეომეტრიული ფორმის, მაგალითად, მართკუთხა, ფუძით $(2b + 1)$ და სიმაღლით $(2h + 1)$. მაშინ ზედა ფორმულა მიიღებს სახეს

$$I_w[i, j] = \frac{1}{(2b+1)(2h+1)} \sum_{-b \leq m \leq b} \sum_{-h \leq n \leq h} I[i + m, j + n].$$

მაგალითად, დაგლუვების ფილტრს, რომელიც გამოსახველი გამოსახულების ყოველ პიქსელს ანიჭებს მოცემული პიქსელის (5×5) ფანჯარაში არსებული 25 პიქსელის ინტენსივობის მნიშვნელობების საშუალოს, ექნება ასეთი სახე:

$$I_w[i, j] = (\sum_{m=-2}^2 \sum_{n=-2}^2 I[i + m, j + n]) / 25.$$

უნდა აღინიშნოს, რომ გამოსახველი გამოსახულების დეფოკუსირება ანუ გაბუნდოვანება ყოველთვის არ არის სასარგებლო. საერთოდ, გამოსახულების ანალიზისას, ცხადია, გვინდა, რომ გამოსახულება ვიზუალურად გამოიყურებოდეს კარგად, ანუ იყოს მკაფიო და მკვეთრი. ერთ-ერთ შემთხვევას, როცა ის სასარგებლოა, წარმოადგენს დისკრეტული ბინარული გამოსახულება.

როგორც ვხედავთ, დაგლუვების ოპერაცია განხორციელდა პიქსელის მართკუთხა მიდამოში თანაბარშეწონილი მნიშვნელობების გამოყენებით. ასეთ ფილტრს **გამასაშუალოებელი ფილტრი** ეწოდება. მაგრამ თანაბარი წონების მინიჭების ნაცვლად უპრიანია პიქსელებისთვის ისეთი წონების მინიჭება,

რომელთა სიდიდე მცირდება ცენტრალურ პიქსელსა $I[i_c, j_c]$ და შესასვლელ პიქსელებს შორის მანძილის ზრდასთან ერთად. ასეთი თვისებით ხასიათდება **გაუსის ფილტრი**. მისი გამოყენებისას $[i, j]$ პიქსელის მნიშვნელობის შეწონვა ხდება წონითი

$$g(x, y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{d^2}{2\sigma^2}}$$

ფუნქციის საშუალებით, სადაც, $d = \sqrt{(i - i_c)^2 + (j - j_c)^2}$ წარმოადგენს მანძილს გამოსასვლელ გამოსახულებაზე არსებულ $[i, j]$ და მიდამოს ცენტრალურ $[i_c, j_c]$ პიქსელებს შორის.

6.5. მედიანური ფილტრი

გასაშუალოების ოპერაცია კარგ შედეგს იძლევა მაშინ, როცა პიქსელის მიდამო ერთგვაროვანია. ასეთ მიდამოში ხმაური აღიწერება ნულოვანი მათემატიკური მოლოდინის მქონე განაწილებით. მაგრამ როცა ასეთი ოპერატორი გამოიყენება ორი ერთგვაროვანი არის საზღვარზე, ანუ იმ ადგილას, სადაც ადგილი აქვს ინტენსივობის მკვეთრ ვარდნას, მაშინ ამ არეების მნიშვნელობების გამოყენებას მიყვავართ საზღვრის წაშლისკენ (გადარეცხვისკენ).

ხმაურისგან გამოსახულების გაწმენდის კიდევ ერთ მეთოდს წარმოადგენს **მედიანური ფილტრაცია**, რომელიც მიეკუთვნება სიგნალების არაწრფივი დამუშავების მეთოდს.

ერთგანზომილებიანი მედიანური ფილტრი წარმოადგენს გამოსახულებაზე მოსრიადე ფანჯარას, რომელიც მოიცავს პიქსელთა კენტ რაოდენობას (3,5,7,...). მედიანური ფილტრაციის შედეგად ფანჯრის ცენტრალური პიქსელის მნიშვნელობა

იცვლება ფანჯარაში მოხვედრილი პიქსელების მნიშვნელობათა მედიანით.

ვთქვათ, $A[i]$ წარმოადგენს n ნამდვილ რიცხვთა დალაგებულ, ზრდის ან კლების მიხედვით, დისკრეტულ მიმდევრობას, სადაც n -კენტია. მაშინ მისი მედიანა წარმოადგენს მიმდევრობის იმ ელემენტს, რომლისთვისაც ამ მიმდევრობაში არსებობს სიდიდით მისი ტოლი ან მასზე ნაკლები $A[(n-1)/2]$ ელემენტი და სიდიდით მისი ტოლი ან მასზე მეტი $A[(n-1)/2]$ ელემენტი. მაგალითად, 5-ელემენტიან ფანჯარაში მოხვდა გამოსახულების პიქსელების შემდეგი მნიშვნელობები: 60, 80, 160, 90, 130. მაშინ მიმდევრობის მედიანა უდრის 90 და შესაბამისად ფანჯრის ცენტრალურ ელემენტში ის ჩაიწერება. უნდა აღინიშნოს, რომ მედიანური ფილტრი ერთ შემთხვევაში შეიძლება აღმოჩნდეს ხმაურის ჩამხშობი, ხოლო მეორე შემთხვევაში კი მის მოქმედებას შეიძლება მოჰყვეს გამოსახულების ხარისხის გაუარესება. თუ ჩვენს მაგალითში პიქსელის მნიშვნელობა - 160 შეესაბამება ხმაურს, მაშინ მედიანური ფილტრაციის შედეგად მოხდება გამოსახულების გაუმჯობესება, და, პირიქით, თუ ის არ შეესაბამება ხმაურს, მაშინ ადგილი ექნება გამოსახულების ხარისხის გაუარესებას. მოვიყვანოთ $A[i]$ მედიანის რამდენიმე თვისება. კერძოდ:

$$\text{med}\{K \cdot A[i]\} = K \text{med}\{A[i]\};$$

$$\text{med}\{K + A[i]\} = K + \text{med}\{A[i]\},$$

$$\text{med}\{A[i] + B[i]\} \neq \text{med}\{A[i]\} + \text{med}\{B[i]\}$$

სადაც, K მუდმივია.

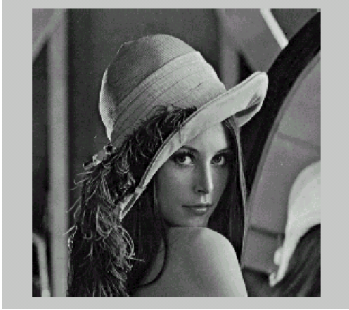
თუ მედიანური ფილტრის გამოყენების ეფექტი უმნიშვნელოა, მაშინ ხდება ფანჯრის ზომის გაზრდა 3 პიქსელიდან 5 პიქსელამდე. ასე იქცევიან მანამ, სანამ ფილტრის მოქმედება არ

იწყებს გამოსახულების გაუარესებას. ზოგადად კი ის არეები გამოსახულებაზე, რომლებიც ფილტრის გავლის შემდეგ არ შეიცვალა, შეუცვლელი რჩება ფილტრის ხელმეორედ გამოყენების დროსაც.

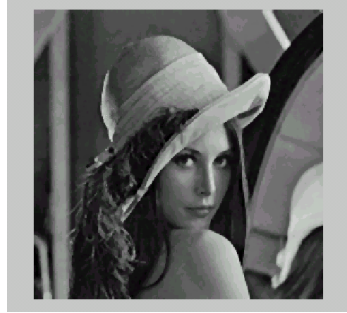
ერთგანზომილებიანი ფილტრის გარდა, შესაძლებელია ორგანზომილებიანი ფილტრის გამოყენებაც, ანუ როცა გამოსახულებაზე გადაადგილება მოხდება ორგანზომილებიანი სასურველი ფორმის ფანჯრის საშუალებით, მაგალითად, მარტკუთხა ან ჯვარედინი. ცხადია, ($L \times L$) ფანჯრის ზომის მედიანური ფილტრი უფრო ეფექტური იქნება ხმაურის ჩასახშობად, ვიდრე ($L \times 1$) ფანჯრის ზომის ერთგანზომილებიანი ფილტრები გამოყენებული მიმდევრობით ჯერ ჰორიზონტალურად და შემდეგ ვერტიკალურად. თუმცა ამ დროს ადგილი აქვს გამოსახულების ხარისხის მნიშვნელოვან გაუარესებას.

მედიანური ფილტრის გამოყენების შედეგები სხვადასხვა ფანჯრის შემთხვევაში ნაჩვენებია 6.21 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) ფანჯარა ზომით (3×3), გ) ფანჯარა ზომით (15×15).

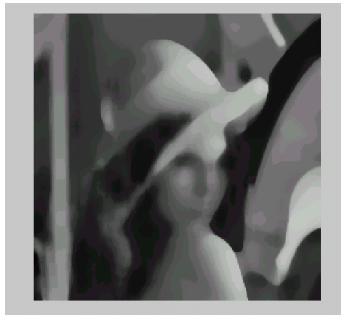
გამოსახულების დაგლუვებას მედიანური ფილტრის გამოყენებით უპირატესობასთან ერთად, რაც მის სიმარტივეში მდგომარეობს, ნაკლიც გააჩნია. კერძოდ, მის გამოყენებას მიუყვართ გამოსახულების დეტალების საზღვრების წაშლამდე.



ა



ბ



ბ

სურ. 6.21

VII თავი. კიდის აღმოჩენა გამოსახულებაზე

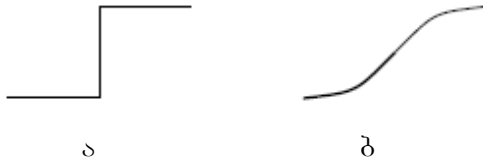
7.1. კიდის არსი გამოსახულებაზე

კიდის აღმოჩენა (edge detection) გამოსახულებაზე წარმოადგენს ძალიან მნიშვნელოვან ნაბიჯს გამოსახულების მახასიათებელი ნიშან-თვისებების გამოვლენის მიმართულებით. ის წარმოადგენს საზღვრის ლოკალიზაციის პროცესს. საზღვარი შეიცავს არსობრივ ნიშან-თვისებებს და მნიშვნელოვან ინფორმაციას გამოსახულების შესახებ. გამოსახულებაზე ის ერთმანეთისგან ყოფს ორ ერთგვაროვან, განსხვავებული სიკაშკაშის მნიშვნელობების მქონე არეს. ამგვარად, საზღვარი გამოსახულებაზე წარმოადგენს ადგილს, სადაც ხდება სიკაშკაშის, ან გამოსახულების სხვა რომელიმე პარამეტრის (ფერი, ტექსტურა) მკვეთრი, ნახტომისებური ლოკალური ცვლილება – წყვეტა.

განსაზღვრება 7.1. საზღვარი გამოსახულებაზე წარმოადგენს ისეთი პიქსელების სიმრავლეს, რომელთა უახლოესი მიდამოში (მეზობლობაში) ადგილი აქვს სიკაშკაშის მნიშვნელობების მკვეთრ, ნახტომისებურ ცვლილებას – ვარდნას.

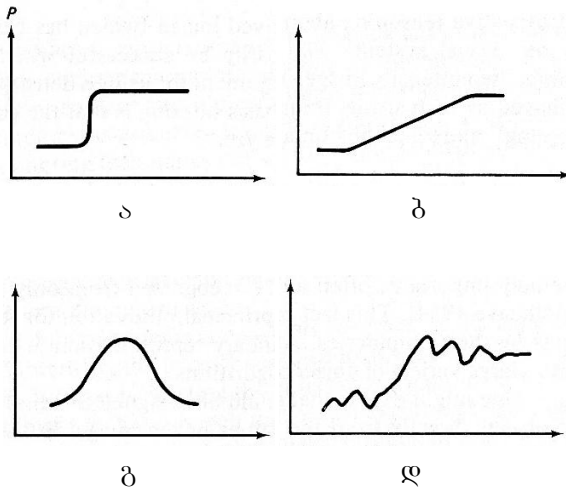
აღნიშნული პიქსელების სიმრავლეს გამოსახულებაზე კიდევს ან კიდის წერტილებს უწოდებენ. ამგვარად, კიდევს გამოსახულებაზე ასახავს პიქსელისა და მისი უახლოესი მეზობლების ლოკალურ თვისებებს. კიდის წერტილების ერთობლიობა ქმნის ობიექტის კონტურს – საზღვარს ორ ერთგვაროვან არეს შორის. კიდის წერტილში ინტენსივობა მკვეთრად იცვლება ობიექტის კონტურის განივი მიმართულებით. აუცილებელი არ არის, რომ კიდევს გამოსახულებაზე ემთხვეოდეს ფიზიკურ კიდევს. თუ გამოსახულების კიდევს ინტენსივობის

წვევების ადგილებია, ფიზიკური კიდევები ზედაპირის წვევების ადგილებს წარმოადგენს. მაგრამ, მაგალითად, ჩრდილების კიდევები არ შეესაბამება ზედაპირის წვევების ადგილებს. იდეალური კიდევ ნაჩვენებია 7.1,ა სურათზე, ხოლო გადარეცხილი – 7.1,ბ სურათზე.



სურ. 7.1

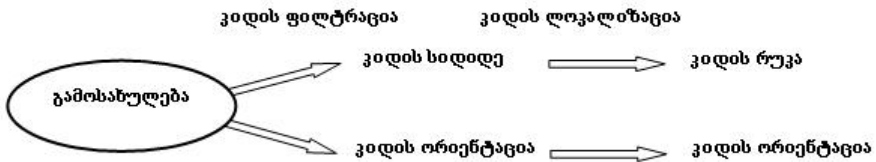
თეორიულად კიდევებს შეიძლება ჰქონდეს სხვადასხვა სახის პროფილი (სურ. 7.2,ა,ბ,გ), თუმცა რეალობაში ხმაურის გავლენის გამო (სურ. 7.2,დ) პროფილი შეიძლება წარმოჩნდეს როგორც სხვადასხვა სახის პროფილებისგან შედგენილი კომპოზიცია.



სურ. 7.2

ზემოთქმულიდან გამომდინარე, ბუნებრივია ინტერესი გამოსახულებაზე საზღვრის აღმოჩენის მიმართ, რადგანაც ის შეიძლება განპირობებული იყოს სხვადასხვა მიზეზით. კერძოდ, სიკაშკაშის მნიშვნელობების მკვეთრი ცვლილება დამახასიათებელია ობიექტის საზღვრებისთვის, როცა, ვთქვათ, მუქი ობიექტი მოთავსებულია ღია ფონზე ან პირიქით. სიკაშკაშის მკვეთრ ცვლილებებს ასევე იწვევს სცენის ობიექტების არეკვლის უნარის ცვლილებებიც, რომლებიც ერთმანეთისგან შეიძლება საკმაოდ განსხვავდებოდეს – მაგალითად, ზოლები ზებრას და ლაქები ლეოპარდის სხეულებზე; სიკაშკაშის ცვლილებები შეიძლება აგრეთვე გამოიწვიოს ობიექტებიდან დაცემულმა ჩრდილებმა; და ბოლოს, სიკაშკაშის ცვლილებებს შეიძლება იწვევდეს ობიექტის ზედაპირის ორიენტაციის ცვლილებებიც.

გამოსახულებაზე კიდეების აღმოჩენის პროცედურა შედგება რამდენიმე ნაბიჯისგან (სურ. 7.3).



სურ. 7.3

როგორც სურათიდან ჩანს, პირველ ნაბიჯზე უნდა ჩატარდეს კიდის ფილტრაციის ოპერაცია (edge filtering), რაც გულისხმობს კიდის სიდიდის (edge magnitude), როგორც ლოკალური კონტრასტის საზომის, და ორიენტაციის (edge orientation) განსაზღვრას, ხოლო შემდეგ კი კიდის ლოკალიზაციას (edge localization).

ამოცანები, რომელთა გადაწყვეტაც საჭიროა კიდის ლოკალიზაციისთვის, შემდეგია:

- კიდეების ხმაურისგან გაწმენდა;
- მცდარი კიდეების მოშორება, წვრილი კონტურების მიღება;
- კიდეების ბინარული რუკის მიღება.

უნდა აღინიშნოს, რომ კიდის ხმაურისგან გაწმენდა შეიძლება განხორციელდეს კიდის ფილტრაციის ოპერაციამდე. კიდის აღმოჩენის ნაბიჯების კონკრეტულ გამოსახულებაზე გამოყენების შედეგები მოცემულია 7.4 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) კიდის სიდიდე, (გ) კიდის ორიენტაცია, (დ) კიდის ბინარული რუკა.



ა



ბ

სურ. 7.4



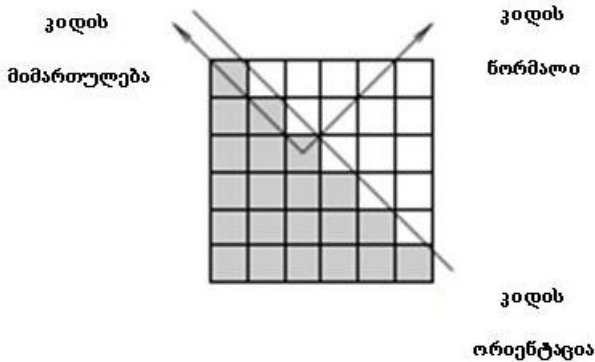
ბ



დ

სურ. 7.4

როგორც ზემოთ აღვნიშნეთ, კოდის აღმოჩენისთვის მნიშვნელოვანია განვსაზღვროთ მისი ორიენტაცია და მიმართულება (სურ. 7.5).



სურ. 7.5

წარმოდგენილ სურათზე მოცემულია:

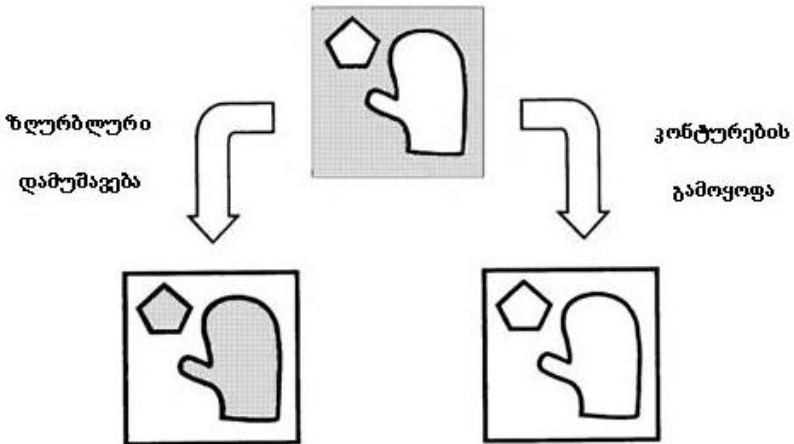
- **კიდის ნორმალი:** კიდის წერტილში სიკაშკაშის მნიშვნელობის მაქსიმალური ცვლილების მიმართულება (ერთეულოვანი ვექტორი კიდის მართობულია);
- **კიდის მიმართულება:** კონტურის მიმართ გავლებული მიმართული მხები (კიდის პარალელური ერთეულოვანი ვექტორი);
- **კიდის ორიენტაცია:** კუთხე (რადიანებში ან გრადუსებში) სიკაშკაშის მნიშვნელობის მაქსიმალური ცვლილების მიმართულებასა და აბსცისათა ღერძს შორის.

ამგვარად, კიდე გამოსახულების (X, Y) სიბრტყეზე წარმოადგენს ლოკალური კონტრასტის არეს, სადაც უწყვეტი ინტენსივობის $f(x, y)$ ფუნქციის გრადიენტის სიდიდე დიდია, ანუ ადგილი აქვს მაღალ კონტრასტს. კიდეების აღმოჩენა გვეხმარება გამოსახულების შესახებ სასარგებლო ინფორმაციის გამოყოფაში.

7.2. კონტურები

როგორც უკვე ვიცით, გამოსახულებაზე სიკაშკაშის მიმართ ერთგვაროვანი არეების გამოსაყოფად, ანუ გამოსახულების ბინარიზაციისათვის, გამოიყენება ზღურბლური დამუშავება (სურ. 7.6). მიღებული არეების გეომეტრიული მახასიათებლები კი წარმოადგენს მნიშვნელოვან ნიშან-თვისებებს გამოსახულებაზე არსებული ობიექტების კლასიფიკაციისთვის და გამოსახულების, როგორც მთლიანის აღქმისთვის. მაგრამ ხშირად საზღვრების საშუალებით გამოყოფილ ერთგვაროვან არეებზე უფრო მნიშვნელოვანია ამ არეების საზღვრები – **კონტურები**. ექსპერიმენტებით დადგენილია, რომ ბიოლოგიური სისტემები მხედველობის ანალიზატორის მიერ გარესამყაროს ობიექტების

აღქმისას ყურადღებას აქცევს არა განსხვავებას ობიექტების სიკაშკაშეებს შორის, არამედ ობიექტების მოხაზულობას – კონტურებს. კონტურების გამოყოფის ამოცანა მდგომარეობს გამოსახულებაზე ობიექტების კონტურების მოხაზულობის პოვნაში ბინარული გამოსახულების აგების გზით.

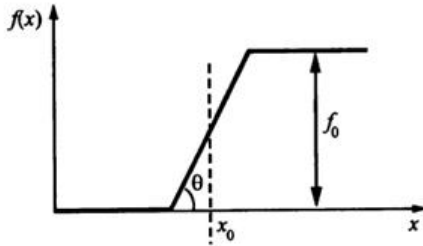


სურ. 7.6

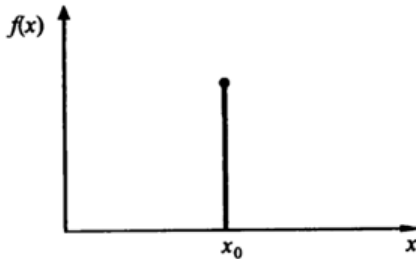
კონტურის არსის გასაგებად განვიხილოთ გამოსახულებაზე კონტურის შემცველი რომელიმე არე. ვთქვათ, სიკაშკაშის მნიშვნელობების ვარდნას ერთგანზომილებიან შემთხვევაში აქვს 7.7,ა სურათზე მოცემული სახე. აქ $f(x)$ სიკაშკაშის ფუნქციაა, ხოლო სიკაშკაშის ვარდნის მახასიათებლებს კი წარმოადგენს: f_0 - ვარდნის სიმაღლე, θ - დახრის კუთხე და x_0 - დახრილი მონაკვეთის ცენტრის კოორდინატი.

სიკაშკაშის ვარდნა მაშინ ჩაითვლება კონტურად, თუ მისი სიმაღლე f_0 და დახრის კუთხე θ გარკვეულ ზღურბლურ მნიშვნელობებს გადააჭარბებს. 7.7,ა სურათზე წარმოდგენილი

შემთხვევა გამოსახულებაზე სიკაშკაშის რეალურ ვარდნას შეესაბამება.



ა



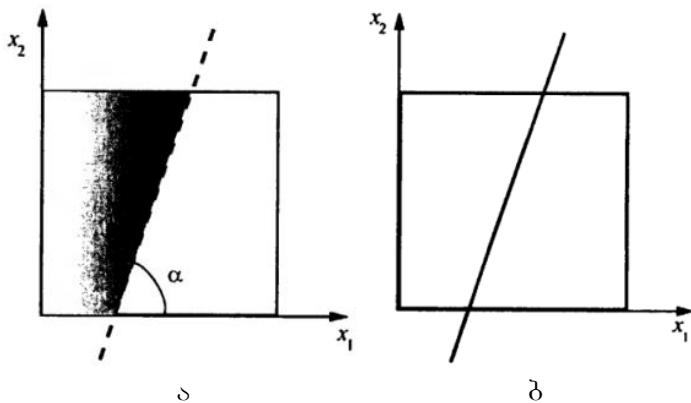
ბ

სურ. 7.7

იდეალური ვარდნის შემთხვევაში კი, კონტურის აღმომჩენმა დეტექტორმა მის არსებობაზე უნდა მიუთითოს დახრილი მონაკვეთის ცენტრში მოთავსებულ მხოლოდ ერთადერთ წერტილში (სურ. 7.7,ბ).

ორგანზომილებიანი სიკაშკაშის ფუნქციის შემთხვევაში კი სიკაშკაშის ვარდნის მახასიათებლებს ემატება კიდევ ერთი მახასიათებელი – ვარდნის ორიენტაცია სიბრტყეზე (კუთხე სიბრტყეზე). 7.8,ა სურათზე ნაჩვენებია შემთხვევა, როცა ლოკალურ არეში გვაქვს სწორხაზოვანი კონტური. კონტურის იდეალურმა აღმომჩენმა კი უნდა მოგვცეს კონტურის უსასრულოდ

წვრილი უწყვეტი წირი იმ არის ცენტრში, სადაც ადგილი აქვს სიკაშკაშის მნიშვნელობების ცვლილებას (სურ. 7.8,ბ), აქ α - კონტურის წირის დახრის კუთხეა x_1 ღერძის მიმართ.

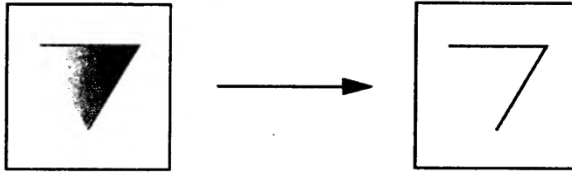


სურ. 7.8

მაგრამ ამავე დროს, რეალურ გამოსახულებებზე კონტურის აღმოჩენისას ყველაფერი იდეალურად არ არის. კერძოდ, შეიძლება წავაწყდეთ შემდეგ შემთხვევებს:

- მიღებული კონტური შეიძლება არ აღმოჩნდეს შეკრული წირი, არამედ ადგილი ჰქონდეს კონტურის წყვეტებს

ასეთი შეუკრავი კონტურის მაგალითი წარმოდგენილია 7.9 სურათზე. ეს იმაზე მიუთითებს, რომ გამოსახულების ამ ადგილებში სიკაშკაშის მნიშვნელობების ცვლილება საკმარისად მკვეთრად არ ხდება.

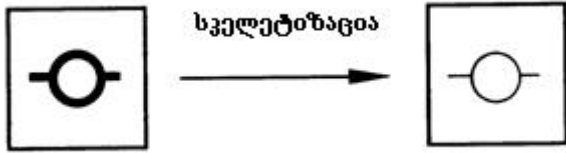


სურ. 7.9

გამოსახულებაზე ხმაურის არსებობის გამო კონტური შეიძლება შეცდომით აღმოჩნდეს იქ, სადაც ობიექტის საზღვარი რეალურად არ გადის. ასეთ ადგილებში რეალური კონტურის მოსაძებნად მიმართავენ გამოსახულების დამატებით დამუშავებას. კერძოდ: საზღვრების მიღვენების პროცედურას, ინტერპოლაციას, აღმოჩენილი კონტურული წირების ცალკეული “ფრაგმენტების (ნაგლეჯების)” სიმრავლეში ბმული კომპონენტების გამოვლენას და ა.შ.

- კონტურების გამოვლენისას მათი გადარეცხვის, ხმაურის ან თვით აღმომჩენი ალგორითმის ნაკლოვანების გამო შეიძლება ადგილი ჰქონდეს არა მარტო წვევებებს, არამედ მიღებული კონტურის წირების სისქე დასაშვებზე მეტი აღმოჩნდეს.

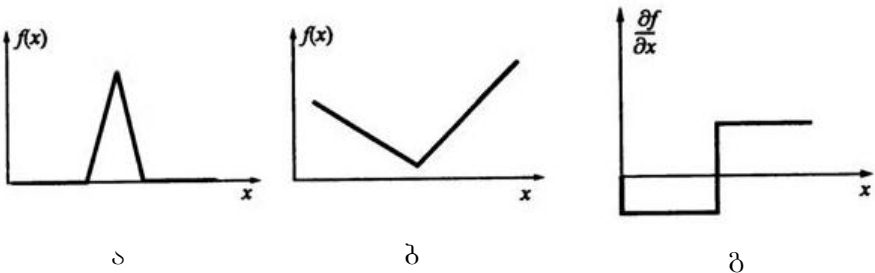
ასეთ შემთხვევაში კონტურის წირის დასავიწროვებლად მიმართავენ ბინარული გამოსახულების დამუშავების სპეციალურ პროცედურებს, რასაც გრაფიკული პრეპარატის “სკელეტიზაციას”- დავიწროებას უწოდებენ. სქელი კონტური და მისი “სკელეტიზაციის” შედეგი ნაჩვენებია 7.10 სურათზე.



სურ. 7.10

- გამოსახულებაზე ხანდახან არსებობს ისეთი არეები, რომელთა საზღვრების წარმომდგენი წირები არ შეესაბამება კონტურის ზემომოყვანილ განსაზღვრებას.

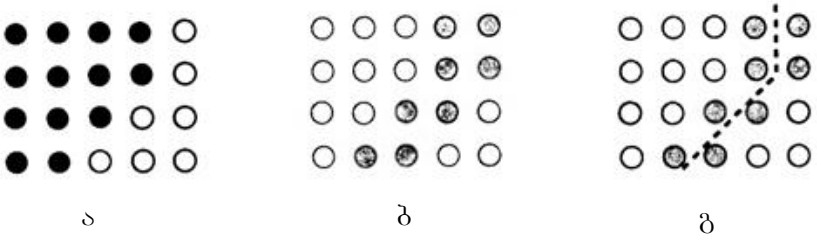
ამას მაშინ აქვს ადგილი, როცა, მაგალითად, ბინარულ გამოსახულებაზე არსებობს ობიექტები, რომელთა საზღვრები ბინარიზაციის შედეგად წარმოდგენილია წვრილი წირების სახით (სურ. 7.11,ა), ან როცა რომელიმე არეში სიკაშკაშის მნიშვნელობების ცვლილებას ტეხილის ფორმა აქვს (სურ. 7.11,ბ). მაგრამ წვრილი წირი თავისთავად წარმოადგენს კონტურს და ამიტომ მის გამოსაყოფად ზღურბლური დამუშავებაც საკმარისია. ხოლო რაც შეეხება ტეხილს, იმისათვის, რომ ის მიეუსადაგოთ კონტურის განსაზღვრებას, საჭიროა სიკაშკაშის f ფუნქციის წინასწარი დიფერენცირება (სურ. 7.11,გ).



სურ. 7.11

- იმის გამო, რომ გამოსახულება ციფრულია, ანუ წარმოადგენს პიქსელების სიმრავლეს $2D$ სივრცეში, ამიტომ არსებობს საზღვრების წარმოდგენის თავისებურება.

არგუმენტის დისკრეტულობის გამო გრაფიკული პრეპარატი – კონტური ბინარიზებულ ციფრულ გამოსახულებაზე მოიცემა ერთეულოვანი და არა ნულოვანი სისქის წირების საშუალებით, როგორც ეს უწყვეტი გამოსახულების შემთხვევაშია, ანუ ის არ არის წარმოდგენილი უსასრულოდ წვრილი წირის სახით (სურ. 7.12,ა,ბ). რა თქმა უნდა, კონტურის წარმოდგენა უფრო მართებული იქნებოდა არა პიქსელების ერთობლიობით შედგენილი წირის, არამედ პიქსელებს შორის გამავალი საზღვრის აღმნიშვნელი წირის საშუალებით (სურ. 7.12,გ წარმოდგენილია პუნქტირით). მაგრამ გარკვეული მიზეზების გამო კონტურის ასეთი სახით წარმოდგენას იშვიათად იყენებენ.



სურ. 7.12

7.3. კედის აღმოჩენის დიფერენციალური მეთოდები

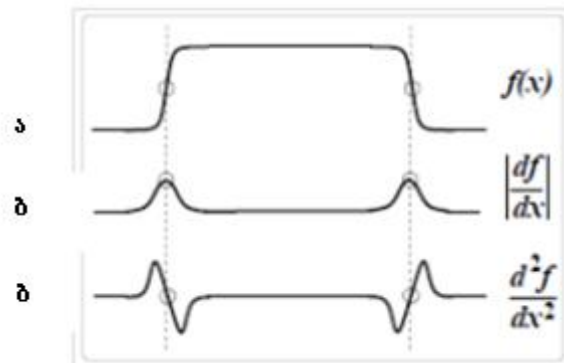
როგორც ზემოთ აღვნიშნეთ, მაღალი კონტრასტის მქონე წერტილების (კედების) გამოვლენა გამოსახულებაზე შესაძლებელია გამოსახულების ლოკალურ არეებში სიკაშკაშის მნიშვნე-

ლობებს შორის სხვაობების გამოთვლით. ასეთი წერტილების სიმრავლეები გამოსახულებაზე განსაზღვრავს სცენის ცალკეულ ობიექტებს შორის საზღვრებს.

კიდეების აღმოჩენის მეთოდები საკმაოდ ბევრია და ერთმანეთისგან განსხვავებული. მიუხედავად ამისა ისინი ძირითადად იყოფიან ორ კატეგორიად:

- **გრადიენტზე დამყარებული:** კიდეების აღმოსაჩენად იყენებენ ინტენსივობის ფუნქციის პირველი რიგის წარმოებულების ექსტრემუმებს. ამ კატეგორიას მიეკუთვნება, მაგალითად, რობერტსის, პრევიტის, სობელის ოპერატორები. მათ მიერ გამოყოფილი ნიშან-თვისებები გამოირჩევა ძალიან მკვეთრი კიდეებით.
- **ლაპლასიანზე დამყარებული:** კიდეების აღმოსაჩენად ინტენსივობის ფუნქციის მეორე რიგის წარმოებულზე ეძებენ ნულზე გადასვლის (მეორე წარმოებულის ნიშნის ცვლილების) წერტილებს. ამ კატეგორიას მიეკუთვნება ოპერატორები, მაგალითად, გაუსიანი, გაუსიანის ლაპლასიანი, გაუსიანების სხვაობა, ქენის დეტექტორი და ა.შ.

გრადიენტული და ლაპლასის ოპერატორების მოქმედების საილუსტრაციოდ 7.13 სურათზე წარმოდგენილია (ა) ერთგანზომილებიანი $f(x)$ სიგნალი, (ბ) მისი პირველი რიგის წარმოებულები და (გ) მეორე რიგის წარმოებულები. როგორც ვხედავთ, (ბ) შემთხვევაში კიდის წერტილები განლაგებულია პირველი წარმოებულის აბსოლუტური მინიმუმების – $\left| \frac{df}{dx} \right|$ მაქსიმუმებში, ხოლო (გ) შემთხვევაში კი მეორე წარმოებულის – $\frac{d^2f}{dx^2}$ ნულზე გადასვლის წერტილებში.



სურ. 7.13

7.4. გრადიენტზე დამყარებული ოპერატორები

გამომდინარე იქიდან, რომ სასაზღვრო წერტილებში (პიქსელებში) აღილი აქვს სიკაშკაშის მნიშვნელობების მკვეთრ ვარდნას, გამოსახულებაზე საზღვრების გამოსაგლენად ყველაზე უფრო მიზანშეწონილია მოცემულ წერტილში ინტენსივობის ფუნქციის გრადიენტის განსაზღვრა. როგორც ცნობილია, გრადიენტული ოპერატორი ∇ წარმოადგენს ვექტორული ტიპის ოპერატორს, რომლის კომპონენტები ორთოგონალური მიმართულებით აღებული პირველი რიგის კერძო წარმოებულებია, ანუ:

$$\nabla = \left[\frac{d}{dx}, \frac{d}{dy} \right].$$

მოცემული უწყვეტი ინტენსივობის $f(x, y)$ ფუნქციისთვის ის შემდეგნაირად ჩაიწერება:

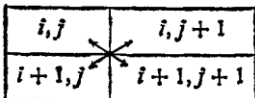
$$\nabla f(x, y) = \left(\frac{df}{dx}, \frac{df}{dy} \right) = (f_x, f_y).$$

ისევე როგორც ყველა ვექტორს, მასაც გააჩნია **სიდიდე (მოდული) $\|\nabla f(x, y)\|$** და **ორიენტაცია $\phi(\nabla f)$** . გამოსახულების მოცემულ წერტილში გრადიენტის სიდიდე გვიჩვენებს განსხვავებას მოცემული და მისი მეზობელი პიქსელის მნიშვნელობებს შორის, ანუ **ლოკალური კონტრასტის** (ვარდნის) სიდიდეს. ხოლო ორიენტაცია კი მოცემული პიქსელის მის მეზობელ პიქსელებთან განსხვავებებს შორის უდიდესი განსხვავების მიმართულებას, რაც, სავარაუდოდ, წარმოადგენს ორიენტაციას, მიმართულს კიდის განივად (კიდის ნორმალის გასწვრივ). უნდა აღინიშნოს, რომ გრადიენტის გამოთვლისას ხშირად ითვლიან მხოლოდ მის სიდიდეს, მაგრამ გასათვალისწინებელია ის, რომ გრადიენტის ორიენტაცია დიდი ინფორმაციის მატარებელია. იმის გამო, რომ ციფრული გამოსახულება არ წარმოადგენს უწყვეტი ტიპის გამოსახულებას, ამიტომ წარმოებულის განსაზღვრა პიქსელთა დისკრეტული სიმრავლის სახით წარმოდგენილ გამოსახულებაზე ხდება მიახლოებით. ორთოგონალური მიმართულებით კერძო წარმოებულების მიახლოებითი გამოთვლისათვის იყენებენ მეზობელ პიქსელებში სიკაშკაშეების მნიშვნელობებს შორის სხვაობებს.

7.4.1. რობერტსის ოპერატორი

გამოსახულების კონტრასტირებისა და ვარდნების აღმოჩენისთვის *რობერტსის* მიერ შემოთავაზებულ იქნა ორგანზომილებიანი დისკრეტული დიფერენცირების მარტივი არაწრფივი ოპერატორი, რომელსაც ასევე რობერტსის ჯვარედინ ოპერატორსაც უწოდებენ. ამ მიდგომის ძირითადი არსი ის არის, რომ მოცემულ პიქსელში გრადიენტის სიდიდის განსაზღვრისთვის, ანუ კიდის (ინტენსივობის ვარდნის) აღმოჩენისთვის ჯერ ხდება ინტენსივობის ფუნქციის წარმოებულების მოძებნა

ორთოგონალური მიმართულებებით, ხოლო შემდეგ კი კიდის საბოლოო აღმოჩენისთვის მათი გაერთიანება. აღვნიშნავთ, რომ მოცემული პიქსელისთვის ორთოგონალურ მიმართულებებზე შეიძლება ავირჩიოთ დიაგონალები ფანჯარაში ზომით (2×2) (სურ. 7.14).



სურ. 7.14

ამგვარად, $I[i, j]$, $i = \overline{1, N}$, $j = \overline{1, M}$, ციფრული გამოსახულების რომელიმე $[i, j]$ პიქსელში გრადიენტის სიდიდე ტოლი იქნება

$$\|\nabla f(i, j)\| = \sqrt{f_x^2 + f_y^2} \approx R[i, j] = \sqrt{(I[i, j] - I[i + 1, j + 1])^2 + (I[i, j + 1] - I[i + 1, j])^2}$$

როგორც ვხედავთ, დისკრეტულ შემთხვევაში გრადიენტის $R[i, j]$ სიდიდის მისაღებად ორთოგონალური მიმართულებით აღებული წარმოებულები აპროქსიმირდებიან მეზობელ პიქსელებში სიკაშკაშეების მნიშვნელობებს შორის სხვაობით. ხშირად გამოთვლების სიჩქარის გაზრდის მიზნით მიმართავენ რობერტსის ოპერატორის გამარტივებას. კერძოდ, კვადრატში აყვანისა და კვადრატული ფესვის ნაცვლად იყენებენ აბსოლუტურ მნიშვნელობებს:

$$F[i, j] = |I[i, j] - I[i + 1, j + 1]| + |I[i, j + 1] - I[i + 1, j]|.$$

ცნობილია, რომ გრადიენტის სიდიდის გამოსათვლელ ზემომოყვანილ ფორმულებს შორის ადგილი აქვს შემდეგ თანაფარდობას:

$$R[i, j] \leq F[i, j].$$

რაც შეეხება გრადიენტის ორიენტაციას, ის ასეთნაირად გამოითვლება

$$\phi(\nabla f) = \arctg(f_x/f_y).$$

აღმოჩნდა, რომ კიდეების პოვნა შეიძლება ე.წ. გამოთვლითი ტრაფარეტების საშუალებითაც. რობერტსის ოპერატორისთვის ტრაფარეტებს ასეთი სახე აქვს (სურ. 7.15)

1	0
0	-1

0	1
-1	0

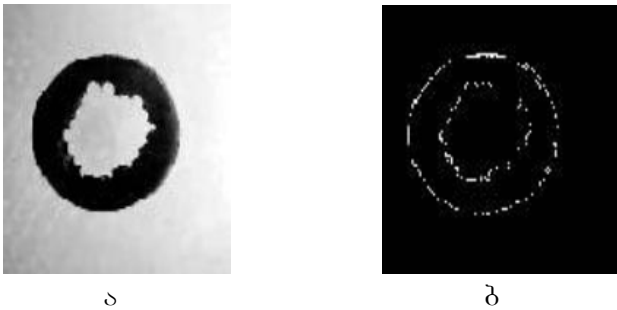
სურ. 7.15

როგორც ვხედავთ, ტრაფარეტი არის ფანჯარა ზომით (2×2) , რომლის უჯრედებში მოთავსებული რიცხვები წარმოადგენს წონებს, რომლებზედაც უნდა გამრავლდეს შესაბამისი სიკაშკაშის მნიშვნელობები. პირველი ტრაფარეტი შეესაბამება გრადიენტის სიდიდის გამოსაანგარიშებელი ფორმულის ფესქვეშა გამოსახულების პირველ წევრს, ხოლო მეორე კი – მეორეს.

თუ რობერტსის ოპერატორით გადავადგილდებით მთელ გამოსახულებაზე ელემენტ-ელემენტ, მივიღებთ გრადიენტულ გამოსახულებას. ხოლო თვით გრადიენტული გამოსახულების მიღების პროცესს სივრცით დიფერენცირებას, კონტურების გა-

მოყოფას, სიმკვეთრის გაზრდას უწოდებენ. უნდა აღინიშნოს, რომ რობერტსის ოპერატორი საუკეთესო შედეგს იძლევა მკვეთრ დაბალხმაურიან გამოსახულებებზე.

რობერტსის ოპერატორის მოქმედების შედეგი ნაჩვენებია 7.16 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) რობერტსის ოპერატორით მიღებული და შემდეგ ბინარიზებული გამოსახულება.



სურ. 7.16

7.4.2. სობელის და პრევიტის ოპერატორები

აღმოჩნდა, რომ გამოსახულებაზე ძლიერი ხმაურის არსებობის შემთხვევაში კიდეების გამოყოფისთვის საჭიროა ინტენსივობის ფუნქციის დაგლუვება, რათა შევამციროთ საშუალოდ ყველაზე ძლიერი ხმაურის გავლენა, მაგრამ ამავე დროს შევინარჩუნოთ ინტენსივობის ვარდნები. გასაშუალოების ამოცანის შესრულებამ მოითხოვა გრადიენტული ოპერატორი, რომელიც იმუშავებდა ფანჯრით (3×3) -ზე. ასეთი არაწრფივი ოპერატორები შემოთავაზებული იქნა სობელის და პრევიტის მიერ.

სობელის მიერ შემოთავაზებული ოპერატორის ფანჯარაში ელემენტების აღნიშვნები მოცემულია 7.17 სურათზე.

A_0	A_1	A_2
A_7	$[i, j]$	A_3
A_6	A_5	A_4

სურ. 7.17

კიდის აღმოსაჩენად გამოსახულების ყოველ $[i, j]$ წერტილში გამოითვლება $G[i, j]$ სიდიდე ფორმულით

$$G[i, j] = \sqrt{X^2 + Y^2}$$

აბ $G[i, j] = |X + Y|,$

სადაც,

$$X = (A_2 + 2A_3 + A_4) - (A_0 + 2A_7 + A_6),$$

$$Y = (A_0 + 2A_1 + A_2) - (A_6 + 2A_5 + A_4).$$

სიდიდეები $A_i, i = \overline{0,7}$ წარმოადგენს ფანჯრის ელემენტების სიკაშკაშეების მნიშვნელობებს, ხოლო 1, 2, და 1 კი წონითი კოეფიციენტებია. შესაბამისად, X და Y სიდიდეებისთვის სობელის ოპერატორის ტრაფარეტები ასე გამოიყურება (სურ. 7.18)

1	2	1	1	0	-1
0	0	0	2	0	-2
-1	-2	-1	1	0	1

სურ. 7.18

ამ ტრაფარეტების ცენტრალური ელემენტები ემთხვევა გამოსახულების $[i, j]$ წერტილს.

რაც შეეხება პრევიტის ოპერატორს, მისი ფანჯარაც ისეთივეა, როგორც სობელის. გამოსახულების ყოველ $[i, j]$ წერტილში პრევიტის ოპერატორის სიდიდე ტოლია

$$G[i, j] = \sqrt{X^2 + Y^2}$$

აბ $G[i, j] = |X + Y|,$

სადაც,

$$X = (A_0 + A_1 + A_2) - (A_6 + A_5 + A_4),$$

$$Y = (A_0 + A_7 + A_6) - (A_2 + A_3 + A_4).$$

სიდიდეები $A_i, i = \overline{0,7}$ წარმოადგენს ფანჯრის ელემენტების სიკაშკაშეების მნიშვნელობებს. აქ წონითი კოეფიციენტები 1-ის ტოლია. X და Y სიდიდეებისთვის კი სობელის ოპერატორის ტრაფარეტებია (სურ. 7.19)

1	-1	1
1	-1	0
1	-1	-1

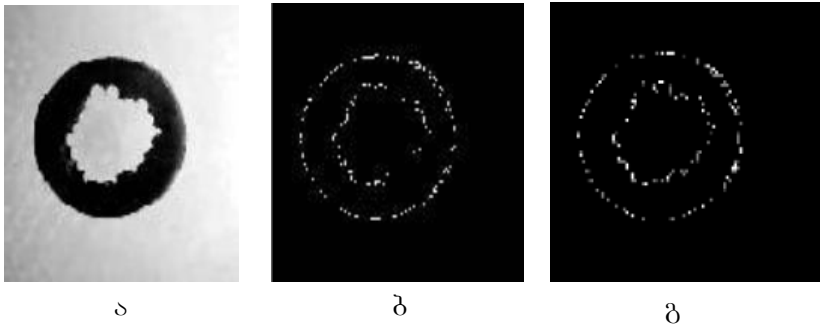
1	1	1
0	0	0
-1	-1	-1

სურ. 7.19

რაც შეეხება გრადიენტის ორიენტაციას, ორივე ოპერატორი გამოირჩევა მის მიმართ დაბალი მგრძნობიარობით.

სობელის და პრევიტის ოპერატორები მსგავსი ოპერატორებია, თუმცა სობელის ოპერატორის მიერ გამოყოფილი კონტური უფრო სქელია, ვიდრე პრევიტის ოპერატორის მიერ. ამ ოპერატორების მოქმედება ნაჩვენებია 7.20 სურათზე, სადაც (ა)

საწყისი გამოსახულებაა, ხოლო (ბ) და (გ) შესაბამისად სობელის და პრევიტის ოპერატორებით მიღებული გამოსახულებებია.



სურ. 7.20

გარდა ზემოაღნიშნული არაწრფივი ოპერატორებისა, არსებობს ასევე მთელი რიგი არაწრფივი ოპერატორები, კერძოდ, უოლისის, როზენფელდის და ა.შ., რომელთა საშუალებით შესაძლებელია კიდევების აღმოჩენა.

გრადიენტული ოპერატორები, ძირითადად, კარგ შედეგებს იძლევა გამოსახულებაზე ერთმანეთისგან იზოლირებული (განცალკევებული) ობიექტების საზღვრების საპოვნელად. მაგრამ მათ ნაკლოვანებებიც გააჩნია. კერძოდ, ობიექტების საზღვრების სიმრუდისა და ჩრდილებით გამოწვეული ინტენსივობის მდოვრე ცვლილების გამო ხანდახან საზღვრები იკარგება. ასევე, მაღალკონტრასტულ არეებზე წარმოიშვება რამდენიმე პიქსელის სისქის მქონე საზღვრები, რომელთა დავიწროება ცალკე ამოცანას წარმოადგენს.

7.5. ლაპლასიანზე დამყარებული ოპერატორები (ლაპლასიანები)

გამოსახულებაზე არსებული კიდევები ყოველთვის არ წარმოადგენს იმ ლოკალურ არეებს, სადაც ადგილი აქვს ინტენსივობის მკვეთს ვარდნას. ხშირად ასეთ არეებში ადგილი აქვს ინტენსივობის მდოვრე – თანდათანობით გადასვლას ერთი მნიშვნელობიდან მეორეში. ასეთ შემთხვევაში გვაქვს აღმაავალი (ზრდადი) გრადიენტის სიდიდე, რომელიც, აღწევს რა პიკს, შემდეგ იწყებს დაშვებას. ამგვარად, იდეალური ვარდნის აღმოჩენა მოითხოვს იმ ოპტიმალური მრუდის პოვნას, რომელიც ოპტიმალური გრადიენტის სიდიდეს მოგვცემს. მაგრამ არსებობს მეთოდები, რომელთა საშუალებით პიკების პოვნა შესაძლებელია უშუალოდ. როგორც ზემოგანხილული ერთგანზომილებიანი სიგნალის მაგალითიდან ჩანს (სურ. 7.13), იდეალური კიდის მოძებნა ნიშნავს ინტენსივობის ფუნქციის წარმოებულის მაქსიმუმის (აღმაავალი კიდე დადებითი ქანობით) ან მინიმუმის (დაღმაავალი კიდე უარყოფითი ქანობით) პოვნას. ამ ექსტრემუმების საპოვნელად ჩვენ ფუნქციის პირველი წარმოებულის გაუტოლეთ ნულს. თუ ახლა მოვახდენთ პირველი წარმოებულის ისევ გაწარმოებას, მივიღებთ ინტენსივობის ფუნქციის მეორე წარმოებულს. აღმოჩნდა, რომ ოპტიმალური კიდეების (გრადიენტის სიდიდის ექსტრემუმების) პოვნა ეკვივალენტურია მეორე რიგის წარმოებულზე იმ ადგილების პოვნისა, სადაც ის ნულის ტოლი ხდება. ეს კარგად ჩანს 7.13 სურათზე. გამოსახულებაზე დიფერენციალური ოპერატორების გამოყენებისას მეორე რიგის წარმოებულის ნულზე გადასვლის წერტილები იშვიათად ემთხვევა ხოლმე ზუსტად პიქსელებს. ჩვეულებრივად ისინი პიქსელებს შორის ხვდება. ჩვენ შეგვიძლია ეს “ნულები” ერთმანეთისგან განვაცალკევოთ ნულზე გადასვლის წერტილზე

ბის მოძებნით ანუ იმ ადგილების პოვნით, სადაც ერთი პიქსელი დადებითია, ხოლო მისი მეზობელი კი – უარყოფითი, ან პირიქით.

ნულზე გადასვლის წერტილებს აქვს ერთი კარგი თვისება – ისინი უზრუნველყოფს შეკრული კონტურების მიღებას გარდა იმ კერძო შემთხვევისა, როცა კონტური გადის გამოსახულების საზღვრებს გარეთ.

ჩვეულებრივად, ნულზე გადასვლის (ნიშნის ცვლილების) წერტილების პოვნაზე დამყარებულ მეთოდებთან დაკავშირებულია ორი პრობლემა:

- ისინი გვაძლევს ორი პიქსელის სისქის კიდევს (კონტურებს) (დადებითი პიქსელი ერთ მხარეს და უარყოფითი პიქსელი მეორე მხარეს) (სურ. 7.21);
- ისინი შეიძლება ძალიან მგრძობიარე იყოს ხმაურის მიმართ.



**ნულზე გადასვლის
წერტილი**

სურ. 7.21

7.5.1. ლაპლასის ოპერატორი

ლაპლასის ოპერატორი (ლაპლასიანი) წარმოადგენს სკალარს, რომლის კომპონენტები ფუნქციის მეორე რიგის წარმოებულებია. უწყვეტი ინტენსივობის $f(x, y)$ ფუნქციისთვის მას ასეთი სახე აქვს

$$\Delta = \nabla^2 f(x, y) = \frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2} .$$

ციფრული გამოსახულებისთვის გამოყენებული ლაპლასიანი წარმოდგენს უწყვეტი ფუნქციის ლაპლასიანის აპროქსიმაციას, ისევე როგორც ეს გრადიენტის შემთხვევაში გვქონდა. ის წრფივი ოპერატორია. ლაპლასიანის საინტერესო თვისებას წარმოადგენს მისი ინვარიანტულობა ბრუნვის მიმართ. კერძოდ, ის არ არის დამოკიდებული წარმოებულების აღების მიმართულების არჩევაზე, მთავარია ისინი ორთოგონალური იყოს. ამის გამო, მეორე რიგის წარმოებულების ჯამი ნებისმიერი ორი ორთოგონალური მიმართულებით ერთმანეთის ტოლია. ამგვარად, ლაპლასიანი არ არის მგრძნობიარე კიდის ორიენტაციის მიმართ, ანუ არ იძლევა სასარგებლო ინფორმაციას კიდის ორიენტაციის შესახებ. ასევე იმის გამო, რომ ჩვენ საქმე გვაქვს მეორე რიგის წარმოებულთან, ის ორმაგად აძლიერებს ხმაურს. მაგრამ ეს უკანასკნელი მარტო ამით არ არის განპირობებული, არამედ ცნობილია, რომ რაც უფრო მაღალია წარმოებულის რიგი, მით უფრო მგრძნობიარეა ოპერატორი.

ვთქვათ, გამოსახულება მოცემულია 7.17 სურათზე წარმოდგენილი ფანჯრის საშუალებით და მოვახდინოთ მისი ცენტრალური ელემენტის ლაპლასიანის შეფასება 4-ბმულიანი მეზობლობის პირობებში. ამისათვის გამოვიყენოთ შემდეგი მიხსლოებები:

$$\begin{aligned} \frac{d^2 f}{dx^2} &\approx \frac{1}{\epsilon^2} (A_7 - 2[i, j] + A_3), \\ \frac{d^2 f}{dy^2} &\approx \frac{1}{\epsilon^2} (A_5 - 2[i, j] + A_1) \end{aligned}$$

სადაც ϵ მანძილია ორი მეზობელი ელემენტის ცენტრებს შორის, ხოლო 1, -2 და 1 წონითი კოეფიციენტებია. ამ ორი მდგენელის გაერთიანებით მივიღებთ ლაპლასიანის შეფასებას ცენტრალური $[i, j]$ პიქსელისთვის.

$$\frac{d^2f}{dx^2} + \frac{d^2f}{dy^2} \approx \frac{4}{\epsilon^2} \left[\frac{1}{4} (A_7 + A_5 + A_3 + A_1) - [i, j] \right].$$

ამგვარად, 4-ბმულიანი მეზობლობისას ფანჯრის ცენტრალურ ელემენტში ლაპლასიანის მიახლოებითი მნიშვნელობის გამოსათვლელად საჭიროა მისი 4 მეზობლის სიკაშკაშეების საშუალო მნიშვნელობას გამოვაკლოთ ინტენსივობის მნიშვნელობა ცენტრალურ ელემენტში. ცხადია, ერთგვაროვან არეში ლაპლასის ოპერატორის მნიშვნელობა 0-ის ტოლია. მაგრამ ამას ადგილი აქვს ისეთ არეებშიც, სადაც სიკაშკაშე წრფივად იცვლება.

7.22 სურათზე წარმოდგენილია გამოსახულების მოცემულ წერტილში ლაპლასიანის გამოსათვლელი ტრაფარეტები 4-ბმულიანი და 8-ბმულიანი მეზობლობის შემთხვევებისთვის

$\frac{1}{4}$	0	-1	0
	-1	4	-1
	0	-1	0

$\frac{1}{8}$	-1	-1	-1
	-1	8	-1
	-1	-1	-1

სურ. 7.22

ლაპლასის ოპერატორის გამოსახულებაზე მოქმედების შედეგი ნაჩვენებია 7.23 სურათზე, სადაც (ა) საწყისი გამოსახულება, (ბ) ლაპლასის ოპერატორის მოქმედების შემდეგ მიღებული გამოსახულება.



ა

ბ

სურ. 7.23

7.6. კიდეების აღმოჩენა გაუსის ფილტრაციით და ლოგ - ფილტრაციით

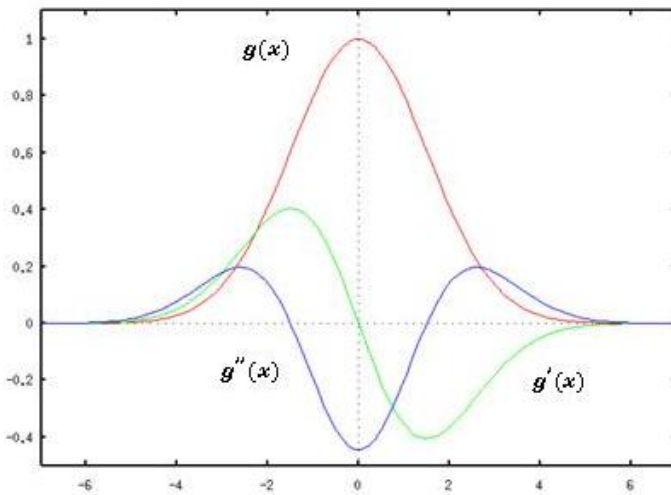
გაუსის ფუნქცია (გაუსიანი) ფართოდ გამოიყენება როგორც მათემატიკაში, ასევე გამოყენებით ამოცანებში. გამოსახულებების ციფრული დამუშავების შემთხვევაში კი – გამოსახულებათა ფილტრაციისთვის. კერძოდ, დაგლუვებისა და კიდეების აღმოჩენის ამოცანებში. განვიხილოთ გაუსის ფუნქციის ის თვისებები, რომლებიც მას სასარგებლოს ხდის აღნიშნული ამოცანების გადაწყვეტის საქმეში.

გაუსის ფუნქციას ორი ცვლადის შემთხვევაში ასეთი სახე აქვს

$$g(x, y) = ce^{-\frac{(x^2+y^2)}{2\sigma^2}},$$

სადაც c მუდმივაა, σ - საშუალოკვადრატული გადახრა. ერთი ცვლადის გაუსის ფუნქცია და მისი პირველი და მეორე რიგის წარმოებულები ნახვენებია 7.24 სურათზე.

აღნიშნული ფუნქციები გაუსის ნორმალური განაწილების სიმკვრივის ფუნქციის მსგავსია. ამასთან, ნორმალური განაწილების დროს c მუდმივა ისე შეირჩევა, რომ განაწილების მრუდის ქვეშ ფართობი 1-ის ტოლი იყოს. ხოლო გამოსახულების ფილტრაციისთვის ტრაფარეტის ფორმირებისას კი c -ს ისეთი მნიშვნელობა შეირჩევა, რომ ტრაფარეტის ყველა ელემენტის მნიშვნელობა მთელი რიცხვი იყოს. როგორც სურათიდან ჩანს, გაუსიანი მოცემულია კოორდინატთა სათავის მიმართ სიმეტრიული ფუნქციის სახით. ამიტომ მისთვის, როგორც ნორმალური განაწილებისთვის, საჭირო არ არის μ მათემატიკური ლოდინის მნიშვნელობის მოცემა.



სურ. 7.24

როგორც ვხედავთ, კოორდინატთა სათავიდან დაშორების ზრდასთან ერთად ფუნქციის მნიშვნელობა 0-კენ მიისწრაფვის. შესაბამისად, ფილტრაციისას გამოსახულების იმ პიქსელების ინტენსივობის მნიშვნელობები, რომლების ცენტრთან ახლოსაა,

უფრო მნიშვნელოვანი იქნება, ვიდრე მისგან უფრო დაშორებული პიქსელების. ამასთან, დასამუშავებელი პიქსელების მიდამოს სიგანე დამოკიდებულია σ -ს სიდიდეზე. ფილტრის ტრაფარეტის წონითი კოეფიციენტების 95% მოდის გაუსიანის ცენტრიდან 2σ მანძილის საზღვრებში. როგორც სურათიდან ჩანს, მეორე წარმოებულს აქვს ამოხნეკილობა უარყოფით ნახევარსიბრტყეში და ორი გვერდითი ამოხნეკილობა დადებით ნახევარსიბრტყეში. ფუნქცია ნულის ტოლი ხდება $-\sigma$ და σ წერტილებში, რომლებიც შეესაბამება $g(x)$ ფუნქციის გადაღუნვის და $g'(x)$ ფუნქციის ექსტრემუმის წერტილებს.

როგორც ზემოთ აღვნიშნეთ, ლაპლასიანი მგრძობიარეა ხმაურის მიმართ, ამის გამო მას ერთნაირი წარმატებით შეუძლია როგორც კიდებობის, ასევე ხმაურის აღმოჩენაც. ამ პრობლემის გადასაჭრელად ჯერ მიმართავენ გამოსახულების დაგლუვებას σ სიგანის გაუსის დაგლუვების ოპერატორით (გაუსის ბირთვით), ხოლო შემდეგ კი კიდებობის აღმოსაჩენად იყენებენ ლაპლასის ოპერატორს. შედეგად მიიღება მეორე რიგის დიფერენციალური ფილტრი (ოპერატორი),

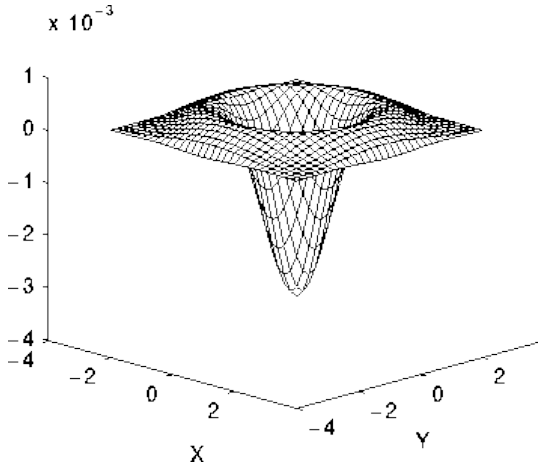
$$\begin{aligned} \text{Log} = \Delta g(x, y) &= \frac{d^2}{dx^2} g(x, y) + \frac{d^2}{dy^2} g(x, y) = \\ &= c \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-(x^2 + y^2)/2\sigma^2}, \end{aligned}$$

რომელსაც გაუსიანის ლაპლასიანი (Laplacian of the Gaussian, LoG) ანუ ლოგ-ფილტრი ეწოდება.

ორი ცვლადის ლოგ-ფილტრს, ცენტრირებულს კოორდინატთა სათავის მიმართ ($\mu = 0$) და გაუსიანის სტანდარტული გადახრით ($\sigma = 1$) აქვს ასეთი სახე

$$\text{Log}(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{\frac{x^2+y^2}{2\sigma^2}},$$

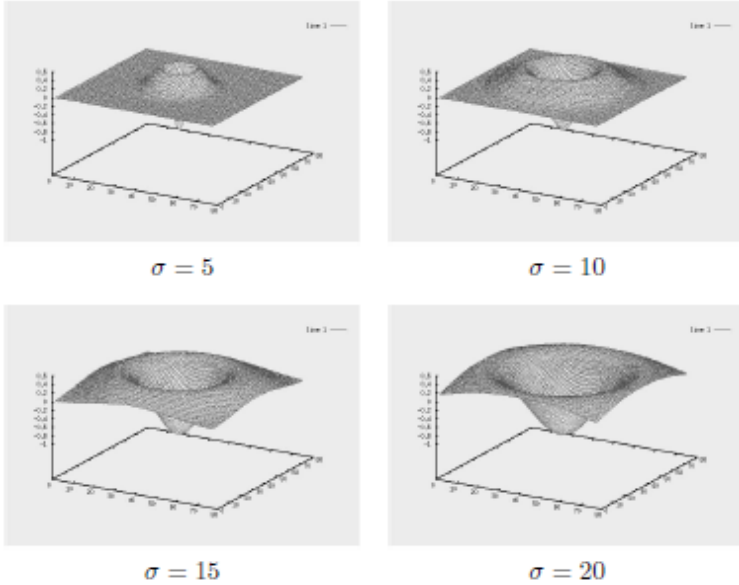
ხოლო მისი გრაფიკი 3D სივრცეში წარმოდგენილია 7.25 სურათზე.



სურ . 7.25

ლოგ-ფილტრის გრაფიკები σ სხვადასხვა მნიშვნელობებისთვის მოცემულია 7.26 სურათზე.

უწყვეტი $\text{Log}(x, y)$ ლოგ-ფილტრის დისკრეტული მიახლოებითი მნიშვნელობის მისაღებად შეიძლება გამოყენებულ იქნეს ტრაფარეტი (სურ. 7.27).



სურ. 7.26

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & -16 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

სურ. 7.27

ამასთან უნდა აღინიშნოს, რომ ტრაფარეტის ზომები შეიძლება ნებისმიერი იყოს, მაგრამ მიუხედავად ამისა, მისი ელემენტების ჯამი აუცილებლად ნულის ტოლი.

ნაბიჯები, რომელთა შესრულებაც საჭიროა კიდევების აღმოსაჩენად ლოგ-ფილტრის გამოყენებით შემდეგია:

- ლოგ-ფილტრის გამოყენება გამოსახულებაზე;

- ნულზე გადასვლის წერტილების აღმოჩენა გამოსახულებაზე;
- ზღურბლის გამოყენება ნულზე გადასვლის წერტილებზე ძლიერი წერტილების (რომელთა შორის განსხვავება დიდია) შენარჩუნების მიზნით.

ეს ბოლო ნაბიჯი საჭიროა იმისთვის, რომ გამოსახულებაზე დარჩნენ ის წერტილები, რომელთა ინტენსივობებს შორის დიდი განსხვავებაა, ანუ სადაც გვაქვს კიდევები, და დაითრგუნოს ნულზე გადასვლის სუსტი წერტილები, რადგანაც ასეთი წერტილები ყველაზე უფრო ხშირად ხმაურს შეესაბამებიან.

7.28 სურათზე წარმოდგენილია ზემოთაღნიშნული ნაბიჯების შესაბამისად განხორციელებული კიდევების აღმოჩენის შედეგი, სადაც (ა) საწყისი გამოსახულებაა, ხოლო (ბ) და (გ) შესაბამისად (15 × 15) და (21 × 21) ზომის ტრაფარეტების გამოყენებით მიღებული გამოსახულებები.



სურ. 7.28

ლოგ-ფილტრის გამოყენება კარგ შედეგს იძლევა ინტენსივობის მნიშვნელობათა სხვაობის ორი ტიპისთვის. კერძოდ, მცირე ზომის ლაქებისთვის, რომელთა ცენტრი ემთხვევა ფილ-

ტრის ტრაფარეტის ცენტრს და ფართო, საფეხუროვანი, ტრაფარეტის ცენტრთან ახლოს მყოფი კიდეებისთვის.

ლოგ-ფილტრის აპროქსიმაცია შესაძლებელია ფილტრით, რომელიც წარმოადგენს ორი განსხვავებული გადახრის მქონე გაუსიანის სხვაობას

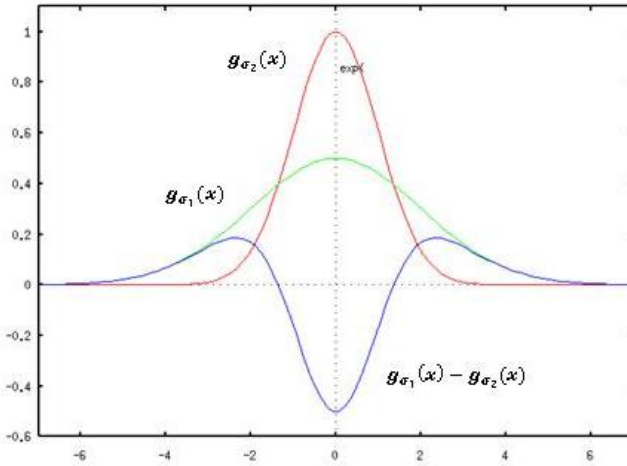
$$g''(x, y) \approx c_1 e^{-\frac{x^2+y^2}{2\sigma_1^2}} - c_2 e^{-\frac{x^2+y^2}{2\sigma_2^2}}.$$

გამოსახულებისთვის ეს ნიშნავს, რომ მისი დაგლუვება ჯერ ხდება σ_1 სიგანის მქონე გაუსიანით, შემდეგ $-\sigma_2$ სიგანის გაუსიანით, ხოლო შემდეგ აიღება მათი სხვაობა. შედეგად მიღებულ ფილტრს (ოპერატორს) **გაუსიანების სხვაობა (Difference of Gaussian, Dog)** ანუ **ლოგ-ფილტრი** ეწოდება. ფორმალურად ლოგ-ფილტრი ასე წარმოდგინება

$$DoG = g_{\sigma_1} - g_{\sigma_2} = c \left[\frac{1}{\sigma_1} e^{-(x^2+y^2)/2\sigma_1^2} - \frac{1}{\sigma_2} e^{-(x^2+y^2)/2\sigma_2^2} \right].$$

ერთი ცვლადის შემთხვევაში ფუნქციები $g_{\sigma_1}(x)$, $g_{\sigma_2}(x)$ და მათი სხვაობა გრაფიკულად წარმოდგენილია 7.29 სურათზე.

იმისათვის, რომ გაუსიანების სხვაობის შესაბამისი მრუდის ცენტრალური ამოზნექილობა დადებით ნახევარსიბრტყეში იყოს, საჭიროა, რომ $\sigma_1 < \sigma_2$. σ_2 მნიშვნელობა უნდა შეირჩეს σ_1 მნიშვნელობის გათვალისწინებით ისეთნაირად, რომ ფილტრის გამოძახილის მნიშვნელობებს შორის არსებული ნულზე გადასვლის წერტილები კორექტულად იყოს განლაგებული და ამასთან ფილტრის უარყოფითი კოეფიციენტების ჯამი გაწონასწორებული იყოს დადებითი კოეფიციენტების ჯამით.



სურ. 7.29

ლოგ-ფილტრის დისკრეტული მიახლოება მიიღება ზემო-წარმოდგენილი უწყვეტი ლოგ-ფილტრის გამოსახულების აპროქსიმაციით. ამასთან, ისევე როგორც ლოგ-ფილტრის შემთხვევაში, აქაც აუცილებელია, რომ ტრაფარეტის ყველა ელემენტის ჯამი ნულის ტოლი იყოს.

ლოგ-ფილტრის გამოყენება გამოსახულებაზე კიდეების აღმოსაჩენად მსგავსად ლოგ-ფილტრისა იგივე ნაბიჯების შესრულებას საჭიროებს.

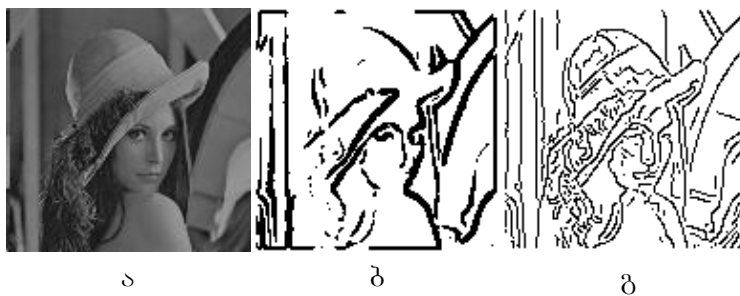
7.6.1. ქენის დეტექტორი

ქენის დეტექტორი (Canny detector) წარმოადგენს კიდეების აღმომჩენ ეფექტურ წრფივ ოპერატორს. მის პირველ და ყველაზე უფრო ცხად თვისებას წარმოადგენს დაბალი ცდომილების კოეფიციენტი. ეს ძალიან მნიშვნელოვანია, რადგანაც ოპერატორმა არ უნდა გამოტოვოს კიდის წერტილი იქ, სადაც ის არის და უნდა გვიოხრას არა, მაშინ როცა ის არ არის. ქენის

ოპერატორის მოქმედების თანამიმდევრობა მრავალსაფეხურიანია. ჯერ ოპერატორი ხმაურის მოცილების მიზნით ახდენს ნახევარტონური გამოსახულების დაგლუვებას σ გადახრის მქონე გაუსის ფილტრით. ამის შემდეგ დაგლუვებულ გამოსახულებაზე გამოიყენება რომელიმე გრადიენტული ოპერატორი მაღალი სიდიდის პირველი რიგის სივრცითი წარმოებულების მქონე არეების (კონტურების) გამოსაყოფად. მიღებულ კონტურებზე ყველაზე მაღალი გრადიენტის სიდიდის მქონე თხემური წერტილების (წვეროების) გამოსაყოფად ალგორითმი მიყვება კონტურს და ნულის ტოლ სიკაშკაშის მნიშვნელობას ანიჭებს ყველა იმ პიქსელს, რომლებშიც გრადიენტის სიდიდე არ აღემატება მოცემული პიქსელიდან გრადიენტის მიმართულების გასწვრივ ორივე მხარეს მყოფ ორ მეზობელ (თითო პიქსელი თითო მხარეს) პიქსელში გრადიენტის სიდიდეს. ასეთ პროცედურას შედეგად მოყვება კონტურის დავიწროვება. აღნიშნულ მეთოდს **არამაქსიმალური ჩახშობა** (nonmaximum suppression) ეწოდება. აღნიშნული პროცედურა იმართება ორი, T_1 (მაღალი) და T_2 (დაბალი) ზღურბლის საშუალებით ($T_1 > T_2$). გრადიენტულ გამოსახულებაზე კონტურზე გადაადგილების პროცესი აუცილებლად უნდა დაიწყოს თხემის იმ წერტილიდან, რომელშიც გრადიენტის სიდიდე მეტია T_1 ზღურბლზე. ამის შემდეგ კონტურის მიყოლა გრძელდება აღნიშნული წერტილის გრადიენტის მიმართულების გასწვრივ ორივე მხარეს მანამ, სანამ არ მივაღწევთ თხემის იმ წერტილამდე, რომელშიც გრადიენტის სიდიდე ნაკლები აღმოჩნდება T_2 ზღურბლზე. ამგვარად, გრადიენტული გამოსახულების ნებისმიერი წერტილი, რომელშიც გრადიენტის სიდიდე მეტია T_1 -ზე, ითვლება კიდის ანუ კონტურის წერტილად. კონტურის წერტილად ითვლება აგრეთვე მოცემულ წერტილთან ბმული ყველა ის წერტილი, რომელშიც გრადიენტის სიდიდე მეტია T_2 -ზე. ჩვეულებრივ,

ზღურბლის დაბალ T_2 მნიშვნელობად იღებენ მაღალი T_1 მნიშვნელობის დაახლოებით ნახევარს.

ქენის ოპერატორის მოქმედების შედეგი ნაჩვენებია 7.30 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) გაუსის ფილტრის გამოყენებით მიღებული გამოსახულება, (გ) კონტურის დაევიწროებით მიღებული გამოსახულება.

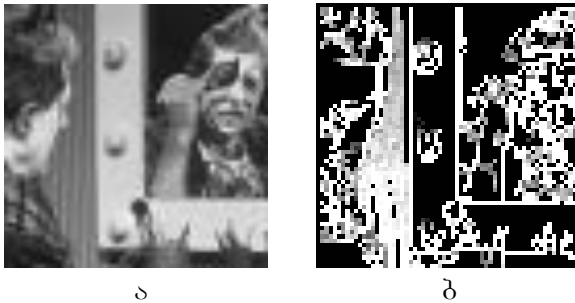


სურ. 7.30

ქენის დეტექტორის ეფექტურობა განპირობებულია სამი პარამეტრის – გამოსახულების დაგლუვების ეტაპზე გამოყენებული გაუსიანის σ სიგანით და კონტურის მიყოლის ეტაპზე გამოყენებული T_1 და T_2 ზღურბლების სიდიდით. გაუსიანის სიგანის გაზრდა ამცირებს დეტექტორის მგრძობიარობას ხმაურის მიმართ. გაუსიანის სიგანის გაზრდით ასევე იზრდება გამოყოფილი კიდის წერტილების ლოკალიზაციის ცდომილება. რაც შეეხება ზღურბლების სიდიდეს, კარგი შედეგის მისაღებად მაღალი ზღურბლი უნდა იყოს საკმაოდ მაღალი, ხოლო დაბალი ზღურბლი კი საკმაოდ დაბალი. დაბალი ზღურბლისთვის დიდი მნიშვნელობის მინიჭება გამოიწვევს ხმაურის შემცველი კიდების დანაწევრებას, ხოლო მაღალი ზღურბლის საკმაოდ დაბალი მნიშვნელობა კი გაზრდის ცრუ და არასასურველი კიდის ფრაგმენტების რაოდენობას.

ქენის ოპერატორის გამოყენებას თან სდევს ერთი პრობლემა. კერძოდ, ეს ეხება შემთხვევას, როცა გრადიენტულ გამოსახულებაზე ერთ წერტილში ერთმანეთს ხედება კიდის სამი სწორხაზოვანი სეგმენტი (მონაკვეთი), ანუ გვაქვს Y-ტიპის შეერთება. ასეთ შემთხვევას ადგილი აქვს მაშინ, როცა კიდე ნაწილობრივ დაფარულია სხვა ობიექტით. აღნიშნულ შემთხვევაში კონტურის მიმყოლი ალგორითმი 3 მონაკვეთიდან ორს განიხილავს როგორც კიდის სწორხაზოვან სეგმენტს, ხოლო მესამეს როგორც წირს, რომელიც უახლოვდება, მაგრამ მთლიანად არ უერთდება აღნიშნულ სეგმენტს.

7.31 სურათზე მოცემულია ქენის ოპერატორის გამოყენების მაგალითი გამოსახულებაზე, რომელზედაც სურათის ქვედა მარცხენა კუთხეში გვაქვს Y-ტიპის შეერთების შემთხვევა; (ა) საწყისი გამოსახულებაა, (ბ) ქენის ოპერატორის მოქმედების შედეგად მიღებული გამოსახულება.



სურ. 7.31

VIII თავი. სცენის გამოსახულებათა ანალიზი

8.1. შესავალი

სცენის გამოსახულების ანალიზი წარმოადგენს გამოსახულებათა ციფრული დამუშავების ყველაზე უფრო მნიშვნელოვან ნაწილს და მოიცავს ამოცანათა ფართო წრეს, რომელთა მიზანია სცენის გამოსახულების ანალიზის საშუალებით მოახდინოს გამოსახულებიდან სასურველი ინფორმაციის (არაგამოსახველობითი ხასიათის) გამოყოფა ავტომატური ან ნახევრადავტომატური (ინტერაქტიული) მეთოდების საშუალებით. ლიტერატურაში მას ასევე სხვა დასახელებებიც აქვს: გამოსახულებიდან მონაცემების გამოყოფა, სცენათა ანალიზი, გამოსახულებათა აღწერა, ავტომატური დეშიფრირება, გამოსახულებათა ამოცნობა და ა.შ.

განსხვავებით გამოსახულებათა დამუშავების სხვა სახეებისგან, სცენის გამოსახულების ანალიზის შედეგს წარმოადგენს არა სურათი, არამედ მისი აღწერა, გამოსახული რიცხვების კრებულის, სიმბოლოთა ჯაჭვის ან გრაფის სახით. გარდა ამისა განსხვავება გამოსახულებათა ანალიზსა და კლასიკურ სახეთა ამოცნობას შორის მდგომარეობს იმაშიც, რომ ანალიზისას ხდება არ მარტო სცენაზე არსებული არეების კლასებად გამოყოფა, არამედ მისი მთავარი მიზანი რთული სცენების აღწერაა, რომელთა მრავალფეროვნება შეიძლება იმდენად დიდი იყოს, რომ არ აღიწეროს წინასწარ მოცემულ ტერმინებში. გარდა ამისა გამოსახულებათა ანალიზის სისტემებში ხშირად გამოიყენება აპრიორული მონაცემები როგორც გამოსახულებაზე არსებული ობიექტების, ასევე მათი ურთიერთმიმართებების შესახებ.

კომპიუტერული ხედვის ამოცანების შესწავლისას ინტერესს წარმოადგენენ ამოცანები, რომლებიც დაკავშირებულია ობიექტების ამოცნობასთან ნიშან-თვისებების გამოყენებით, ფონიდან ობიექტის გამოყოფის ხერხებთან, რაც კომპიუტერულ ხედვაში ცნობილია როგორც სეგმენტაციის პროცესი და მათ ინტერპრეტაციასთან. მხედველობის სისტემაში ანალოგიური პროცესების შესწავლამ წარმოაჩინა გამოსახულების წინასწარი დამუშავების მიზნები და მათი კავშირი რეალური გარესამყაროს შინაგანი წარმოდგენის ცნებასთან. ეს პროცესები წინ უსწრებს სცენის ცალკეულ ობიექტებად დანაწევრებას და მათი ამოცნობის პროცედურებს.

ზემოთქმულიდან გამომდინარე, სცენის გამოსახულების ანალიზის პროცესი მოიცავს ორ ეტაპს:

- **გამოსახულების წინასწარი დამუშავება**, რაც გულისხმობს გამოსახულების ისეთ წინასწარ დამუშავებას, რომელიც გააადვილებს მისი ანალიზისათვის საჭირო საუკეთესო (ინფორმაციული) ნიშან-თვისებების გამოყოფას; მაგალითად, საწყისი გამოსახულების ხარისხის გაუმჯობესება, ხმაურისგან გაწმენდა, სიკაშკაშის შკალის გარდაქმნა და ა.შ.
- **სეგმენტაცია**, რომლის არსი დაკავშირებულია მხედველობის ერთ-ერთ ცენტრალურ საკითხთან, ფონიდან ობიექტის გამოყოფის პრობლემასთან და რომლის მიზანია გამოსახულების (როგორც სახის) დანაწევრება ცალკეულ ობიექტებად (ქვესიმრავლეებად), ანუ ერთნაირი თვისებების მქონე არეებად, რომელთა არჩევა განისაზღვრება ან რაიმე კონკრეტული მიზნიდან გამომდინარე, ან ამ არეების შესაბამისობით რეალურ ობიექტებთან ან მათ ნაწილებთან.

სცენათა ანალიზის უმარტივესი გამოყენების სფეროს წარმოადგენს გამოსახულებაზე არსებული ობიექტების კლასიფიკაცია გამოსახულების რაოდენობრივი მახასიათებლების (ნიშან-თვისებების) საფუძველზე. რაც გულისხმობს იმას, რომ ობიექტის თვისებების განსაზღვრის შემდეგ უკვე შესაძლებელია მოვახდინოთ მისი მიკუთვნება რომელიმე წინასწარ ცნობილი ობიექტების კლასისთვის. **კლასის** ქვეშ იგულისხმება რაღაც კატეგორია, განსაზღვრული მასში შემავალი ელემენტების თვისებათა ერთობლიობით. ამ ამოცანის გადაწყვეტაში ძირითადი როლი ენიჭება იმ მახასიათებლების ანუ ნიშან-თვისებების გამოყოფას, რომელთა ერთობლიობით აღიწერება გამოსახულების ობიექტები. ამიტომ ნიშან-თვისებების შერჩევისა და მათი მნიშვნელობების განსაზღვრის ეტაპი ერთ-ერთი მნიშვნელოვანია კლასიფიკაციის ამოცანაში.

8.2. სცენის გამოსახულებათა ამოცნობა: ზოგადი მოდელი

კომპიუტერული ხედვის გამოყენების სხვადასხვა შემთხვევაში სცენა ორგანზომილებიანია. ასეთი სცენის მაგალითებს წარმოადგენს: დოკუმენტები, დიდი სიმაღლიდან გადაღებული დედამიწის ზედაპირის ფოტოსურათები (როცა რელიეფი შეიძლება უგულებელყოფილი იქნეს), ნიშუშები, რომლებზედაც დაკვირვება ხორციელდება მიკროსკოპის საშუალებით (როცა ხედვის სიღრმე უმნიშვნელოა) და დეტალების დამუშავებული ბრტყელი ზედაპირები. არსებობს მთელი რიგი ამოცანებისა, სადაც გამოსახულებები განიხილება როგორც იმ ინფორმაციის წყარო, რომლის საფუძველზეც საჭირო ხდება გარკვეული გადაწყვეტილების მიღება. მაგალითად, სამედიცინო დიაგნოსტიკაში, სადაც ადამიანის რომელიმე ორგანოს შემცველი სამედიცინო გამოსახულება ანალიზდება ავადმყოფ-

ფობის აღმოჩენის მიზნით ან დაქტილოგრამები (აღამიანის თითის ანაბეჭდების გამოსახულებები), რომლებიც გამოიყენება პიროვნების დადგენის მიზნით და ა.შ. ასეთი ამოცანების გადასაწყვეტად გამოიყენება სახეთა ამოცნობის თეორია. ამ თეორიის მიხედვით ზემოაღწერილი ამოცანების გადაწყვეტისას სახეს წარმოადგენს სწორედ გამოსახულება. ასეთ შემთხვევაში სახეთა ამოცნობის ამოცანაა გამოსახულებათა კლასიფიკაცია გარკვეული პირობების საფუძველზე, რომლის თანახმადაც ის გამოსახულებები, რომლებიც სახეთა ერთ კლასს მიეკუთვნება, ხასიათდება სიახლოვის მაღალი ხარისხით. უფრო კონკრეტულად ეს ნიშნავს საკვლევ გამოსახულებაზე გამოთვლილი ნიშან-თვისებების კლასიფიკაციას. ამავე დროს არსებობს ამოცანები, სადაც საჭირო ხდება თვით გამოსახულებაზე არსებული ობიექტების პოვნა და ამოცნობა. აქედან გამომდინარე, ყველა ამოცანა, რომელიც წარმოიშობა გამოსახულებაზე სახეთა ამოცნობის მეთოდების გამოყენებისას, იყოფა ორ ჯგუფად:

- გამოსახულებათა კლასიფიკაცია ანუ ამოცნობა;
- გამოსახულებაზე ობიექტების (ლოკალური არეების) პოვნა და ამოცნობა.

ამოცანათა პირველ ჯგუფში ამოცნობის ამოცანა წყდება მთელი გამოსახულებისთვის, ანუ ამოცნობის პროცესში ხდება გამოსახულების მიკუთვნება რამდენიმე წინასწარ ცნობილი კლასიდან ერთ-ერთისადმი. ეს ნიშნავს, რომ ამოცნობის ამოცანის ამოხსნა წარმოადგენს ასეთი სახის ასახვის რეალიზაციას:

გამოსახულება ⇒ კლასის ნომერი.

მაგრამ, თავის მხრივ, სახეთა კლასიფიკაცია ნიშნავს ნიშან-თვისებათა სივრცის ასახვას გადაწყვეტილებათა სივრცეში. აქედან გამომდინარე, ამოცნობის პროცესი შეიცავს ორ ეტაპს:

- ნიშან-თვისებათა შერჩევა და მოწესრიგება;
- კლასიფიკაცია.

აღნიშნულის გათვალისწინებით, ზემონახსენები ასახვის რეალიზაციისთვის საჭიროა ორი სახის ასახვის რეალიზაცია:

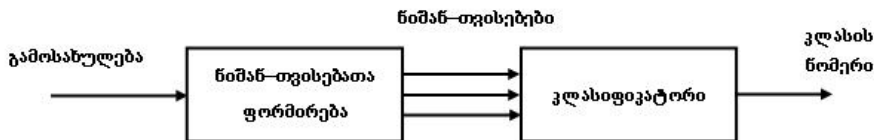
გამოსახულება ⇒ ნიშან-თვისებები

და

ნიშან-თვისებები ⇒ კლასი.

რეალურად ეს ნიშნავს კლასში შემავალი გამოსახულებისთვის კლასის აღმნიშვნელი ჭდის (ნომრის, სიმბოლოს) მინიჭებას. აქედან გამომდინარე, გამოსახულებების კლასიფიკაცია წარმოადგენს გამოსახულებებისთვის ჭდეების მინიჭების პროცესს მათი ნიშან-თვისებების აღწერის საფუძველზე.

გამოსახულებების კლასიფიკაციის ამოცანის გადაწყვეტის პროცესის ზოგადი სქემა ნაჩვენებია 8.1 სურათზე.



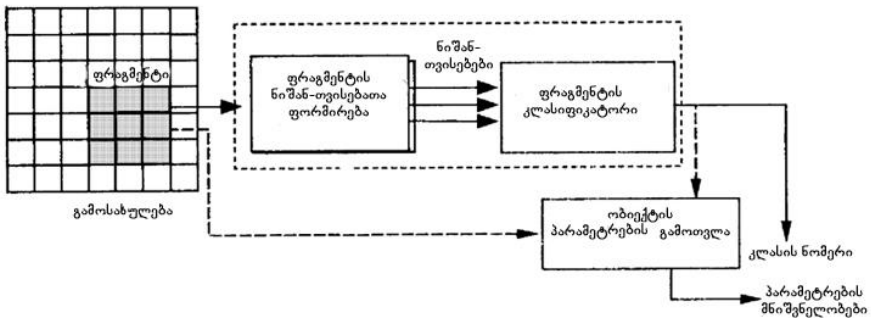
სურ. 8.1

პირველი ჯგუფის მაგალითებს წარმოადგენს ადამიანის პირისახის ამოცნობა ფოტოსურათის საშუალებით, თითის ანაბეჭდების ამოცნობა, ნაბეჭდი და ხელნაწერი სიმბოლოების ამოცნობა, დაავადებათა დიაგნოსტიკა სამედიცინო გამოსახულებების საშუალებით და ა.შ.

ამოცანათა მეორე ჯგუფში ამოცნობის პროცესი წარმოადგენს გამოსახულების დამუშავების უფრო ფართო ამოცანის ნაწილს, კერძოდ, დაკავშირებულს ამოსაცნობი გეომეტრიული ობიექტების ძიებასთან მთელ გამოსახულებაზე. ამ შემთხვევაში ობიექტებს წარმოადგენენ გამოსახულების ნებისმიერ ადგილზე განლაგებული შედარებით მცირე ზომის ლოკალური არეები (ფრაგმენტები). ამ დროს ანალიზდება გამოსახულებაზე არსებული ყველა ფრაგმენტი. თითოეული მათგანისთვის ხდება ნიშან-თვისებათა ფორმირება და შემდგომ კი კლასიფიკაცია. ამასთან, ინფორმაცია იმის შესახებ: არის თუ არა ობიექტები გამოსახულებაზე, რა რაოდენობის, როგორია მათი ორიენტაცია, ზომები და ა.შ., ხშირ შემთხვევაში უცნობია. სწორედ ობიექტის ზოგიერთი მახასიათებლის არცოდნა გამოსახულებაზე ობიექტების მოძიებისა და ამოცნობის ამოცანას მათემატიკურ და გამოთვლით ასპექტში უფრო რთულს ხდის პირველი ჯგუფის ამოცანებთან შედარებით. ასეთი ამოცანის ამონახსნს წარმოადგენს არა მარტო მოძიებული ობიექტისთვის შესაბამისი კლასის განსაზღვრა, არამედ მისი მახასიათებლებისაც: გამოსახულებაზე მდებარეობის, ზომის, ფერის, ორიენტაციის და ა.შ. გამოსახულებაზე ობიექტების მოძიებისა და ამოცნობის სქემა მოცემულია 8.2 სურათზე.

მეორე ჯგუფის მაგალითებს წარმოადგენს: აერო-კოსმოსური სურათების დეშიფრირება, ტექსტების ავტომატური კითხვა, ლოკალური პათოლოგიების მოძიება სამედიცინო გამოსახულებებზე და ა.შ.

როგორც ვხედავთ, გამოსახულებათა კლასიფიკაციის ამოცანის გადაწყვეტის სქემა წარმოადგენს ძიებისა და ამოცნობის ამოცანის გადაწყვეტის სქემის შემადგენელ ელემენტს. მეორე ჯგუფის ამოცანები ფართო გაგებით მიეკუთვნება მაღალი დონის ამოცანათა ჯგუფს. კერძოდ, ამოცანებს, დაკავშირებულს დაკვირვებული გამოსახულების ან სცენის ანალიზის ამოცანასთან. მოცემულ შემთხვევაში სცენის ანალიზის ქვეშიგულისხმება გამოსახულებაზე არსებული ობიექტების სრული აღწერა მათი ადგილმდებარეობის და ურთიერთგანლაგების მითითებით. თუმცა შეიძლება ადგილი ჰქონდეს საწინააღმდეგო შემთხვევასაც, როდესაც მეორე ჯგუფის ამოცანების გადაწყვეტის მეთოდები გამოიყენება როგორც შემადგენელი ელემენტები პირველი ჯგუფის ამოცანების გადაწყვეტისას. ამის მაგალითს წარმოადგენს ადამიანის პირისახის ამოცნობის ამოცანა ფოტოსურათების საფუძველზე, რომელიც მდგომარეობს პირისახეზე მკვეთრად გამოხატული ინტერესის არეების: თვალის, ცხვირის, ტუჩების და ა.შ., მოძიებაში.



სურ. 8.2

ამოცნობის პროცესის პირველ ეტაპზე ნიშან-თვისებათა შერჩევისა და მოწესრიგების კრიტერიუმი განისაზღვრება მათი

მნიშვნელოვნებით სახეთა დახასიათებისთვის. მეორე ეტაპი კი მდგომარეობს გადაწყვეტილების მიღებაში, რომლითაც ხდება სახის მიკუთვნება ამა თუ იმ კლასისადმი, რისთვისაც გამოიყენება მკაცრი მათემატიკური მეთოდები. ამ ამოცანას ასრულებს **კლასიფიკატორი** – მოწყობილობა ან ალგორითმი, რომლის შესასვლელს წარმოადგენს ობიექტის აღწერა, ხოლო გამოსასვლელს კი – კლასის ჭდე.

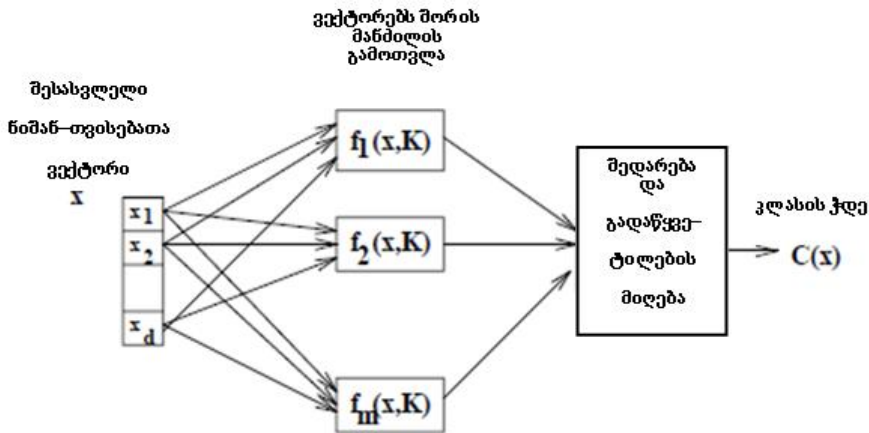
განვიხილოთ კლასიფიკაციის ზოგადი მოდელის შემადგენელი ელემენტები. ამ მოდელში გვაქვს სიმრავლე, რომელიც შედგება ობიექტების წინასწარ ცნობილი m კლასისაგან. კლასები წარმოდგენილია ან რაიმე აღწერის ფორმაში ან ყოველი კლასისთვის ობიექტების მაგალითების სიმრავლის (ეტალონური (training) სახეების) სახით. მაგალითად, სიმბოლოების კლასიფიკაციისას ყოველი სიმბოლოსთვის არსებობს ან მისი გარეგნული აღწერა, ან სახეების კრებული. იმ ობიექტთა სიმრავლეს, რომელთა მიკუთვნებაც რომელიმე წინასწარ ცნობილი კლასისთვის ვერ მოხერხდა, **ამოვარდნების** ანუ **არაკლასიფიცირებულ ობიექტთა ქვესიმრავლე** ეწოდება.

გარესამყაროში არსებული რეალური ფიზიკური ობიექტის აღქმისთვის საჭიროა მისი ისეთნაირი წარმოდგენა ციფრული მონაცემების სახით, რომელიც გამოსადეგი იქნება მანქანური დამუშავებისათვის. ასეთ მოწყობილობას წარმოადგენს **გადამწოდი**. იმის გამო, რომ ჩვენ შევისწავლით კომპიუტერულ ხედვას, შესაბამისად, გადამწოდი უნდა ახდენდეს მონაცემების ორგანიზაციის მასივების გენერირებას.

გადამწოდისაგან მიღებული მონაცემებიდან ობიექტის დახასიათებელ ნიშან-თვისებათა გამომყოფს უწოდებენ **ექსტრაქტორს** (extractor) ან **პრეპროცესორს** (preprocessor). მისი ფუნქციაა ამ მონაცემებში კლასიფიკაციისთვის სასარგებლო ინფორმაცია აღმოაჩინოს. ის მოთავსებულია გადამწოდსა და

კლასიფიკატორს შორის და მის ფუნქციებში ასევე შეიძლება შედიოდეს, მაგალითად, გამოსახულების ფილტრაცია ან სმაურისგან გაწმენდა. ჩვეულებრივად, ნიშან-თვისებათა ექსტრაქტორი რეალიზდება პროგრამულად.

კლასიფიკატორის ფუნქცია ის არის, რომ გადამწოდებისაგან მიღებული საწყისი მონაცემების საფუძველზე, რომლებიც წარმოდგენილია ნიშან-თვისებათა ვექტორების სახით, ობიექტს მიანიჭოს წინასწარ ცნობილი m კლასიდან: $C_1, C_2, \dots, C_{m-1}, C_m = C_r$, სადაც C_r არის ამოვარდნების კლასი, ერთერთი კლასის ჭდე. კლასიფიკაციის პროცესის შესაბამისი ბლოკ-სქემა მოცემულია 8.3 სურათზე. საწყის მონაცემებს ობიექტის კლასიფიკაციისათვის წარმოადგენს d -განზომილების ნიშან-თვისებათა \mathbf{x} ვექტორი.



სურ. 8.3

კლასიფიკაციის სისტემაში ყოველ შესაძლო კლასს შეესაბამება ბლოკი, რომელშიც ინახება მოცემული კლასის შესახებ სწავლების პროცესში მიღებული გარკვეული \mathbf{K} ცნობე-

ბი და ასევე დამუშავების გარკვეული მახასიათებლები. გამოთვლების შედეგები m -ური ბლოკიდან გადაცემა გადაწყვეტილების მიმღებ ბლოკს, რომელიც წინა ეტაპზე გამოთვლილი სიდიდეების შედარების საფუძველზე იღებს გადაწყვეტილებას, თუ რომელ კლასს უნდა მიეკუთვნოს მოცემული ობიექტი.

8.2.1. კლასიფიკაციის სისტემის ცდომილებების შეფასება

ისევე როგორც ყველა სისტემას, კლასიფიკაციის სისტემის ფუნქციონირებისას ადგილი აქვს ცდომილებებს.

კლასიფიკაციის სისტემის ცდომილების კოეფიციენტი (error rate) წარმოადგენს ერთ-ერთ რიცხვით მახასიათებელს იმისა, თუ რამდენად კარგად წყვეტს სისტემა მასზე დაკისრებულ ამოცანას. სხვა მახასიათებლებს წარმოადგენს **კლასიფიკაციის სისტემის სიჩქარე** (დროის ერთეულოვან ინტერვალში დამუშავებული ობიექტების რაოდენობა) და **ღირებულება**, რომელიც შეიცავს აპარატურული და პროგრამული უზრუნველყოფის და სისტემის დაპროექტების ხარჯებს. რაც შეეხება სისტემის წარმადობას, ის დამოკიდებულია როგორც ცდომილებათა რაოდენობაზე, ასევე ამოვარდნების რიცხვზე. თუ სისტემა ყველა შემავალ ობიექტს აკლასიფიცირებს ამოვარდნების კლასში, მაშინ ეს იმას ნიშნავს, რომ სისტემა ცდომილებებს არ უშვებს, მაგრამ ამავე დროს იგი უსარგებლოა.

განსაზღვრება 8.1. *კლასიფიკაციის ცდომილებას ადგილი აქვს ყოველთვის, როცა კლასიფიკატორი შესასვლელ ობიექტს აკუთვნებს C_i კლასს, მაშინ როცა ჭეშმარიტ კლასს წარმოადგენს C_j , ამასთან $i \neq j$ და $C_i \neq C_r$, სადაც C_r - ამოვარდნების კლასია.*

განსაზღვრება 8.2. კლასიფიკაციის სისტემის ემპირიული ცდომილების კოეფიციენტი ტოლია დამოუკიდებელ ტესტურ მონაცემებზე დაშვებული ცდომილებების რაოდენობა გაყოფილი კლასიფიკაციის ცდათა რაოდენობაზე.

განსაზღვრება 8.3. კლასიფიკაციის სისტემის ემპირიული ამოვარდნების კოეფიციენტი ტოლია დამოუკიდებელ ტესტურ მონაცემებზე დაშვებული ამოვარდნების რაოდენობა გაყოფილი კლასიფიკაციის ცდათა რაოდენობაზე.

განსაზღვრება 8.4. დამოუკიდებელი ტესტური მონაცემები წარმოადგენს ცნობილი კლასების ეტალონურ ობიექტებს, ამოვარდნების კლასის ობიექტების ჩათვლით, რომლებიც არ იყენებენ გამოყენებული ნიშან-თვისებათა და კლასიფიკაციის ალგორითმების შემუშავებისას.

აღნიშნული ცდომილებები შეიძლება გამოყენებულ იქნეს პრაქტიკაში კლასიფიკაციის სისტემის წარმადობის ტესტირებისთვის. ამასთან, მკაცრად უნდა ვაკონტროლოთ, რომ სისტემის პროექტირების დროს გამოყენებული ეტალონური სახეები და სატესტო სახეები წარმოადგენდეს იმ ობიექტების წარმომადგენელ (რეპრეზენტაციულ) სახეებს, რომელსაც სისტემა მომავალში დაამუშავებს. ასევე, სისტემის ტესტირებისთვის გამოყენებული სახეები დამოუკიდებელი უნდა იყოს პროექტირების დროს გამოყენებული სახეებისგან. იმ შემთხვევაში, თუ საწყისი მონაცემები იმეორებს რომელიმე თეორიულ განაწილებას, ასეთ შემთხვევაში სისტემის წარმადობის შეფასებისთვის შესაძლებელია გამოთვლილ იქნეს თეორიული ცდომილების ალბათობა და არა ემპირიული ცდომილების კოეფიციენტი ტესტირების მონაცემებზე დაყრდნობით.

დავუშვათ, რომ ჯიბის კომპიუტერის ხელნაწერი სიმბოლოების ამომცნობი მოდული სწორად ამოიცნობს მომხმარებლის მიერ შეყვანილი სიმბოლოების 95%-ს. იმის გამო, რომ მომხმარებელს თვითონ შეუძლია შემდეგ მოახდინოს შემავალი დოკუმენტის რედაქტირება, 5%-იანი ცდომილების კოეფიციენტი წარმოადგენს დასაშვებ სიდიდეს. მაგალითად, მან უნდა გაითვალისწინოს, რომ ყურადღება უნდა მიაქციოს სიმბოლო 'ზ'-ის ზედა ნაწილის უფრო აკურატულად დაწერას, რათა სისტემას ის არ აერიოს სიმბოლო 'ნ'-ში. მაგრამ, როცა საქმე გვაქვს საბანკო სისტემასთან, რომელიც კითხულობს ხელნაწერ ციფრებს დეპოზიტური ბლანკიდან, 5%-იანი ცდომილება შეიძლება მიუღებელი იყოს.

8.2.2. მცდარი ამოქმედებები და მცდარი გამოტოვებები

ცდომილებების განსაკუთრებულ კლასს წარმოადგენს **ორი კლასის პრობლემა** (two-class problems). იგი წარმოიშობა მაშინ, როცა ორი კლასი ერთმანეთს არსობრივად ეწინააღმდეგება: ა) კარგი ობიექტი და ცუდი ობიექტი; ბ) ობიექტი იმყოფება გამოსახულებაზე და ობიექტი გამოსახულებაზე არ იმყოფება, ან გ) ადამიანს აღმოუჩინეს D დაავადება და არ აღმოუჩინეს. ასეთ შემთხვევებში ცდომილებებს განსაკუთრებული მნიშვნელობა აქვს და ამასთან ისინი არასიმეტრიულებია.

აღნიშნული შემთხვევებიდან საილუსტრაციოდ ყველაზე მეტად (გ) შემთხვევა გამოდგება. თუ სისტემა შეცდომით ამბობს, რომ ადამიანს აქვს D დაავადება, მაშინ ასეთ ცდომილებას **მცდარი ამოქმედება** ან **პოზიტიური ცდომილება** ეწოდება. ხოლო თუ სისტემა შეცდომით ამბობს, რომ ადამიანს არ აქვს D დაავადება, მაშინ ასეთ ცდომილებას **მცდარი გამოტოვება** ან **ნეგატიური ცდომილება** ეწოდება. ცხადია,

მცდარი ამოქმედების შემთხვევაში ადამიანმა დამატებითი ანალიზები უნდა ჩაიტაროს, ხოლო მცდარი გამოტოვების შემთხვევაში დიაგნოზი დაიკარგება და ადამიანი არ განიკურნება, რაც ადამიანისთვის შეიძლება საშიში აღმოჩნდეს. ამგვარად, აღნიშნულ ცდომილებებს განსხვავებული ფასი ადევს. შესაბამისად, კლასიფიკაციის სისტემას უნდა ჰქონდეს თვისება გადაწყვეტილების მიღებისას მოახდინოს მცდარი გამოტოვებების მინიმიზაცია მცდარი ამოქმედებების რაოდენობის გაზრდის ხარჯზეც კი.

ნაკლებად დრამატულია (ა) შემთხვევა. მის მაგალითს წარმოადგენს სისტემა, რომელიც ერთმანეთისგან არჩევს დაზიანებულ და დაუზიანებულ ალუბლებს. ამ შემთხვევაში, მცდარ ამოქმედებებს შედეგად მოჰყვება ის, რომ ალუბალი მოხდება ნამცხვარში, და არა საცალო ვაჭრობის ობიექტში. რაც შეეხება (ბ) შემთხვევას, აქ მცდარ ამოქმედებებს შეიძლება მოჰყვეს ზედმეტი ენერჯის დახარჯვა სინათლის ანთებაზე იმ მიზეზით, რომ სცენაზე ფაქტობრივად მოძრაობა არ ხდება, ხოლო სისტემა აჩვენებს, რომ გზაზე ავტომობილმა ჩაიარა, რასაც სინამდვილეში ადგილი არ ჰქონია.

8.3. კავშირი სისტემის სიზუსტესა და ამოკრების სისრულეს შორის

დოკუმენტების ან გამოსახულებების მოძიების (document retrieval, DR) ამოცანებში საჭიროა, რომ მომხმარებლის მოთხოვნებში არსებული ნიშან-თვისებების საფუძველზე ვიპოვოთ ინტერესის ობიექტები – C_1 კლასი, რომლებსაც შესაძლოა მოჰყვეს ასევე მცირე რაოდენობის, მომხმარებლისთვის არასაინტერესო ობიექტებიც – C_2 კლასი. მაგალითად, მომხმარებელმა შეიძლება შეუკვეთოს სურათები, რომლებზეც გამოსახული იქნება მზის

ჩასვლის სცენა ან ცხენი. კლასიფიკაციის სისტემის წარმადობა, რომლის დანიშნულებაც მომხმარებლის მოთხოვნების მომსახურებაა, ხასიათდება **სიზუსტით** და **ამოკრების სისრულით**.

განსაზღვრება 8.5. *დოკუმენტების მოძიების სისტემის სიზუსტე* ტოლია მოძიებული რელევანტური (C_1 კლასის ჭეშმარიტი ელემენტები) დოკუმენტების რაოდენობა გაყოფილი მოძიებული დოკუმენტების (C_1 კლასის ელემენტებს დამატებული C_2 კლასის ელემენტები (მცდარი ამოქმედებები)) საერთო რიცხვზე.

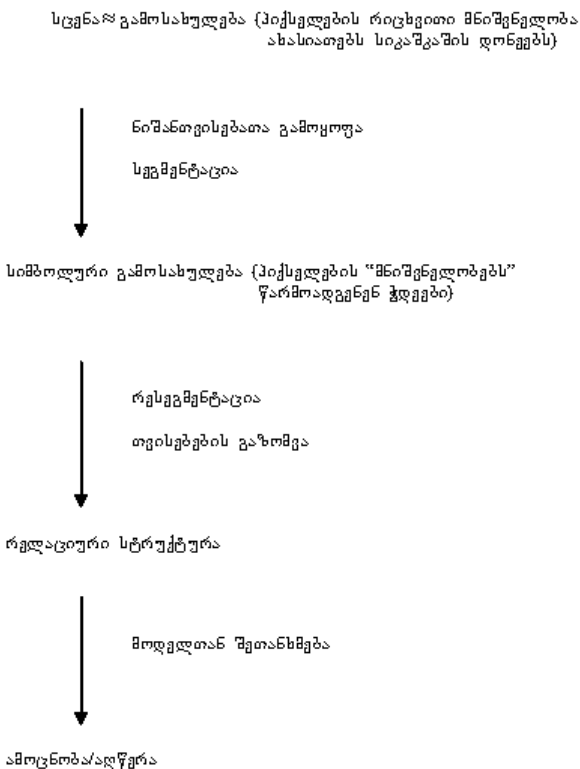
განსაზღვრება 8.6. *დოკუმენტების მოძიების სისტემის ამოკრების სისრულე* ტოლია სისტემის მიერ მოძიებული რელევანტური დოკუმენტების რიცხვი გაყოფილი რელევანტური დოკუმენტების საერთო რიცხვზე მონაცემთა ბაზაში.

მაგალითისთვის განვიხილოთ მონაცემთა ბაზა, რომელშიც ინახება 200 გამოსახულება მზის ჩასვლის სცენით და რომელთა მოძიებაზეც არის მომხმარებლის მოთხოვნა. ვთქვათ, საძიებო (კლასიფიკაციის) სისტემამ ძიების შედეგად ამ 200 რელევანტური გამოსახულებიდან იპოვა 150 სურათი და კიდევ მომხმარებლისთვის არარელევანტური (არასაინტერესო) 100 სურათი. მაშინ კლასიფიკაციის ამ ოპერაციის სიზუსტე ტოლი იქნება $150/250=60\%$, ხოლო ამოკრების სისრულე კი $150/200=75\%$. რა თქმა უნდა, თუ სისტემა დააბრუნებს უკან მონაცემთა ბაზის ყველა სურათს, მაშინ ის მიაღწევს 100% ამოკრების სისრულეს, მაგრამ შედეგად სისტემის სიზუსტე ძალიან დაბალი იქნება. მეორე მხრივ, თუ კლასიფიკაციის სისტემა მომართულია მცდარი ამოქმედებების დაბალ კოეფიციენტზე, მაშ-

ინ სისტემის სიზუსტე მაღალი იქნება, მაგრამ ამოკრების სისრულე ასეთ შემთხვევაში შეიძლება დაბალი აღმოჩნდეს.

8.4. სცენის გამოსახულებაზე ორგანოზომილებიანი ობიექტების ამოცნობის ზოგადი სტრუქტურა

ორგანოზომილებიანი სცენის ციფრულ გამოსახულებაზე ობიექტების ამოცნობის ზოგადი სტრუქტურა წარმოდგენილია 8.4 სურათზე.



სურ. 8.4

გამოსახულებაზე ობიექტი შეიძლება განისაზღვროს როგორც ურთიერთდაკავშირებული ნაწილების ერთობლიობა, რომელთა თვისებები (მაგალითად სიკაშკაშის დონეები, ტექსტურა, ზომები, ფორმა) და მიმართებები (მაგალითად შეფარდებითი ურთიერთმდებარეობა, შეფარდებითი განზომილება და ა.შ.) აკმაყოფილებს მოცემულ შეზღუდვებს. აქედან გამომდინარე, გამოსახულებაზე ობიექტების ამოსაცნობად საჭიროა გამოსახულებაზე მოიძებნოს ის არეები, რომლებიც შეესაბამება ობიექტის ნაწილებს.

პროცედურას, რომელიც გამოიყენება გამოსახულების ნაწილების მოსაძებნად, ეწოდება **სეგმენტაცია** ან **ნიშან-თვისებათა გამოყოფა**. ნიშან-თვისებათა გამოყოფის არსი ის არის, რომ გამოსახულებაზე გამოიყოს განსაკუთრებული ტიპის ლოკალური არეები, მაგალითად მკვეთრად გამოხატული კიდეები, სადაც ადგილი აქვს სიკაშკაშის დონის ნახტომისებურ ცვლილებას, წრფეები, მრუდები, ლაქები, კუთხეები და ა.შ. სეგმენტაციის საშუალებით ხდება პიქსელების მახასიათებელი ქვესიმრავლეების გამოყოფა, ანუ გამოსახულების დანაწევრება ისეთ ბმულ არეებად, რომლებიც ხასიათდება თვისებათა ერთგვაროვნებით.

სეგმენტაციისა და ნიშან-თვისებათა გამოყოფის პროცესები შეიძლება განხილულ იქნეს როგორც პიქსელებისთვის განსაკუთრებული იდენტიფიკატორების – **ჭდეების** მინიჭების პროცესი, რომელიც განსაზღვრავს მათი გარკვეული კლასისადმი მიკუთვნებას. ჭდეებად შეიძლება გამოყენებულ იქნეს როგორც მთელი რიცხვები, ასევე სიმბოლოები. ამგვარად, სეგმენტაციის პროცესის შედეგს წარმოადგენს არა პიქსელების, არამედ სიმბოლოების სახით აღწერილი გამოსახულება. ეს სიმბოლოები თავის მხრივ შეიძლება ისევ დაჯგუფებული იქნეს უკვე ახალ სიმრავლეებად სხვადასხვა კრიტერიუმების

საფუძველზე. ამ პროცესს **რესეგმენტაცია** ეწოდება. მაგალითად, პიქსელთა სიმრავლე შეიძლება დანაწევრდეს ბმულ არეებად; ერთმანეთისგან შეიძლება განვასხვაოთ რომელიმე არის ცალკეული ნაწილები (ჩაზნეკილი/ამოზნეკილი, წაგრძელებული და ა.შ.). ასევე სხვადასხვა გეომეტრიულ კრიტერიუმებზე დაყრდნობით შეიძლება რესეგმენტირებული იყოს ბმული რკალები და მრუდები (ან არის საზღვრები).

გამოსახულების რესეგმენტაციის შედეგად მიღებული ცალკეული ნაწილების კრებული საშუალებას გვაძლევს გავზომოთ გამოსახულების ნაწილების სხვადასხვა თვისებები და მათ შორის მიმართებები. ამ ნაწილების გეომეტრიული თვისებები (მაგალითად, ფართი, წაგრძელებულობა და ა.შ.) დამოკიდებულია მხოლოდ ამ ნაწილების შემადგენელი პიქსელების სიმრავლეზე, მაშინ, როცა სხვა თვისებები დამოკიდებულია ამ პიქსელების სიკაშკაშის დონეზე (მაგალითად, მათი საშუალო შეხურულობა).

ზემოთქმულიდან გამომდინარე, სცენის გამოსახულებაზე ობიექტების ამოცნობა შეიძლება განხილულ იქნეს როგორც გამოსახულებაზე ისეთი ქვესიმრავლეების პოვნის ამოცანა, რომელთა თვისებები და მიმართებები აკმაყოფილებს მოცემულ შეზღუდვებს.

გამოსახულებაზე არსებული ობიექტის ნაწილები თავიანთი თვისებებითა და მიმართებებით შეიძლება წარმოდგენილ იქნეს **რელაციური სტრუქტურის**, ანუ **მონიშნული გრაფის** საშუალებით (ობიექტის გრაფი), რომლის კვანძები წარმოადგენს ობიექტების ცალკეულ ნაწილებს, კვანძების ჭდეები – თვისებების მნიშვნელობებს, ხოლო რკალების ჭდეები მიმართებების მნიშვნელობებს. ამგვარად, გამოსახულებაზე ობიექტების ამოცნობა გულისხმობს გამოსახულების ნაწილების (ქვესიმრავლეების) გამოყოფას (ნიშან-თვისებათა გამოყოფის მეთოდე-

ბისა და სეგმენტაციის საშუალებით), მათი თვისებების გამოთვლას და მიმართებების განსაზღვრას და ამის საფუძველზე რელაციური სტრუქტურის აგებას. სხვა სიტყვებით, ამ გრაფის იმ ქვეგრაფების მოძიებას, რომლებიც მოცემული ქვეგრაფის მსგავსია. გამოსახულებაზე ობიექტების ამოცნობის ზემოაღწერილი პრინციპი გულისხმობს, რომ სცენა შეიძლება განხილულ იქნეს როგორც ორგანზომილებიანი. მაგრამ ეს პრინციპი გამოდგება სამგანზომილებიანი შემთხვევისთვისაც, როცა ცნობილია ობიექტის ორიენტაცია (მაგალითად, თუ ობიექტი მდგრადად არის მოთავსებული ბრტყელ ზედაპირზე და მას ზემოდან დავეყურებთ). ასეთ დროს სირთულე მდგომარეობს იმაში, რომ გამოსახულების ხილული ნაწილები შეიძლება არ შეესაბამებოდეს ობიექტის ნაწილებს. ამ დროს ობიექტები სცენაზე შეიძლება ერთმანეთს ფარავდეს და ადგილი ჰქონდეს მათ დამახინჯებას. ასეთი რამ შესაძლებელია 2D შემთხვევაშიც კი სეგმენტაციისას წარმოშობილი შეცდომების გამო.

8.5. ობიექტების წარმოდგენა ნიშან-თვისებათა ვექტორის საშუალებით

კლასიფიკაციის პროცესში ობიექტების აღსაწერად გამოიყენება მათი წარმოდგენა **გაზომვათა ანუ ნიშან-თვისებათა ვექტორების** სახით. ნიშან-თვისება წარმოადგენს დამოუკიდებელ გაზომვად სიდიდეს. ამოცნობის პროცესში ობიექტის აღსაწერად საჭიროა ავირჩიოთ ობიექტის მახასიათებელი d გაზომვადი სიდიდეები და წარმოვადგინოთ ისინი როგორც ერთი წერტილი d -განზომილების გაზომვათა ან ნიშან-თვისებათა სივრცეში. ნიშან-თვისებათა ვექტორი სწორედ ასეთი სიდიდეების ერთობლიობას წარმოადგენს. ვთქვათ, გაზომვათა (ნიშან-თვისებათა) სივრცე არის d განზომილების. მაშინ ასეთ სივრცეში

ნებისმიერი საკლასიფიკაციო A ობიექტის შესაბამისი ვექტორის i -ურ კოორდინატს ექნება ერთი და იგივე არსი. მაგალითად, ნებისმიერი ობიექტის აღმწერი ვექტორის პირველი კომპონენტი შეიძლება იყოს ობიექტის ფართი, მეორე – ობიექტის პერიმეტრი, მესამე – ხვრელების რაოდენობა ობიექტში და ა.შ. რაც შეეხება კომპონენტების რიცხვით მნიშვნელობებს, ისინი უფრო მოსახერხებელია წარმოვადგინოთ ნამდვილი რიცხვების სახით. კლასიფიკაციის ამოცანის ამოხსნისთვის ხდება ნიშანთვისებათა სივრცის დანაწევრება. ითვლება, რომ ნიშანთვისებათა სივრცის ის ნაწილი, რომელშიც აღმოჩნდება უცნობი ობიექტის ნიშანთვისებათა ვექტორი, განსაზღვრავს ამ ობიექტის მიკუთვნებას შესაბამისი კლასისადმი.

ვექტორთა მსგავსების რაოდენობრივი შეფასებისათვის კი შეიძლება ევკლიდური ან სხვა რომელიმე მეტრიკის გამოყენება.

განსაზღვრება 8.7. *d -განზომილების ნიშანთვისებათა სივრცეში ნიშანთვისებათა \mathbf{x}_1 და \mathbf{x}_2 ვექტორებს შორის ევკლიდური მანძილი ტოლია*

$$\|\mathbf{x}_1 - \mathbf{x}_2\| = \sqrt{\sum_{i=1}^d (\mathbf{x}_1[i] - \mathbf{x}_2[i])^2}$$

კლასიფიკაციის თეორიასა და პრაქტიკაში ასევე ძალიან მნიშვნელოვანია იმის განსაზღვრა, თუ რომელი ნიშანთვისებებია ინფორმაციული ამოცნობისთვის. მაგალითისთვის განვიხილოთ ხელნაწერი სიმბოლოების ამოცნობის ამოცანა. დაეუშვათ, რომ ცალკეული სიმბოლოების ერთმანეთისგან იზოლირება შესაძლებელია ბმული კომპონენტების მარკირების ალგორითმის გამოყენებით ან მომხმარებლისთვის მოთხოვნის დაყე-

ნებით, რომ სიმბოლოების ჩაწერა მოხდეს ცხადად გასარჩევ მართკუთხა არეებში. ჩავთვალოთ ასევე, რომ ბინარული გამო-სახულების სახით წარმოდგენილი თითოეული სიმბოლოსთვის V თავში გამოყენებული მეთოდების საშუალებით შესაძლებე-ლია შემდეგი ნიშან-თვისებების გამოთვლა:

- სიმბოლოს ფართობი (შავი ფერის პიქსელების რაოდენობა);
- აღმწერი მართკუთხედის სიმაღლე და სიგანე;
- ხვრელების რაოდენობა სიმბოლოში;
- სიმბოლოს მაფორმირებელი შტრიხების რაოდენობა;
- სიმბოლოს პიქსელთა სიმრავლის სიმძიმის ცენტრი;
- სიმბოლოს პიქსელებზე გამავალი იმ მთავარი ღერძის მიმართულება, რომლის მიმართაც ინერციის მომენტი იღებს უმცირეს მნიშვნელობას;
- სიმბოლოს პიქსელთა მეორე რიგის მომენტები მთავარი ღერძების (ინერციის უდიდესი და უმცირესი მომენტების მქონე ღერძების) მიმართ.

შემდეგ ინტუიციური მოსაზრებებისა და ეტალონური სიმბოლოების ანალიზის საფუძველზე შეიძლება სიმბოლოების ზემოვანხილული ნიშან-თვისებებისგან ცხრილის შედგენა (ცხ. 8.1). აღნიშნულ ცხრილში მოცემულია 10 სხვადასხვა სიმბოლოს 8 ნიშან-თვისების მნიშვნელობები. იმისთვის, რომ ცხრილში ნიშან-თვისებათა რიცხვითი მონაცემები უფრო ზუსტი გახდეს, საჭიროა გავითვალისწინოთ მონაცემები თითოეული სიმბოლოს დიდი რაოდენობის სახეებისგან. ამის შემდეგ უკვე შეიძლება კლასიფიკაციისთვის საჭირო გადაწყვეტილების მიღების პროცედურის ფორმულირება, ან პროტოტიპების იმ კრე-

ბულის ფორმირება, რომელიც შეპირისპირებისთვის გამოდგება.

განსაზღვრება 8.8. ცალკეული ობიექტის შეპირისპირების პროცესს ობიექტის პროტოტიპთან (მოდელთან) ან კლასის აღწერასთან ვერიფიკაცია ეწოდება.

ცხ. 8.1

სიმბოლო	ფართობი	სიმაღლე	სიგანე	ნახევრების რაოდენობა	შტრიხების რაოდენობა	სიმძიმის ცენტრი (Cx,Cy)	მთავარი ღერძი	ინერციის მომენტის მინიმ. მნიშვნელობა
'A'	საშუალო	დიდი	3/4	1	3	(1/2,2/3)	90	საშუალო
'B'	საშუალო	დიდი	3/4	2	1	(1/3,1/2)	90	დიდი
'8'	საშუალო	დიდი	2/3	2	0	(1/2,1/2)	90	საშუალო
'0'	საშუალო	დიდი	2/3	1	0	(1/2,1/2)	90	დიდი
'1'	მცირე	დიდი	1/4	0	1	(1/2,1/2)	90	მცირე
'W'	დიდი	დიდი	1	0	4	(1/2,2/3)	90	დიდი
'X'	დიდი	დიდი	3/4	0	2	(1/2,1/2)	?	დიდი
'^'	საშუალო	მცირე	1/2	0	0	(1/2,1/2)	?	დიდი
'_'	მცირე	მცირე	2/3	0	1	(1/2,1/2)	0	მცირე
'γ'	დიდი	დიდი	2/3	0	1	(1/2,1/2)	60	მცირე

ჩვენ მიერ სქემატურად წარმოდგენილი გადაწყვეტილების მიღების პროცედურა მოდელურ მაგალითს წარმოადგენს და სინამდვილეში შორსაა გადაწყვეტილების მიღების იმ პროცედურისგან, რომელიც შეძლებს გამართულად მუშაობას ხელნაწერი სიმბოლოების ამოცნობის რეალურ სისტემაში. მაგალითად, ძალიან ძნელია შტრიხების რაოდენობის საიმედო აღწერა და გამოთვლა.

8.6. კლასიფიკატორის რეალიზაციის ალგორითმები

სახეთა კლასიფიკაციის კლასიკური მიდგომის თანახმად საკლასიფიკაციო უცნობი ობიექტი წარმოიდგინება ელემენტარული ნიშან-თვისებების ვექტორის სახით. ასეთ მიდგომაზე

დამყარებული ამოცნობის სისტემა შეიძლება აგებულ იქნეს სხვადასხვა ხერხით. კერძოდ, ნიშან-თვისებათა ვექტორები სისტემისთვის შეიძლება ცნობილი გახდეს სწავლების პროცესის შედეგად ან პროგნოზირებული იყოს რაღაც მოდელების საფუძველზე. შემდგომში ჩვენ განვიხილავთ ამოცნობის სისტემის სწავლების შედეგად ფორმირებული ეტალონური სახეების მონაცემთა ბაზის გამოყენების ორ ალტერნატიულ მეთოდს.

8.6.1. კლასიფიკაცია უახლოესი საშუალო მნიშვნელობის მიხედვით

მონაცემთა კლასიფიკაციის მარტივი ალგორითმი, კლასიფიკაცია უახლოესი საშუალო მნიშვნელობის მიხედვით (the nearest class mean) მდგომარეობს თითოეულ კლასში შემავალი ეტალონური მონაცემების (სახეების) დაჯგუფებაში ამ კლასის მათემატიკური ლოდინის (საშუალო მნიშვნელობის) ვექტორის, ანუ სიმძიმის ცენტრის გამოყენებით

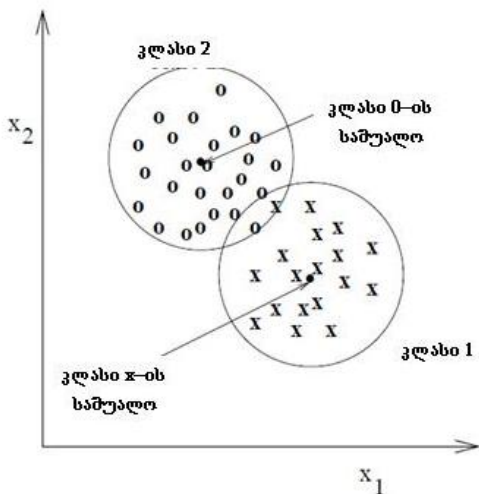
$$\bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ij} ,$$

სადაც \mathbf{x}_{ij} წარმოადგენს ნიშან-თვისებათა j -ურ ეტალონურ ვექტორს i -ური კლასისთვის, ხოლო n_i - ეტალონური ვექტორების რაოდენობაა i -ურ კლასში. უცნობი ობიექტი ნიშან-თვისებათა \mathbf{x} ვექტორით მიეკუთვნება i -ურ კლასს, თუ ის მნიშვნელოვნად უფრო ახლოსაა i -ური კლასის მათემატიკური ლოდინის ვექტორთან, ვიდრე სხვა კლასების მათემატიკური ლოდინის ვექტორებთან. თუ \mathbf{x} ვექტორი არც ერთი კლასის მათემატიკური ლოდინის ვექტორთან არ არის საკმაოდ ახლოს, მაშინ ის მიეკუთვნება ამოვარდნების კლასს. კლასიფიკაციის

ასეთი ალგორითმი მარტივი და საკმაოდ სწრაფია და გამოდგება ისეთი ამოცანებისთვის, სადაც თითოეულ კლასში მოხვედრილი ობიექტები, ერთი მხრივ, კომპაქტურადაა განლაგებული და, მეორე მხრივ, საკმაოდ დაშორებულია სხვა კლასებისგან. ასეთი კლასიფიკაციის მაგალითი, რომელშიც გვაქვს ორგანზომილებიანი ნიშან-თვისებათა ვექტორები, განაწილებული ორ კლასში, ნაჩვენებია 8.5 სურათზე, სადაც ერთი კლასის ობიექტები აღნიშნულია \times სიმბოლოთი, ხოლო მეორესი – 0.

იმის გამო, რომ თითოეულ კლასში არსებობს წერტილები, რომლებიც თანაბარი მანძილითაა დაშორებული ორივე კლასის ცენტრებისგან, კლასიფიკაციის ცდომილების კოეფიციენტი არანულოვანი იქნება. მოცემული ალგორითმისთვის 8.3 სურათზე წარმოდგენილი კლასიფიკაციის ფუნქციონალური ბლოკ-სქემა შემდეგნაირად შეიძლება იყოს ინტერპრეტირებული: i -ური ფუნქციონალური $f_i(\mathbf{x}, \mathbf{K})$ ბლოკი ითვლის მანძილს უცნობ შესასვლელ \mathbf{x} ვექტორსა და შესაბამისი ეტალონური სახეების კლასის მათემატიკური ლოდინის ვექტორს შორის. ეტალონური სახეები (ნიმუშები) წარმოადგენს თითოეული კლასის შესახებ \mathbf{K} აპრიორულ ინფორმაციას (ცნობებს).

სირთულეები წარმოიშობა მაშინ, როცა ეტალონური სახეების კლასებს აქვს რთული სტრუქტურა. 8.5 სურათზე ნაჩვენებია შემთხვევა, როცა სახეების კლასები კარგადაა ერთმანეთისგან გამოყოფილი (განმხილვადი), მაგრამ გამომდინარე მათი სტრუქტურიდან, ცენტრთან სიახლოვეზე დამყარებული კლასიფიკაცია ცუდ შედეგს მოგვცემს. ამას აქვს რამდენიმე მიზეზი. პირველი მიზეზი ის არის, რომ მეორე კლასი (სახეები აღნიშნულია 0-ით) მრავალმოდულია.

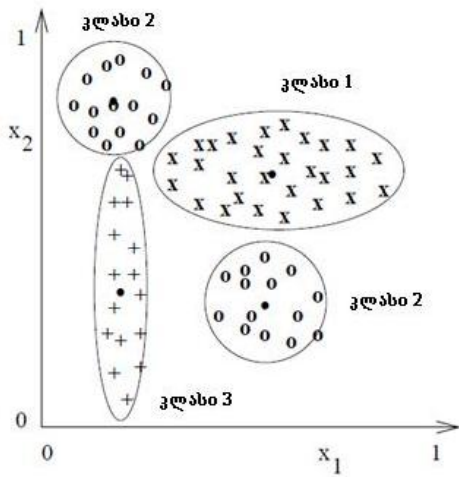


სურ. 8.5

მასში შემავალი სახეები განლაგებულია ორ ცალ-ცალკე მყოფ კომპაქტურ კლასტერებში, რომლებიც ცუდად აღიწერება ერთი საშუალო მნიშვნელობით, რომელიც წარმოადგენს ორ მოდას შორის მოთავსებულ წერტილს. პირველი კლასის (აღნიშნულია x -ით) რამდენიმე სახე უფრო ახლოსაა მეორე კლასის ცენტრთან, ვიდრე პირველის. სურათზე წარმოდგენილი სახეების შესწავლით ჩანს, რომ მეორე კლასში გვაქვს ორი მოდა და ამიტომ ის შეიძლება დავყოთ ორ ქვეკლასად, რომელთაც ექნება თავიანთი საშუალო მნიშვნელობები. ამის გაკეთება შეიძლება ორგანზომილებიანი წერტილოვანი გრაფიკის საშუალებით (სურ. 8.6), მაგრამ უფრო მეტი განზომილების შემთხვევაში სახეების სტრუქტურის გაგება ძალიან ძნელი იქნება.

მეორე მიზეზი მომდინარეობს პირველი და მესამე კლასების წაგრძელებულობიდან. ცხადია, მესამე კლასის სახეები, რომელთაც აქვს x_2 კოორდინატის დიდი მნიშვნელობები, უფრო ახლოსაა მეორე კლასის საშუალო მნიშვნელობასთან,

ვიდრე მესამეს. ანალოგიურად, მცირე x_1 კოორდინატის სიდიდის მქონე პირველი კლასის (X) სახეები ახლოს იქნება მეორე კლასის ერთ-ერთ ქვეკლასის საშუალო მნიშვნელობასთან იმ შემთხვევაშიც კი, თუ მას დაეყოფთ ორ მოდად. აღნიშნული პრობლემა ზოგიერთ შემთხვევაში შეიძლება თავიდან აცილებულ იქნეს მანძილის გამოთვლის წესის შეცვლის გზით, თუ მხედველობაში მივიღებთ სახეების განსხვავებულ განაწილებებს განზომილებების სხვადასხვა მიმართულებით.



სურ. 8.6

უცნობი ობიექტის ნიშან-თვისებათა \mathbf{x} ვექტორსა და კლასის მათემატიკური ლოდინის \mathbf{x}_c ვექტორს შორის “მოდიფიცირებული” მანძილი შესაძლებელია გამოთვლილ იქნეს ევკლიდური მანძილის კომპონენტების მასშტაბირებით C კლასის საშუალოკვადრატული σ_i გადახრის (კლასის ელემენტების “გაბნევის” მახასიათებლის) გამოყენების გზით ნიშან-თვისებათა

სივრცის თითოეული i -ური კოორდინატის მიმართულების გასწვრივ.

განსაზღვრება 8.8. *შკალირებული (მასშტაბირებული) ანუ ნორმირებული ევკლიდური მანძილი (scaled Euclidean distance) \mathbf{x} ვექტორსა და კლასის მათემატიკური ლოდინის \mathbf{x}_c ვექტორს შორის ტოლია*

$$\|\mathbf{x} - \mathbf{x}_c\| = \sqrt{\sum_{i=1}^d ((\mathbf{x}[i] - \mathbf{x}_c[i])/\sigma_i)^2}$$

შკალირება თითქმის ყოველთვისაა აუცილებელი, როცა სხვადასხვა კოორდინატა ღერძების გასწვრივ გამოიყენება სხვადასხვა განზომილების ერთეულები. მაგალითად, მოვახდინოთ ავტომანქანების კლასიფიკაცია ნიშან-თვისებების, $\mathbf{x}[1]$ = მანძილი მეტრებში და $\mathbf{x}[2]$ = წონა კილოგრამებში, მიხედვით. ასეთ შემთხვევაში მეორე ნიშან-თვისება შკალირების გამოყენების გარეშე მნიშვნელოვნად შეამცირებს ავტომობილების სიგრძეებს შორის განსხვავების გავლენას კლასიფიკაციის შედეგზე.

8.6 სურათზე წარმოდგენილი მაგალითის შემთხვევაში, როცა მეორე კლასის ორი მოდისთვის გვაქვს მათემატიკური ლოდინის ცალ-ცალკე მნიშვნელობები, ასეთ შემთხვევაში x_1 და x_2 ნიშან-თვისებების ცალ-ცალკე, კლასზე დამოკიდებული შკალირება მოგვცემს კლასიფიკაციის კარგ შედეგს. მაგრამ ამოცანების უმეტესობა ასეთი მარტივი არ არის. თუ ელიფსები, რომლებიც წარმოადგენს კლასებში სახეების განაწილებას, არ არის გაზომვათა სივრცის ღერძების გასწვრივ განლაგებული, როგორც ეს ნაჩვენებია სურათზე 8.6., მაშინ უცნობ ვექტ-

ტორსა და კლასის მათემატიკური ლოდინის ვექტორს შორის მანძილის კორექტული გამოთვლისთვის საჭიროა კოორდინატების გარდაქმნა. ეს საკითხი განხილული იქნება მოგვიანებით, ბაიესის კლასიფიკაციის შესწავლის დროს. უფრო რთულ ამოცანასთან გვაქვს საქმე, როცა გაზომვათა d -განზომილებიან სივრცეში სახეების სიმრავლის საზღვარი მრუდწირულია.

8.6.2. კლასიფიკაცია უახლოეს მეზობლამდე მანძილის მიხედვით

სახეთა კლასიფიკაციის უფრო მოქნილ, მაგრამ ამასთან რეალიზაციის დროს თვალსაზრისით უფრო ხანგრძლივ მეთოდს წარმოადგენს ალგორითმი, რომელშიც უცნობი ნიშან-თვისებათა \mathbf{x} ვექტორის მიკუთვნება ხდება იმ კლასისადმი, რომლის ცალკე აღებულ რომელიმე ეტალონურ სახესთან ეს ვექტორი უახლოესი მანძილითაა დაშორებული. კლასიფიკაციის ამ მიდგომას **უახლოესი მეზობლის წესი** (nearest neighbor rule) ეწოდება. ეს ალგორითმი ეფექტურია მაშინაც კი, როცა კლასებს d -განზომილებიან გაზომვათა სივრცეში რთული სტრუქტურა აქვს და როცა კლასები ერთმანეთს ფარავს. ასეთი მიდგომა არ მოითხოვს რაიმე წინასწარი ვარაუდების არსებობას გაზომვათა სივრცეში ნიშან-თვისებათა ვექტორების განაწილების მოდელების შესახებ. ალგორითმი იყენებს მხოლოდ ცნობილ ეტალონურ სახეებს.

ვთქვათ, ყოველი კლასისთვის ცნობილია შესაბამისი სახეები. აღვნიშნოთ i -ური კლასის j -ური სახის ნიშან-თვისებათა ვექტორი \mathbf{x}_{ij} . უახლოესი მეზობლის წესის მიხედვით \mathbf{x} უცნობი ვექტორის კლასიფიკაცია მდგომარეობს ამ ვექტორისადმი უახლოესი სახის მოძებნაში და იმ კლასის განსაზღვრაში, რომ-

ელსაც ის უნდა მიეკუთვნოს. რომელიმე k (კლასი) და l (სახე) ინდექსებისთვის ეს ნიშნავს, რომ თუ

$$|\mathbf{x}_{kl} - \mathbf{x}| < |\mathbf{x}_{ij} - \mathbf{x}|, \forall i, j,$$

მაშინ უცნობი \mathbf{x} მიეკუთვნება k კლასს.

არსებობს მონაცემთა ისეთი სტრუქტურები, რომლებიც გამოსათვლელი მანძილების რაოდენობის შემცირების შესაძლებლობას იძლევა. ასეთი სტრუქტურების მაგალითს წარმოადგენს მონაცემთა ხისებური და ბადისებური სტრუქტურები.

კლასიფიკაციის წარმოდგენილი მიდგომის დროს საუკეთესო საკლასიფიკაციო გადაწყვეტილების მიღება შეიძლება განხორციელდეს მონაცემთა ბაზაში უახლოესი k ნიშან-თვისებათა ვექტორების შემოწმებით. $k > 1$ -ზე მნიშვნელობები d -განზომილებიან გაზომვათა სივრცეში უზრუნველყოფს საუკეთესო ამონაკრებს ვექტორთა განაწილებისათვის. ეს კი განსაკუთრებით სასარგებლოა კლასტერთა გადაკვეთის არეებში. ასეთ არეში მოხვედრილი უცნობი ვექტორი შეიძლება მიეკუთვნოს იმ კლასს, რომლის წარმომადგენელიც მასთან უახლოესი აღმოჩნდა. ასეთი შემთხვევების დროს კლასიფიკაცია არსობრივად შემთხვევითია. კვლევებით დადგენილია, რომ სახეთა რაოდენობის უსასრულო გაზრდისას შეცდომების კოეფიციენტი ზღვარში, მაშინაც კი როცა $k = 1$, შეცდომების ოპტიმალურ კოეფიციენტს აჭარბებს არა უმეტეს ორჯერ. მართალია, თეორიულად შესაძლებელია უკეთესი შედეგის მიღება, როცა $k > 1$, ანუ თუ განვიხილავთ რამდენიმე უახლოეს მეზობელს, მაშინ უცნობ ნიშან-თვისებათა \mathbf{x} ვექტორი მიეკუთვნება იმ კლასს, რომელიც ყველაზე უფრო ხშირია k უახლოეს მეზობლებს შორის. მაგრამ k -ს დიდი მნიშვნელობების გამოყენების ეფექტურობა დამოკიდებულია იმაზე, არის თუ არა სახეების

შესაბამისი რაოდენობა გაზომვათა სივრცის ყოველ განსახილველ არეში, რათა შევწყვიტოთ უახლოესი მეზობლის ძიება უცნობი \mathbf{x} ვექტორისგან ძალიან დაშორებულ სახეებს შორის. ორი კლასის ამოცანაში, როცა გვაქვს $k = 3$ შემთხვევა, ჩვენ შეგვეძლო მოგვეხდინა \mathbf{x} -ის, როგორც იმ კლასის ელემენტის კლასიფიცირება, რომლის 2 ან 3 სახე \mathbf{x} -ის ახლოსაა განლაგებული. ორ კლასზე მეტის შემთხვევაში კი შესაძლო კომბინაციათა რაოდენობა მეტი ხდება და გადაწყვეტილების მიღებაც უფრო რთულია.

ამ მიდგომით კლასიფიკაციისას მნიშვნელოვან პრობლემას წარმოადგენს გამოთვლები. თუ მაგალითების რაოდენობა დიდია, მაშინ საჭიროა დიდი მეხსიერება. ასევე, თუ გაზომვათა სივრცე არ არის დანაწევრებული, მაშინ იძულებული ვიქნებით, რომ გამოვთვალოთ მანძილები უცნობ ვექტორსა და ყველა სახეს შორის. ამ მეთოდის დიდ უპირატესობას სხვებთან შედარებით წარმოადგენს ის, რომ კლასტერებს შეიძლება ჰქონდეს რთული ფორმა. მათთვის არ არის აუცილებელი, რომ იყვნენ ბრუნვის ღერძის მიმართ სიმეტრიული ან ამოზნექილი.

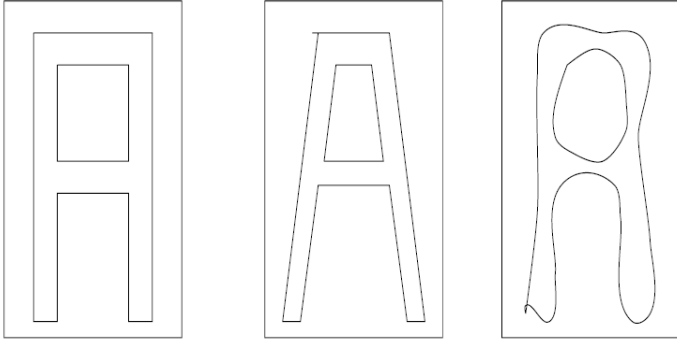
8.7. ამოცნობის სტრუქტურული მეთოდები

ხშირად ობიექტების ამოცნობისთვის შეიძლება საკმარისი არ აღმოჩნდეს ობიექტის მახასიათებელი უბრალო რიცხვითი ან სიმბოლური ნიშან-თვისებების გამოყენება, მაგალითად, სიმბოლოების ამოცნობის ამოცანაში. განვიხილოთ 2 სიმბოლოს შემთხვევა (სურ. 8.7). ორივე მათგანი ხასიათდება: ერთნაირი აღმწერი მართკუთხედებით, ნახვრეტების და შტრიხების (მონაკვეთების) ერთნაირი რაოდენობით, ერთნაირი სიმძიმის ცენტრებით, სტრიქონების და სვეტების გასწვრივ ერთმანეთის ტოლი მეორე რიგის მომენტებით და ასევე იმით, რომ ამ

ბების ფორმირების მიზნით. ასეთი მიდგომა დაედო საფუძვლად სახეთა სტრუქტურული ამოცნობის მეთოდებს.

სახეთა სტატისტიკური ამოცნობისას ობიექტის შესახებ არსობრივი ინფორმაცია წარმოიდგინება ნიშან-თვისებათა ვექტორების საშუალებით, რომლებიც ჩვეულებრივ წარმოადგენს ობიექტის აღმწერი რაღაც ელემენტარული სიდიდეების ვექტორებს, მაგალითად, რიცხვებს ან ლოგიკურ სიდიდეებს. ეს სიდიდეები წარმოადგენს ობიექტის გლობალური თვისებების გაზომვის შედეგებს, როგორცაა, მაგალითად, ფართობი ან ინერციის მომენტები. ჩვენ მიერ წარმოდგენილი სიმბოლოების ამოცნობის მაგალითი კი უფრო რთულია, რადგანაც მასში, გარდა ზემოაღნიშნულისა, ასევე ყოველი სიმბოლოსთვის იზომება ნახვრეტების და შტრიხების რაოდენობაც. ეს კი თავის მხრივ მოითხოვს ნახვრეტების მოძიების პროცედურის არსებობას და სეგმენტაციის ალგორითმს სიმბოლოს შტრიხებად დანაწევრებისთვის.

სიმბოლოთა სტრუქტურული ამოცნობისას კი ობიექტის შესახებ არსობრივი ინფორმაცია წარმოიდგინება არა მარტო გლობალური ნიშან-თვისებების, არამედ სიმბოლოთა ელემენტარული ნაწილების ერთობლიობის, მათი ატრიბუტების და ურთიერთმიმართებების სახით. 8.8 სურათზე ნაჩვენებია სამ სხვადასხვა სახეში წარმოდგენილი 'A' სიმბოლო, რომელთაც თითქმის ერთნაირი სტრუქტურა აქვს. თითოეული მათგანი შეიძლება წარმოვიდგინოთ როგორც შედგენილი შემდეგი 4 შტრიხისგან – მონაკვეთისგან: ორი ჰორიზონტალური და ორი ვერტიკალური ან დახრილი. ამასთან თითოეულ სიმბოლოს აქვს ნახვრეტი ზედა ნაწილში და უბე ქვემოთ, რომლებიც ერთმანეთისგან გამოყოფილია ჰორიზონტალური მონაკვეთით – ჩამკეპით.

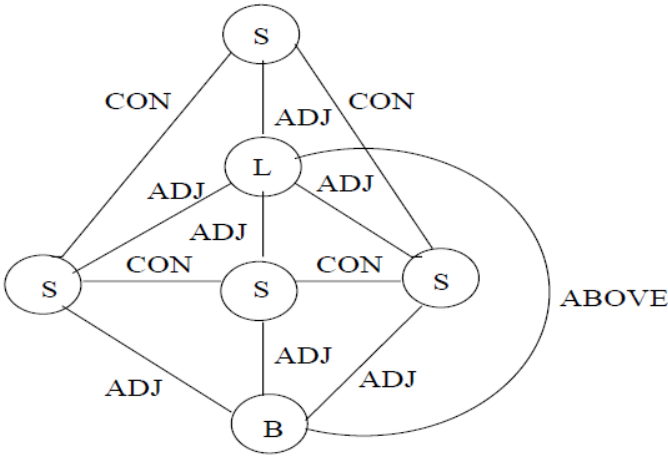


სურ. 8.8

თუ პრიმიტივებს შორის არსებობს ბინარული მიმართება, მაშინ ობიექტის არსის სტრუქტურული აღწერა შეიძლება წარმოდგენილი იქნეს გრაფის საშუალებით. დავუშვათ, რომ სიმბოლოების შემადგენელ მონაკვეთებს, უბეებს და ნახვრეტებს შორის შესაძლებელია, რომ არსებობდეს სახეთა ამოცნობისთვის სასარგებლო შემდეგი მიმართებები:

- CON (კავშირი): მიუთითებს ცალკეულ მონაკვეთებს შორის კავშირზე;
- ADJ (მომიჯნაობა): მიუთითებს, რომ მონაკვეთი ესაზღვრება ნახვრეტს ან უბეს;
- ABOVE (ზემოთ): მიუთითებს, რომ ერთი ნახვრეტი (ნახვრეტი ან უბე) მოთავსებულია მეორის ზემოთ.

8.9 სურათზე წარმოდგენილია 'A' სიმბოლოს სტრუქტურული აღწერა გრაფის საშუალებით ზემომოყვანილი ბინარული მიმართებების გამოყენებით (S – შტრიხი, L – ნახვრეტი, B – უბე).



სურ. 8.9

ხშირად სახეთა სტრუქტურული ამოცნობის რეალიზაცია შესაძლებელია გრაფთა შეპირისპირების (matching) ალგორითმების საშუალებით. ამავე დროს, ორ პრიმიტივს შორის არსებული მიმართება თავად შეიძლება განხილულ იქნეს როგორც ელემენტარული ნიშან-თვისება, შედიოდეს ნიშან-თვისებათა ვექტორში როგორც კომპონენტი და მხედველობაში იქნეს მიღებული სტატისტიკური გადაწყვეტილების მიღებისას. ერთ-ერთი ასეთი უმარტივესი ხერხი მდგომარეობს იმაში, რომ მოცემული სახისთვის შეიძლება გამოთვლილ იქნეს მოცემულ მიმართებათა რაოდენობა ორ სხვადასხვა ტიპის ნიშან-თვისებას შორის (მაგალითად, უბე მდებარეობს ჰორიზონტალური შტრიხის ქვემოთ). ამის შემდეგ ამ მიმართებათა რაოდენობის აღმრიცხავი მრიცხველის სიდიდე გამოიყენება როგორც ნიშან-თვისება ამოცნობისთვის.

სახეთა ამოცნობის სტრუქტურული მეთოდები სასარგებლოა ისეთი რთული სახეების ამოცნობისთვის, რომლებიც

თავის მხრივ შედგება უფრო დაბალი დონის ბევრი სახისგან. სარგებლიანობა ჩანს განსაკუთრებით მაშინ, როცა სცენაზე რამდენიმე ობიექტია განთავსებული.

8.8. უზუსტობათა მატრიცა

უზუსტობების მატრიცა გამოიყენება კლასიფიკაციის ექსპერიმენტების შედეგების წარმოსადგენად. 8.2 ცხრილში წარმოდგენილია უზუსტობათა მატრიცის ჰიპოტეტური მაგალითი ციფრების ამოცნობისას. ცხრილის i -ური სტრიქონისა და j -ური სვეტის გადაკვეთაში მოთავსებული ელემენტი ტოლია კლასიფიკაციის ექსპერიმენტების იმ რიცხვისა, რომელთა დროსაც i -ურ კლასში შემავალი ობიექტი მცდარად იყო კლასიფიცირებული როგორც j -ური კლასის ობიექტი. დიაგონალური $i = j$ ელემენტები კი შეესაბამება ჭეშმარიტ (სწორ) კლასიფიკაციას. ცხადია, იდეალურ კლასიფიკაციას შეესაბამება მატრიცა, რომლის ყველა არადიაგონალური ელემენტი 0-ის ტოლია. როცა არადიაგონალური ელემენტების მნიშვნელობები დიდია, ეს ნიშნავს, რომ კლასები არასწორადაა არჩეული და საჭირო ხდება ნიშან-თვისებათა პროცედურებისა და კლასიფიკაციის პროცესის დამატებითი ანალიზი. ტესტირების პროცედურის დროს უზუსტობების მატრიცა გვიჩვენებს იმ მოსალოდნელი შეცდომების ტიპებსა და კოეფიციენტებს, რომელთაც ადგილი ექნება მომდევნო ექსპერიმენტების დროს. 8.2 ცხრილში ნაჩვენებ მაგალითში ექსპერიმენტების დროს 1000 შესასვლელი ვექტორიდან 7 მიეკუთვნა ამოვარდნების კლასს (R). მე-9 კლასის 3 შესასვლელ ვექტორს არასწორად მიეკუთვნა მე-4 კლასის ჭდე, ხოლო მე-4 კლასის ორი ვექტორი არასწორად არის კლასიფიცირებული როგორც მე-9 კლასის ელემენტები. ამგვარად, ექსპერიმენტის დროს ჯამში 25 შესასვლელი ვექტორი

არასწორად იყო კლასიფიცირებული. გამომდინარე იმ დაშვებიდან, რომ სატესტო მონაცემები დამოუკიდებელია სასწავლო მონაცემებისაგან, ჩვენ შეგვიძლია გამოვთვალოთ ამოვარდნების ემპირიული კოეფიციენტი, რომელიც ტოლია $7/1000=0,007$ და ცდომილების საერთო კოეფიციენტი $25/1000=0,025$. რაც შეეხება მე-9 კლასის ელემენტებს, მათი ცდომილების კოეფიციენტი უდრის $5/100=0,05$.

ცხ. 8.2

ამოცნობის სისტემის მიერ იდენტიფიცირებული j-ური კლასი

	0'	1'	2'	3'	4'	5'	6'	7'	8'	9'	R	
i-ური ობიექტის ჰემმარიტი კლასი	0'	97	0	0	0	0	0	1	0	0	1	1
	1'	0	98	0	0	1	0	0	1	0	0	0
	2'	0	0	96	1	0	1		1	0	0	1
	3'	0	0	2	95	0	1	0	0	1	0	1
	4'	0	0	0	0	98	0	0	0	0	2	0
	5'	0	0	0	1	0	97	0	0	0	0	2
	6'	1	0	0	0	0	1	98	0	0	0	0
	7'	0	0	1	0	0	0	0	98	0	0	1
	8'	0	0	0	1	0	0	1	0	96	1	1
	9'	1	0	0	0	3	0	0	0	1	95	0

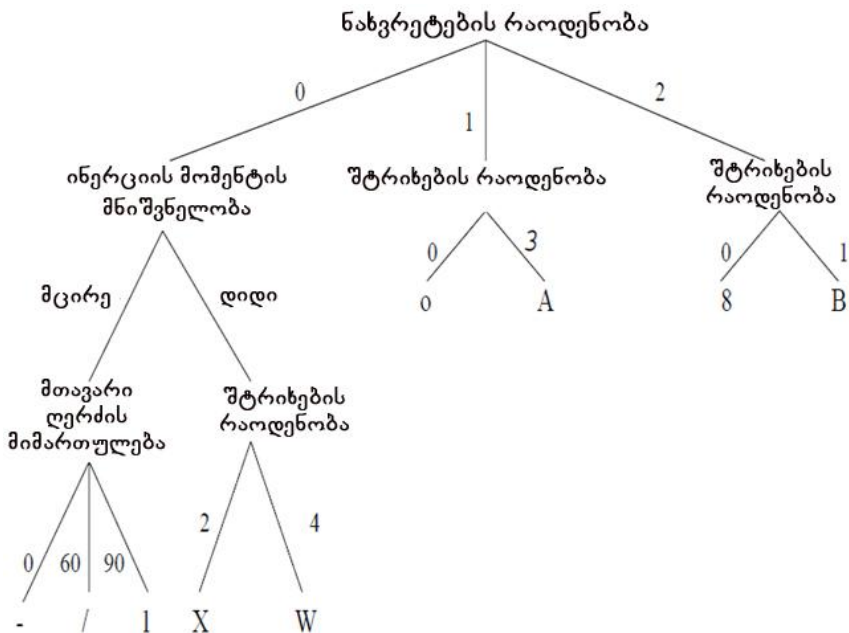
8.9. გადაწყვეტილებათა ხე

სახეთა ამოცნობის რთული ამოცანების გადაწყვეტისას, როცა საქმე გვაქვს სხვადასხვა ნიშან-თვისებების დიდ რაოდენობასთან და გვიხდება უცნობ ნიშან-თვისებათა ვექტორის შედარება მრავალ ეტალონური სახეების ნიშან-თვისებათა ვექტორთან, ამოცნობის პროცესი დიდ დროს მოითხოვს. ზოგიერთ

შემთხვევაში ასეთი შედარება დაუშვებელია, რადგან იგი შეიძლება დაკავშირებული იყოს დიდ დანახარჯებთან და დროსთან. მაგალითად, სამედიცინო დიაგნოსტიკაში, სადაც ნიშანთვისებათა გაზომვა შეიძლება მოითხოვდეს ძვირადღირებულ ლაბორატორიულ ანალიზებს და დანახარჯებს. ასეთ შემთხვევებში ობიექტების კლასიფიკაციისთვის მიმართავენ გადაწყვეტილების მიღების თანმიმდევრულ პროცედურას. გადაწყვეტილებათა ხის გამოყენება საშუალებას იძლევა, რომ მოვახდინოთ ნიშან-თვისებათა გამოყოფის და კლასიფიკაციის ეტაპების ურთიერთმონაცვლეობა.

განსაზღვრება 8.9. *გადაწყვეტილებათა ხე წარმოადგენს სტრუქტურას, რომლის საშუალებითაც ხდება კლასიფიკაციის პროცესის რეალიზება.*

აღნიშნული სტრუქტურა გადაწყვეტილების მიღებისთვის დროის ყოველ მომენტში იყენებს ერთ ან რამდენიმე ნიშან-თვისებას. გადაწყვეტილებათა ხე, რომელიც ასახავს კლასიფიკაციის პროცედურას ჩვენ მიერ განხილულ ხელნაწერი სიმბოლოების ამოცნობის ამოცანაში (ცხ. 8.1), მოცემულია 8.10 სურათზე. ამ ხის კვანძები წარმოადგენს ნიშან-თვისებათა ვექტორის სხვადასხვა ნიშან-თვისებებს. თითოეულ კვანძს განშტოებებით თავისი ნიშან-თვისების თითოეული შესაძლო მნიშვნელობისთვის აქვს თითო შვილობილი კვანძი. გადაწყვეტილების მიღების პროცედურა შვილობილ კვანძს ირჩევს მოცემული ნიშან-თვისების მნიშვნელობის საფუძველზე უცნობ ნიშან-თვისებათა ვექტორში. შვილობილი კვანძი შეიძლება მიუთითებდეს სხვა ნიშან-თვისების შემოწმებაზე ან შეიძლება აღმოჩნდეს ფურცლოვანი კვანძი, რომელიც შეიცავს გადაწყვეტილების ხეზე უკვე გავლილი გზის შედეგად მიღებულ კლასიფიკაციის შედეგს.

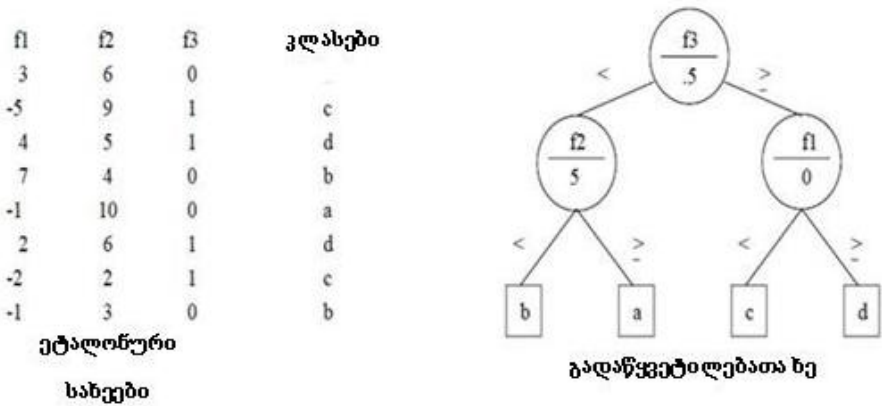


სურ. 8.10

განსაზღვრება 8.10. ბინარული გადაწყვეტილების ხე წარმოადგენს ბინარული ხის სახით წარმოდგენილ მონაცემთა სტრუქტურას, რომლის თითოეულ კვანძთან დაკავშირებულია გადაწყვეტილების არჩევის ფუნქცია. გადაწყვეტილების არჩევის ფუნქცია გამოიყენება ნიშან-თვისებათა უცნობი ვექტორის მიმართ და იგი განსაზღვრავს, მიმდინარე კვანძის რომელი შეილობილი კვანძი უნდა იქნეს შემდეგ ნაბიჯზე დამუშავებული – მარცხენა თუ მარჯვენა.

მარტივ შემთხვევაში, როცა ნიშან-თვისებათა მნიშვნელობები რიცხვებს წარმოადგენს, გადაწყვეტილების არჩევის ფუნქცია კვანძში უბრალოდ ერთმანეთს ადარებს უცნობი

ნიშან-თვისებათა ვექტორის კონკრეტული ნიშან-თვისების სიდიდეს ზღურბლთან და ირჩევს მარცხენა შვილობილ კვანძს, თუ ნიშან-თვისების აღნიშნული სიდიდე ნაკლებია ზღურბლზე, ან მარჯვენა შვილობილ კვანძს წინააღმდეგ შემთხვევაში. ამასთან, ხის ყოველ განშტოებად კვანძში ინახება მხოლოდ შესადარებელი ნიშან-თვისების ტიპი და ზღურბლი. თითოეულ ფურცლოვან კვანძში ინახება სახეთა კლასის ჭდე. იმ შემთხვევაში, როცა ხის შემოვლის პროცედურა მიაღწევს ფურცლოვან კვანძს, უცნობი ნიშან-თვისებათა ვექტორი კლასიფიცირდება როგორც სახეთა შესაბამისი კლასის ელემენტი. გადაწყვეტილებათა ხის სტრუქტურა, აგებული ეტალონური მონაცემების კორექტული კლასიფიკაციისთვის, მოცემულია 8.11 სურათზე.

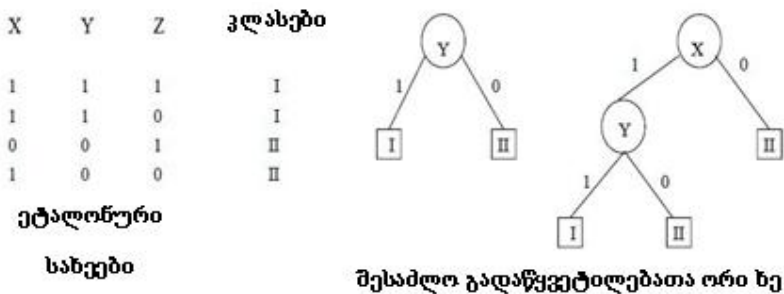


სურ. 8.11

სურათზე წარმოდგენილი გადაწყვეტილების ხე აგებულია ხელით. მოცემული ეტალონური მონაცემები უბრალოდ წარმოადგენს მოდელურ მაგალითს. მაშინ როცა რეალურ ამოცანებში გაცილებით მეტი რაოდენობის ნიშან-თვისებაა და კიდევ

უფრო მეტი ეტალონური სახე. მაგალითად, სამედიცინო დიაგნოსტიკაში რამდენიმე ასეული ნიშან-თვისებაა და რამდენიმე ათასი ეტალონური სახე. ასეთ შემთხვევაში გადაწყვეტილების ხის ხელით აგება უკვე შეუძლებელია. ამიტომ გადაწყვეტილების ხის ფორმირებისთვის მიმართავენ გადაწყვეტილების ხის ფორმირების ავტომატურ პროცედურას. ამასთან, მოცემული ეტალონური სახეების სიმრავლისთვის შეიძლება არსებობდეს ერთზე მეტი მაკლასიფიცირებელი გადაწყვეტილების ხე. ამიტომ, ძალიან მნიშვნელოვანია ისეთი ნიშან-თვისებების არჩევა, რომლებიც გარკვეული კრიტერიუმით საუკეთესო ხეს იძლევა.

ჩვეულებრივ, უპირატესობას ანიჭებენ მარტივ ანუ უმცირესი დონეების (შესასრულებელი შედარებების ნაკლები რაოდენობის) მქონე ხეებს. საილუსტრაციოდ განვიხილოთ ეტალონური მონაცემები და მათი კლასიფიკაციისთვის შესაძლო გადაწყვეტილებათა 2 ხე (სურ. 8.12).



სურ. 8.12

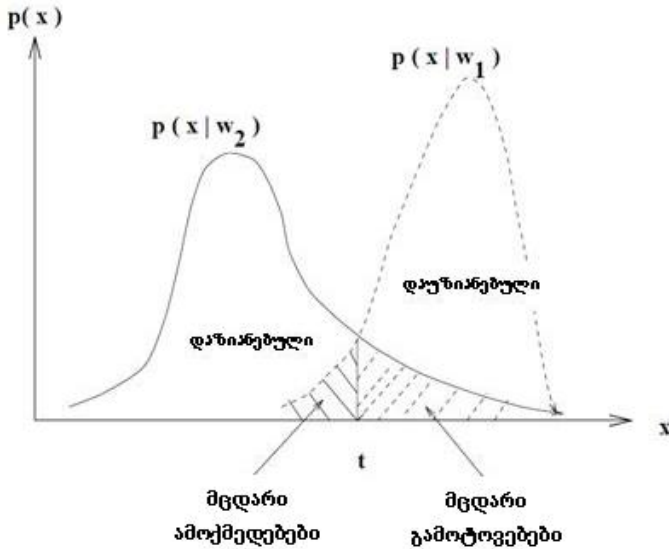
ორივე ხე საშუალებას იძლევა, რომ ეტალონურ სახეთა სიმრავლე დავეოთ ორ კლასად – კლასი I და კლასი II, როგორც ეს კარგად ჩანს თვით ეტალონური მონაცემებიდან. მაგრამ მარცხენა ხეს მარტივი სტრუქტურა აქვს და ნიშან-თვისებათა

ვექტორის კლასიფიკაციისთვის ის ერთ შედარებას ითხოვს, მაშინ როდესაც მეორე, უფრო დიდი ზომის მარჯვენა ხე შედარებების უფრო მეტ რიცხვს საჭიროებს.

8.10. გადაწყვეტილების მიღების ბაიესის წესი

ახლა განვიხილოთ გადაწყვეტილების მიღების ისეთი ამოცანა, როცა ალბათობათა განაწილებების შესახებ წინასწარ ანუ აპრიორულადაა ცნობილი და გვსურს განვხორციელოთ ისეთი კლასიფიკაცია, როცა მოსალოდნელი ცდომილების კოეფიციენტი უმცირესი იქნება.

დავუშვათ გვაქვს მუქი წითელი ფერის ალუბლის ინფრასითელი გამოსახულება და მასზე განხორციელდა ერთადერთი გაზომვა x , რომლის საფუძველზეც უნდა განისაზღვროს, ალუბალი დაზიანებულია თუ არა. ამასთან, ვთქვათ დაუზიანებელი ალუბალი მიეკუთვნება ω_1 კლასს, ხოლო დაზიანებული – ω_2 . ასევე დავუშვათ, რომ ჩვენ წინასწარ გავაანალიზეთ ბევრი დაზიანებული და დაუზიანებელი დიდი რაოდენობის ალუბლის ზედაპირზე არსებული არეები, ასე რომ ჩვენ უკვე წინასწარ გვაქვს განაწილების ფუნქციები (სურ. 8.13). ამ სურათზე მარჯვენა $p(x|\omega_1)$ მრუდი აღწერს გასაზომი x სიდიდის განაწილებას, აღებულს დიდი რაოდენობის დაუზიანებელი ალუბლების ზედაპირის არეებიდან, მარცხენა $p(x|\omega_2)$ მრუდი კი გასაზომი x სიდიდის განაწილებას, აღებულს დიდი რაოდენობის დაზიანებული ალუბლების ზედაპირის არეებიდან. ამავე დროს განხორციელდა მონაცემთა ისეთი ნორმირება, რომ თითოეული მრუდის ქვეშ არსებული ფართობი 1.0 ტოლი იყოს. ამის გამო, თითოეული მრუდი შეიძლება ჩაითვალოს როგორც ალბათობათა სიმკვრივე.



სურ. 8.13

სურათიდან ჩანს, რომ აღნიშნული მრუდები ერთმანეთს ნაწილობრივ ფარავს. ამის მიზეზი შემდეგში მდგომარეობს: ალუბლის ზედაპირის შეჭყლექილი არე შეიცავს წყალს, რომელიც ინფრაწითელ გამოსხივებას უფრო მეტად შთანთქავს, ვიდრე დაუზიანებელი არე. შესაბამისად ნაკლები არეკვლის მქონე არეების არსებობა უფრო ალბათურია დაზიანებული ალუბლებისთვის. იმის გამო, რომ წყლის შემცველობა ვარირებს, იცვლება ალუბლის ელფერიც და სწორედ ამის გამო განაწილების მრუდები კვეთს ერთმანეთს. გადაკვეთის არე შეესაბამება შემთხვევას, როცა ზოგიერთი მუქი ფერის დაუზიანებელი ალუბალი ისევე აირეკლავს ინფრაწითელ გამოსხივებას, როგორც ზოგიერთი ღია ფერის შეჭყლექილი ალუბალი.

თუ დაზიანებული და დაუზიანებელი ალუბლების გამოჩენა თანაბარალბათურია და თუ ყველა კლასიფიკაციის ცდო-

მიღების ფასი ერთნაირია, მაშინ შეიძლება მიღებულ იქნეს ω_1 გადაწყვეტილება, როცა $x > t$, და ω_2 წინააღმდეგ შემთხვევაში (t გადაწყვეტილების მიღების ზღურბლია).

გადაწყვეტილების მიღების ასეთი სტრატეგიის შემთხვევაში t ზღურბლის მარჯვნივ მდებარე დაშტრიხული არე წარმოადგენს მცდარი გამოტოვებების (არ ხდება სწორი გადაწყვეტილების მიღება) გაორმაგებულ კოეფიციენტს: ეს არის x სიდიდის დიდი მნიშვნელობის გაზომვის ალბათობა დაზიანებული ალუბლისთვის. იმის გამო, რომ თითოეული კლასის აპრიორული ალბათობა ტოლია 0.5 – დაშტრიხული არის ფართობი უდრის მცდარი გამოტოვებების გაორმაგებულ კოეფიციენტს და ამიტომ თითოეული ალბათობათა სიმკვრივე ისე უნდა იყოს მასშტაბირებული, რომ ერთი მრუდის ქვეშ არსებული საერთო ფართობი 0.5-ის ტოლი იყოს. t ზღურბლის მარცხნივ მდებარე დაშტრიხული არე წარმოადგენს მცდარი ამოქმედებების გაორმაგებულ ალბათობას. ეს არის ალბათობა იმისა, რომ კარგი ალუბალი მიეკუთვნება დაზიანებულებს, როცა $x < t$. იმის გამო, რომ კლასიფიკაციის სისტემის შესასვლელზე დაზიანებული და დაუზიანებელი ალუბლების ალბათობები თითქმის თანაბარია, ამიტომ სინამდვილეში თითოეული მრუდი მთელი ალბათობის 0.5-ს წარმოადგენს. ამის გამო მონიშნული არეების ფართობი ორჯერ მეტია მათ რეალურ ზომებთან შედარებით. კლასიფიკაციის ჯამური ცდომილება ტოლია ორივე მრუდის ქვეშ არსებული დაშტრიხული არეების ფართობების ჯამისა. უნდა აღინიშნოს, რომ გადაწყვეტილების მიღების t ზღურბლის მარცხნივ ან მარჯვნივ გადაწვეთ იზრდება დაშტრიხული არე და შესაბამისად იზრდება კლასიფიკაციის ცდომილებაც.

ჩვენ მიერ განხილული მაგალითი ასახავს იმ განსაკუთრებულ შემთხვევას, როცა გვაქვს ორი თანაბარალბათური

კლასი ცდომილების ერთნაირი ფასით. მაგრამ რეალურად კლასების რაოდენობა მეტია და ამასთან მათი აპრიორული ალბათობებიც შეიძლება განსხვავებული იყოს. ამიტომ განვიხილოთ კლასიფიკაციის ისეთი შემთხვევა, როცა გვაქვს m კლასი, ხოლო რაც შეეხება ცდომილებებს, გავამარტივოთ შემთხვევა და ჩავთვალოთ, რომ ისინი თანაბარია. გადაწყვეტილების მიღების ბაიესის პროცედურის მიხედვით ობიექტი უნდა მიეკუთვნოს იმ კლასს, რომელსაც ის ყველაზე მაღალი ალბათობით მიეკუთვნება.

განსაზღვრება 8.11. *ბაიესის კლასიფიკატორი დაკვირვებული ნიშან-თვისებების საფუძველზე ობიექტს აკუთვნებს იმ კლასს, რომელსაც ის ყველაზე მაღალი ალბათობით მიეკუთვნება.*

იმისათვის, რომ მოცემული გაზომილი x სიდიდით გამოვთვალოთ ალბათობები, საჭიროა ვიცოდეთ შემდეგი განაწილებები:

- პირობითი ალბათური განაწილება: $p(x|\omega_i)$, $\forall \omega_i$ კლასისთვის;
- აპრიორული ალბათობა: $P(\omega_i)$, $\forall \omega_i$ კლასისთვის;
- უპირობო ალბათური განაწილება: $p(x)$.

თუ ყველა ω_i კლასი ხასიათდება ყველა შესაძლო შემთხვევის მომცველი საკუთარი არათავსებადი ალბათობებით, მაშინ ბაიესის წესი შეიძლება გამოყენებული იყოს თითოეული კლასის აპოსტერიორული ალბათობების გამოსათვლელად ამ კლასების აპრიორული ალბათობებისა და x სიდიდის თითოეული კლასისთვის პირობითი ალბათური განაწილებით

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{p(x)} = \frac{p(x|\omega_i)P(\omega_i)}{\sum_{j=1}^m p(x|\omega_j)P(\omega_j)} \quad (8.1)$$

ახლა დავუბრუნდეთ კლასიფიკატორის ფუნქციონალურ ბლოკ-სქემას (სურ. 8.3). მის თითოეულ ბლოკში ჩვენ ვითვლით $f_i(x, \mathbf{K}) = P(\omega_i|x)$, რომლის გამოთვლაც ბაიესის წესის (8.1) მიხედვით შეიძლება, როგორც $p(x|\omega_i)P(\omega_i)/p(x)$. რადგანაც $p(x)$ ერთნაირია ყველა ბლოკისთვის, ამიტომ ის შეიძლება მხედველობაში არ მივიღოთ და საკლასიფიკაციო ω_i გადაწყვეტილება x სიდიდის რომელიმე ω_i კლასისადმი მიკუთვნების შესახებ მივიღოთ $p(x|\omega_i)P(\omega_i)$ -ის მაქსიმალური მნიშვნელობის მიხედვით.

ბაიესის კლასიფიკატორის პროექტირებისას აუცილებელია გვქონდეს \mathbf{K} ცნობა (ინფორმაცია) თითოეული კლასის აპრიორული $P(\omega_i)$ ალბათური და კლასის $p(x|\omega_i)$ პირობითი ალბათური განაწილების შესახებ. ამ სიდიდეების ცოდნა საშუალებას იძლევა ამოცნობის პროცესში მივიღოთ ოპტიმალური გადაწყვეტილება. აქ სირთულეს წარმოადგენს აპრიორული ალბათობების განსაზღვრა. მაგალითად, როგორ განვსაზღვროთ იმის ალბათობა, რომ დამხარისხებულ მოწყობილობაში შემავალი ნებისმიერი ალუბალი დაზიანებული აღმოჩნდება? თუ ამ ხდომილების ალბათობა დამოკიდებულია ამინდზე და მკრეფავთა ჯგუფზე, მაშინ ამ პირობების ყოველი ცვლილების დროს აუცილებელი ინფორმაციის მისაღებად შეიძლება საჭირო გახდეს ეტალონური მონაცემების დიდი რაოდენობა.

პრაქტიკული ამოცანების გადაწყვეტისას აუცილებელია, რომ $p(x|\omega_i)$ რაღაცნაირად გამოვთვლოთ. ამისთვის იყენებენ ემპირიულ მეთოდს, რომელიც გულისხმობს x გაზომილი სიდიდეების დიაპაზონის ინტერვალებად დანაწევრებას, თითოეულ ამ ინტერვალში სახეებს შორის x სიდიდის გამოჩენის

სისშირის გამოთვლას, და ამ გამოთვლილი სისშირეების მასივში ან ჰისტოგრამაში შენახვას. ამასთან, უნდა აღინიშნოს, რომ აუცილებელია ალბათობის ფუნქციის ნორმირება ისეთნაირად, რომ x -ის ყველა შესაძლო მნიშვნელობების ჯამი ტოლი იყოს 1.0.

იმ შემთხვევაში, თუ x -ის განაწილება რომელიმე ერთ-ერთი ცნობილი პარამეტრული მოდელის მსგავსია, მაშინ შეიძლება, რომ განაწილება წარმოვადგინოთ მისი მახასიათებელი პარამეტრების მცირე რაოდენობის საშუალებით. ჩვეულებრივად იყენებენ პუასონის განაწილებებს – ექსპონენციალურს და ნორმალურს (გაუსის).

8.11. გამოსახულებების ნიშან-თვისებები

გამოსახულების ნიშან-თვისება ეწოდება მის უმარტივეს განმასხვავებელ თვისებას ან მახასიათებელს. არსებობს ბუნებრივი ტიპის ნიშან-თვისებები, რომელთა არსი მდგომარეობს იმაში, რომ მათი დადგენა შესაძლებელია გამოსახულების ვიზუალური ანალიზის შედეგად. ნიშან-თვისებათა მეორე ჯგუფს კი წარმოადგენს ე.წ. ხელოვნური ნიშან-თვისებები, რომლებიც მიიღება გამოსახულების სპეციალური დამუშავების ან გაზომვების შედეგად. ბუნებრივ ნიშან-თვისებებს მიეკუთვნება: ინტენსივობა, ობიექტის კოორდინატები გამოსახულებაზე, გამოსახულების ცალკეული არეების ტექსტურა, ობიექტების კონტურების ფორმა. ხოლო ხელოვნურს – სიკაშკაშეების განაწილების ჰისტოგრამები, სივრცითი სისშირეების სპექტრები.

როგორც ზემოთ აღვნიშნეთ, გამოსახულებათა კლასიფიკაცია მოითხოვს შესაბამისი ნიშან-თვისებების შერჩევას. მათი ფორმირება წარმოადგენს პირველ ეტაპს ნებისმიერ სახეთა ამოცნობის სისტემისთვის და აქედან გამომდინარე, ამ სისტე-

მის ფუნქციონირების ხარისხიც მათ შერჩევაზეა დამოკიდებული. ნიშან-თვისებათა ფორმირების პროცესი რთულ ამოცანას წარმოადგენს, რადგანაც მათი საშუალებით გამოსახულების აღწერა ანუ ნიშან-თვისებათა კრებულის აგება ჯერ-ჯერობით ევრისტიკულ პროცედურას წარმოადგენს (არ არის ფორმალური) და ბევრად არის დამოკიდებული თვით მკვლევარის გამოცდილებასა და კვალიფიკაციაზე.

გამოსახულებათა თვისებების შესწავლისას სასურველია, რომ ცალკე იქნეს შესწავლილი გამოსახულებაზე ის ნიშან-თვისებები, რომლებიც ბუნებრივი ხასიათისაა, და ცალკე ნიშან-თვისებები, რომლებიც წარმოიშობა გამოსახულებაზე გარკვეული ტიპის ქვესიმრავლეების (ობიექტების) გამოყოფისას. ეს უკანასკნელები წარმოადგენს ქვესიმრავლეების ნიშან-თვისებებს. ინტერესი მათ მიმართ განპირობებულია იმით, რომ ხშირად საჭიროა გამოსახულებაზე არსებული ობიექტების ზომებისა და კონფიგურაციის განსაზღვრა. გამოსახულებებზე ობიექტების გამოყოფა ანუ გამოსახულებათა დანაწევრება ქვესიმრავლეებად სეგმენტაციის ამოცანას წარმოადგენს. ხოლო ამ ქვესიმრავლეების გეომეტრიული თვისებები, დაკავშირებული ობიექტების ტოპოლოგიასთან, მეტრიკულ თვისებებთან, ფორმასთან და ა.შ. განისაზღვრება გეომეტრიული ნიშან-თვისებების საშუალებით.

ასე რომ, სახეთა ამოცნობის ყოველი ახალი სისტემის აგებისას საჭირო ხდება როგორც საკვლევი გამოსახულების, ასევე მასზე არსებული ობიექტების სპეციფიკის გათვალისწინება. მიუხედავად ამისა, წლების განმავლობაში დაგროვილმა ცოდნამ შესაძლებელი გახადა, რომ გამოყოფილიყო ნიშან-თვისებათა ის ძირითადი ჯგუფები, რომლებიც წარმატებით გამოიყენება გამოსახულებათა აღწერისა და ამოცნობისთვის. ქვემოთ განვიხილავთ ამ ჯგუფებს ცალ-ცალკე.

8.11.1. სიკაშკაშის ნიშან-თვისებები

განათებულობა წარმოადგენს გამოსახულების უმნიშვნელოვანეს ნიშან-თვისებას, რომელიც გამოიხატება ისეთ სიდიდეებში, როგორცაა სიკაშკაშე, გამოსხივების სპექტრალური ინტენსივობა, ფერის კოორდინატები და ა.შ. მათ სიკაშკაშის ნიშან-თვისებები ეწოდება. მათი გაზომვა შეიძლება როგორც გამოსახულების ცალკეულ წერტილებში, ასევე მათ მიდამოში. მაგალითად, გამოსახულების რომელიმე (j, k) წერტილის $(2w + 1) \times (2w + 1)$ მიდამოში (ფანჯარაში) საშუალო სიკაშკაშის მნიშვნელობა განისაზღვრება ფორმულით

$$\bar{Y}(j, k) = \left[\frac{1}{(2w + 1)^2} \sum_{m=-w}^w \sum_{n=-w}^w Y(j + m, k + n) \right].$$

სიკაშკაშის ნიშან-თვისებების გაზომვა განსაკუთრებით მნიშვნელოვანია მაშინ, როცა აუცილებელი ხდება გამოსახულებაზე გამოსახული ობიექტების გამოყოფა და ასევე მათი მარკირება.

8.11.2. ალბათური ნიშან-თვისებები

ამ ჯგუფის ნიშან-თვისებების არსი გამომდინარეობს გამოსახულების სტატისტიკური აღწერიდან. კერძოდ, გამოსახულებაზე ინტენსივობის ფუნქცია განიხილება როგორც სტაციონარული შემთხვევითი პროცესის ან პროცესების (ფერადი გამოსახულების შემთხვევაში) რეალიზაცია. ასეთ შემთხვევაში გამოსახულების ნიშან-თვისებებს წარმოადგენს შემთხვევითი პროცესის რიცხვითი მახასიათებლები, რომელთაგან ყველაზე უფრო ხშირად გამოიყენება

- **სიკაშკაშის მახასიათებლები**, როგორცაა გამოსახულებაზე სიკაშკაშის მნიშვნელობების განაწილების ჰისტოგრამა, სიკაშკაშის ფუნქციის საწყისი α_k და ცენტრალური μ_k მომენტები;
- **გამოსახულების ტექსტურული მახასიათებლები**, რომლებსაც მიეკუთვნება შემთხვევითი პროცესის კორელაციური თვისებების განმსაზღვრელი მახასიათებლები, როგორცაა გამოსახულების კორელაციის კოეფიციენტი, გამოსახულების კორელაციის ფუნქცია და ა.შ.;
- **გამოსახულების ენერგეტიკული მახასიათებლები**, რომელთაც მიეკუთვნება მათი ენერგეტიკული სპექტრის ანათვლები;
- **სპექტრალური ნიშან-თვისებები**, რომელთა გამოსაყოფად გამოიყენება გამოსახულების გარდაქმნის სპექტრალური მოდელი. ასეთ გარდაქმნებს, რომლებიც გამოსახულების სპექტრს იძლევა მიეკუთვნება კარუნენ-ლოუვის გაშლა, ფურიე-გარდაქმნა (გამოსახულების ფურიე ნიშან-თვისებები), კოსინუს გარდაქმნა (კოსინუსური ნიშან-თვისებები), რადონის გარდაქმნა.

8.11.3. ჰისტოგრამული ნიშან-თვისებები

გამოსახულებაზე სიკაშკაშეების განაწილების ჰისტოგრამა შეიძლება გამოყენებული იქნეს ნიშან-თვისებათა გარკვეული კლასის შესაქმნელად. ჰისტოგრამის ფორმა საკმაოდ ინფორმატივას იძლევა გამოსახულების თვისებების შესახებ. მაგალითად, თუ ჰისტოგრამას აქვს ვიწრო ფორმა, ეს ნიშნავს, რომ გამოსახულება ხასიათდება დაბალი კონტრასტულობით, ანუ საქმე გვაქვს სიკაშკაშეების დონეების სიმცირესთან, ხოლო ბიძოდა-

ლური ჰისტოგრამის შემთხვევაში გამოსახულება ხასიათდება მკვეთრი სიკაშკაშეების მქონე არეების არსებობით.

8.11.4. გეომეტრიული ნიშან-თვისებები

ამ ჯგუფს მიეკუთვნება ის ნიშან-თვისებები, რომლებიც დაკავშირებულია გამოსახულებაზე არსებული ობიექტების გეომეტრიულ მახასიათებლებთან (მაგალითად, ფართობი, სიმძიმის ცენტრი, ექსტრემალური წერტილები, მანძილი, კუთხეები, მომენტები და ა.შ.). კერძოდ, ისინი შეიძლება წარმოადგენდეს გამოსახულების გარკვეული თვისებების მახასიათებლებს:

- გამოსახულებაზე არსებული ობიექტის გეომეტრიული ზომები ჰორიზონტალური ან ვერტიკალური მიმართულებით;
- მანძილი ობიექტის ორ ყველაზე უკიდურესად დაშორებულ წერტილს შორის;
- ობიექტის პერიმეტრი და ფართობი;
- ობიექტის კომპაქტურობა (შეფარდება მის ფართობსა და პერიმეტრს შორის);
- გამოსახულებაზე არსებული ობიექტის შესაბამისი გეომეტრიული ფიგურების (წრეწირი, მრავალკუთხედი და ა.შ.) რიცხვითი მახასიათებლები.

ობიექტის გეომეტრიული მახასიათებლების რაოდენობრივი შეფასების შედეგად მიღებული მნიშვნელობები წარმოადგენს მეტრიკულ ნიშან-თვისებებს. გამოსახულების მეტრიკული მახასიათებლები ეფუძნება მანძილის გაზომვას გამოსახულებაზე არსებული ობიექტის წერტილებს შორის. გამოსახულებ-

ის ანალიზისას გამოყენებული მეტრიკების უმრავლესობას აქვს სახე:

- ევკლიდური მეტრიკა

$$d_E = [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2};$$

- აბსოლუტური მანძილი

$$d_m = |x_i - x_j| + |y_i - y_j|;$$

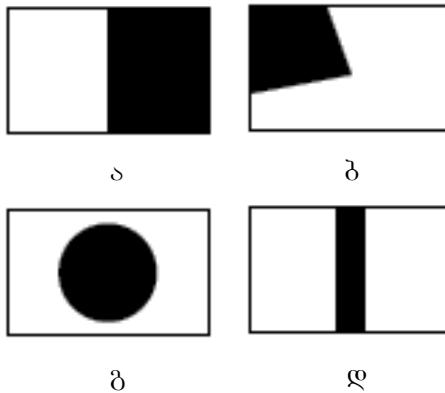
- მაქსიმალური მანძილი

$$d_x = \max\{|x_i - x_j|, |y_i - y_j|\}.$$

დისკრეტული გამოსახულების შემთხვევაში კოორდინატთა სხვაობები წარმოადგენს მთელ რიცხვებს, ხოლო ევკლიდური მანძილი კი ჩვეულებრივად არ წარმოადგენს მთელ რიცხვს, რაც შეცდომის მიზეზი ხდება რიცხვის დამრგვალებისას ან ჩამოჭრისას ციფრული დამუშავების დროს.

გეომეტრიული ნიშან-თვისებებიდან ძალიან სპეციფიკურს და მნიშვნელოვანს წარმოადგენს ობიექტის კონტურის გეომეტრიასთან დაკავშირებული ნიშან-თვისებები. როგორც ვიცით, ადამიანი ობიექტებს მათი კონტურის მიხედვით აღიქვამს. აქედან გამომდინარე, კონტურის აღწერა გამოსახულების ამოცნობისთვის სასარგებლო მახასიათებლებს იძლევა. ობიექტის კონტურთან დაკავშირებულ ლოკალურ ნიშან-თვისებებს წარმოადგენს:

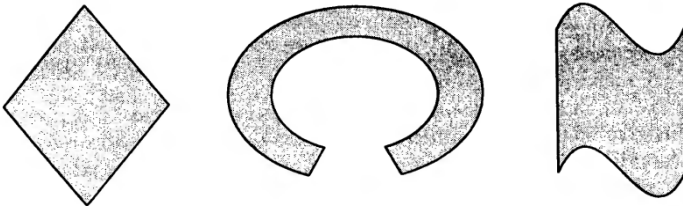
- **კიდე** – სიკაშკაშის მკვეთრი ვარდნის ადგილი კონტურის განივი მიმართულებით (სურ. 8.14,ა);
- **კუთხე** – კონტურის მკვეთრი მობრუნების (მიმართულების ცვლილების) ადგილი (სურ. 8.14,ბ).
- **ლაქა** – გამოსახულებაზე შედარებით მცირე ზომის არე, რომელშიც სიკაშკაშის მნიშვნელობა დაახლოებით მუდმივია და ამასთან მნიშვნელოვნად განსხვავდება მისი მიდამოს სიკაშკაშის მნიშვნელობებისგან (სურ. 8.14,გ);
- **მონაკვეთი** – გამოსახულებაზე სიკაშკაშისა და სიგანის დაახლოებით მუდმივი მნიშვნელობის მქონე ვიწრო წაგრძელებული არე. სიკაშკაშის მნიშვნელობების განაწილებას მონაკვეთის პერპენდიკულარული მიმართულებით აქვს U-ს (ან ი-ს) მაგვარი ფორმა (სურ. 8.14,დ).



სურ. 8.14

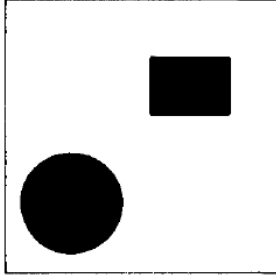
8.11.5. ტოპოლოგიური ნიშან-თვისებები

აღნიშნულს მიეკუთვნება ის ნიშან-თვისებები, რომლებიც ახასიათებს გამოსახულებაზე არსებული ობიექტის **ტოპოლოგიურ თვისებებს**. ამ უკანასკნელის ქვეშ იგულისხმება ის თვისებები, რომლებიც ინვარიანტულია ტოპოლოგიური ანუ ჰომეომორფული გარდაქმნების მიმართ. კერძოდ, ურთიერთცალსახა უწყვეტი (პირდაპირ და უკუ) გარდაქმნები. ამის გამო ტოპოლოგიური ნიშან-თვისებები აპრიორულად ინვარიანტულია გეომეტრიული გარდაქმნების ფართო სპექტრის მიმართ. ნიშან-თვისებათა აღნიშნული თანხვედრის გამო ტოპოლოგიური ნიშან-თვისებები საკმაოდ არაინფორმაციულია და შედეგად ვიღებთ იმას, რომ ობიექტები, რომლებსაც სხვადასხვა გეომეტრია აქვს, ტოპოლოგიური აღწერის თვალსაზრისით ერთმანეთისგან არ განიხრევა. ასეთი ობიექტები ნაჩვენებია 8.15 სურათზე.



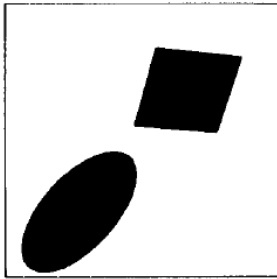
სურ. 8.15

რაც შეეხება ინვარიანტულობას გარდაქმნის მიმართ, მის მაგალითად შეიძლება წარმოვიდგინოთ ე.წ. “ელასტიკური” დეფორმაცია. კერძოდ, რეზინის ფურცლის გაჭიმვა, რომელზედაც დახატულია წინასწარ მოცემული ფორმის მქონე ობიექტი (სურ. 8.16).



სურ. 8.16

გაჭიმვის შედეგად ვლუბლობთ ამ ობიექტის გარკვეულ სივრცით დეფორმაციას (სურ. 8.17). ისეთი გარდაქმნები, რომლის დროსაც მოხდება რეზინის ფურცლის გაწყვეტა ან მისი ერთი ნაწილის შეერთება მეორესთან, დაუშვებელია. ცხადია, მანძილი არ წარმოადგენს ტოპოლოგიურ თვისებას, რადგანაც რეზინის ფურცლის დაჭიმვისას ის შეიძლება იცვლებოდეს.



სურ. 8.17

ასევე არ წარმოადგენს ტოპოლოგიურ თვისებას წირების პერპენდიკულარობა და პარალელურობა. ამავე დროს უნდა აღინიშნოს, რომ არ არსებობს გაჭიმვის ისეთი ოპერაციები, რომელთა მოქმედების შედეგად მოხდება ობიექტის ბმულობის გაზრდა ან შემცირება.

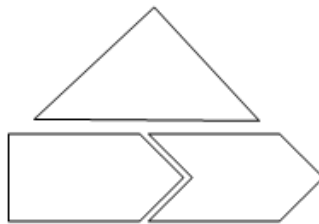
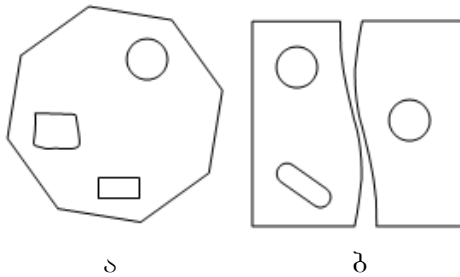
ტოპოლოგიური ნიშან-თვისებების რაოდენობა დიდი არაა და ამასთან მათი გამოთვლა საკმაოდ რთულია. განვიხილოთ ზოგიერთი მათგანი.

- **ობიექტის ბმული კომპონენტების რაოდენობა** – ეს არის ობიექტის შემადგენელი ბმული კომპონენტების ის მინიმალური რაოდენობა, რომელთაგან თითოეულში ნებისმიერი ორი წერტილი შეიძლება შეერთებულ იქნეს მთლიანად ამ კომპონენტშივე არსებული წირით. 8.18 სურათზე ნაჩვენებია შესაბამისად: (ა) ერთი, (ბ) ორი და (გ) სამი ბმული კომპონენტისგან შემდგარი ობიექტები;
- **ობიექტში “ნახვრეტების” რაოდენობა** – ახასიათებს ბმული კომპონენტების იმ რაოდენობას, რომლებიც ობიექტის შიგნითაა, მაგრამ მას არ მიეკუთვნება;
- **ეილერის რიცხვი** – გამოითვლება როგორც სხვაობა ობიექტის ბმული კომპონენტების რაოდენობასა და მასში ნახვრეტების რაოდენობას შორის.

ვთქვათ, მოცემული გვაქვს ერთი ბმული კომპონენტისგან შედგენილი ობიექტი, რომელიც შეიცავს 3 ნახვრეტს (სურ. 8.18,ა). ცხადია, ამ ფიგურის ტოპოლოგიური გარდაქმნის შედეგად მასში ნახვრეტების რიცხვი არ შეიცვლება. ცნობილია ის ფუნდამენტური თანაფარდობა, რომელიც არსებობს ბმული კომპონენტების C რაოდენობასა და ნახვრეტების H რაოდენობას შორის. მას **ეილერის რიცხვი** ეწოდება და აქვს ასეთი სახე

$$E = C - H.$$

რადგანაც C და H წარმოადგენენ ტოპოლოგიურ თვისებებს, ამიტომ E -ც ტოპოლოგიური თვისებაა. ჩვენ მიერ წარმოდგენილი მაგალითის შემთხვევაში $C = 1$ და $H = 3$. აქედან გამომდინარე ეილერის რიცხვი $E = -2$.



ბ
სურ. 8.18

არსებობენ ასევე ტოპოლოგიური თვისებები, რომლებიც დაკავშირებული არიან დისკრეტიზებული ობიექტების ამოზნექილობასთან.

ზემოაღნიშნული ტოპოლოგიური თვისებები შეიძლება გამოყენებულ იქნეს როგორც სიმბოლოები გამოსახულებაზე არსებული ობიექტის ამოცნობისას. მაგრამ გარდა ამისა, ტოპოლოგიური თვისებები შეიძლება გამოყენებულ იქნეს ასევე სხვადასხვა გეომეტრიული სიდიდეების, მაგალითად, დისკრეტული გამოსახულების ქვესიმრავლის პერიმეტრისა და ფარ-

თობის გამოსათვლელად. კერძოდ, ქვესიმრავლის ფართობი ამ შემთხვევაში ზუსტად ტოლია მისი ელემენტების (პიქსელების) რიცხვისა, ხოლო პერიმეტრის გამოთვლისათვის საჭიროა მხედველობაში მივიღოთ ის ფაქტი, რომ ის დაახლოებით პროპორციულია ქვესიმრავლის გარშემოწერილობის ფართისა. ანალოგიურად იზომება წირებისგან შედგენილი ქვესიმრავლის რკალის სიგრძე. ზოგადად, უწყვეტი გამოსახულების რკალის პერიმეტრი და სიგრძე შეუძლებელია ზუსტად გაზომილ იქნეს, მაგრამ შესაძლებელია მათი აპროქსიმაცია მასში ჩაწერილი მრავალკუთხედით. ცხადია, ამ დროს ასეთი ხერხით გამოთვლილი რკალის სიგრძე იზრდება ან უცვლელი რჩება (ყოველ შემთხვევაში არ მცირდება) ამ მრავალკუთხედის გვერდების სიგრძის შემცირებასთან ერთად. ასეთ შემთხვევაში ამბობენ რკალის “ ϵ -სიგრძის” შესახებ, სადაც ϵ წარმოადგენს რკალის აპროქსიმაციისთვის მასში ჩახაზული მრავალკუთხედის უმცირესი გვერდის სიგრძეს. ემპირიულად დადგენილ იქნა, რომ რაიმე რკალის “ ϵ სიგრძე” ხშირად შეიძლება აპროქსიმირებულ იქნეს ϵ -ის მაჩვენებლიანი ფუნქციით $L_\epsilon = K\epsilon^{-d}$, სადაც $d \geq 0$; წრფისთვის $d = 0$, ხოლო ძალიან დაკლაკნილი მრუდებისთვის კი $d \cong 0.25$; როგორც ვხედავთ, აქ d შეიძლება გამოყენებულ იქნეს როგორც სწორხაზოვნების საზომი. ამავე მიზნისთვის შეიძლება გამოყენებულ იქნეს რკალის სიგრძის ფართობა მის ქორდასთან.

გამოსახულებაზე რაიმე ქვესიმრავლის სიმაღლე ან სიგანე მარტივად შეიძლება განვსაზღვროთ, თუ გამოვთვლით მანძილს არის ყველაზე მაღლა და ყველაზე დაბლა მყოფ ჰორიზონტალურ წირებს შორის ან ყველაზე უფრო მარჯვენა და ყველაზე უფრო მარცხენა ვერტიკალურ წირებს შორის. მსგავსად გამოითვლება ქვესიმრავლის წაგრძელებულობა მოცემული მიმართულებით.

ნიშან-თვისებათა შერჩევისას მარტივი მიდგომა მდგომარეობს იმაში, რომ ნიშან-თვისებებად შეიძლება ავირჩიოთ ყველა პიქსელისგან შემდგარი მოუწესრიგებელი კრებული, მაგრამ ასეთი მიდგომა ძალიან რთული გამოთვლითი სირთულეების გამო არაპრაქტიკულია.

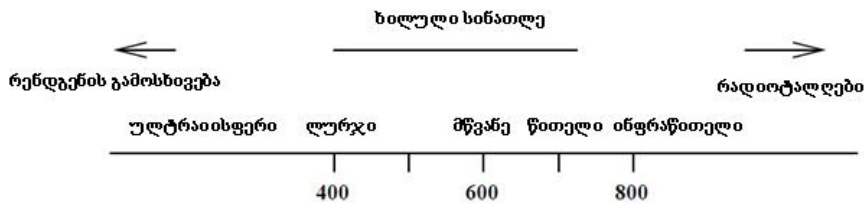
IX თავი. ფერი და განათებულობა

ფერის აღქმას დიდი მნიშვნელობა აქვს ადამიანისთვის, რადგანაც რეალური ბუნება ფერთა მრავალსახეობით ხასიათდება. თავის მხრივ ფერის აღქმა განისაზღვრება როგორც სინათლის ფიზიკური ბუნებით, ასევე სისტემის – თვალი-ტვინი თავისებურებებით. ადამიანი ფერის შესახებ ინფორმაციის საფუძველზე ახდენს გარესამყაროს ობიექტების ამოცნობას, დროის აღქმას და ა.შ. ტექნოლოგიურმა პროგრესმა ხელი შეუწყო ფერადი გამოსახულებების მიღებისა და დამუშავების მოწყობილობების შექმნასა და განვითარებას. ამან კი თავის მხრივ განაპირობა ფერადი გამოსახულებების მანქანური დამუშავების გამოყენებითი სისტემების შექმნა. ფერადი გამოსახულების შემთხვევაში გამოსახულების ყოველ პიქსელში გვაქვს გაზომვის რამდენიმე მნიშვნელობა, რაც ზოგიერთ ამოცანაში ობიექტების კლასიფიკაციის შესაძლებლობასაც კი იძლევა გამოსახულების რთული დამუშავების გარეშე. ამ თავში ჩვენ შევისწავლით ფერადი გამოსახულებების დამუშავების მხოლოდ ძირითად საკითხებს, დაკავშირებულს, მაგალითად, ობიექტების განათებულობასთან, ფერით მონაცემებთან, რასაც დიდი მნიშვნელობა აქვს კომპიუტერული ხედვის ალგორითმების გამოყენებით სცენების ინტერპრეტაციისთვის.

9.1. სინათლისა და ფერის ფიზიკური თვისებები

როგორც ვიცით, ელექტრომაგნიტური ველის ხილული სპექტრი მოიცავს დიაპაზონს დაახლოებით (400 – 700)ნმ. ფარგლებში (სურ. 9.1). ეს ნიშნავს, რომ λ ტალღის სიგრძის მქონე ელექტრომაგნიტური გამოსხივება აღნიშნული დიაპაზონიდან

ადამიანის მხედველობით ორგანოზე ზემოქმედების შედეგად იწვევს მასზე ფერის შეგრძნებას.



სურ. 9.1

ცნობილია, რომ სინათლე ვაკუუმში ვრცელდება (3×10^8) მ/წმ სიჩქარით. ეს გვაძლევს შესაძლებლობას გამოვთვალოთ, მაგალითად, ლურჯი ფერის სინათლის სიხშირე, რომელიც (0.75×10^{15}) ჰც-ის ტოლია. რენტგენის გამოსხივების სიხშირე ათასჯერ მეტია, ხოლო რადიოტალღების კი – მილიარდჯერ ნაკლები.

სინათლის სხივი ობიექტის ზედაპირზე დაცემის შემდეგ აირეკლება და ეცემა კამერის სინათლის მგრძნობიარე გადამწოდს ან ადამიანის თვალის რეცეპტორებს. ამ დროს ფერის მიმართ წარმოშობილი მგრძნობიარობა დამოკიდებულია შემდეგ სამ ფაქტორზე:

- **ენერგეტიკული სპექტრი** – აღწერს ობიექტის ზედაპირზე დაცემული სინათლის ენერგიის განაწილებას ტალღის სიგრძეების მიხედვით;
- **ობიექტის ზედაპირის სპექტრალური არეკვლის უნარი** – განსაზღვრავს დაცემული სინათლის სპექტრის გარდაქმნას არეკვლილ სპექტრში;

- **გადამწოდის სპექტრალური მგრძობიარობა**, რომელზედაც ეცემა ობიექტის ზედაპირიდან არეკვლილი სინათლე.

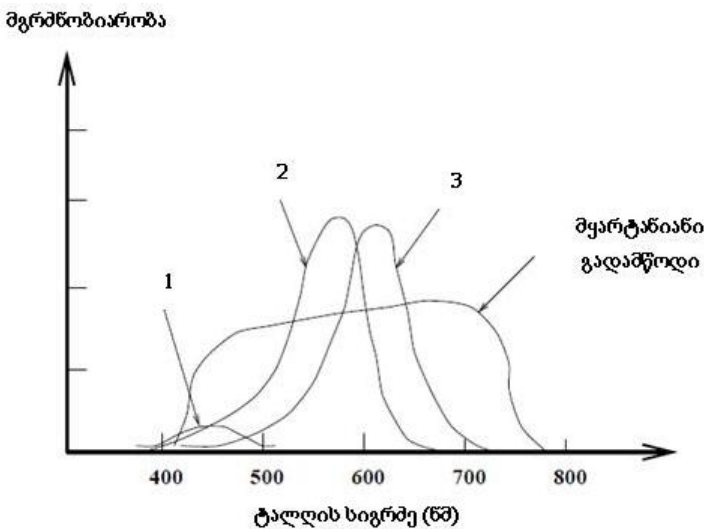
განსაზღვრება 9.1. თეთრი ფერის სინათლე წარმოადგენს ხილულ სპექტრში შემავალი ყველა სიგრძის ტალღის სინათლის კომპონენტების ნარევს, რომელთაც თითქმის ერთნაირი ენერგიები გააჩნიათ.

გარდა ზემოაღნიშნული ფაქტორებისა, რომლებიც განსაზღვრავს ადამიანის მიერ ფერის მგრძობიარობას, არსებობს ასევე დამატებითი ფაქტორებიც. ისინი დაკავშირებულია როგორც ფიზიკურ, ასევე აღქმის თავისებურებებთან. სხვადასხვა მასალისგან დამზადებული ობიექტების ზედაპირები სხვადასხვანაირად აირეკლავს მათზე დაცემულ სინათლის სხივებს. მაგალითად, მქრქალ ზედაპირზე დაცემული სინათლის სხივები აირეკლება ყველა მიმართულებით თანაბრად. ასევე დაცემული სხივების ენერგია (ინტენსივობა) დამოკიდებულია მნათი ენერგიის წყაროსა და ზედაპირის ელემენტებს შორის მანძილზე. რაც მეტია ეს მანძილი, მით ნაკლებია მათზე დაცემული სინათლის სხივების ენერგია. აქედან გამომდინარეობს, რომ ერთი და იმავე მასალისგან დამზადებული ზედაპირების გამოსახულებაზე პიქსელებში ინტენსივობის მნიშვნელობები შეიძლება განსხვავებული იყოს გამოსხივების წყაროსა და ობიექტების ზედაპირებს შორის დაცემული სხივების გასწვრივ განსხვავებული მანძილების გამო. ობიექტის ზედაპირიდან არეკვლილი და გადამწოდზე დაცემული სინათლის სხივის ენერგიის განსაზღვრისათვის, გარდა მანძილისა, კიდევ უფრო დიდი მნიშვნელობა აქვს ზედაპირის ორიენტაციას სინათლის წყაროს მიმართ, ანუ θ კუთხეს დაცემულ სხივსა და ზე-

დაპირის მიმართ დაცემის წერტილში აღმართულ ნორმალს შორის.

9.2. ფერების RGB სისტემა

მხედველობის ანალიზატორების ფოტორეცეპტორები მგრძობიარეა ტალღის სიგრძის მხოლოდ გარკვეული დიაპაზონის მიმართ. ამასთან, თვით ამ დიაპაზონში მათი მგრძობიარობა არათანაბარია. ეს ნაჩვენებია 9.2 სურათზე, სადაც მოცემულია ადამიანის თვალის ბადურაზე არსებული სამი განსხვავებული ტიპის კოლების მგრძობიარობის მრუდები (1, 2, 3).



სურ. 9.2

კოლებების თითოეული ეს ტიპი შეიცავს სხვადასხვა ქიმიურ პიგმენტს, რომელთაც გააჩნია განსხვავებული მგრძობიარობა სხვადასხვა ტალღის სიგრძის მიმართ. ეს კარგად ჩანს 9.2 სურ-

რათზე წარმოდგენილი მრუდებიდან. კერძოდ, მრუდი 1 შეესაბამება იმ კოლბებს, რომლებიც მგრძობიარეა ლურჯი ფერის მიმართ ტალღის სიგრძეებით (400 – 500)ნმ. მრუდი 2 შეესაბამება იმ კოლბებს, რომლებიც ძალიან მგრძობიარეა მწვანე ფერის მიმართ და უფრო ნაკლებად ლურჯი და წითელი ფერების მიმართ, ხოლო მრუდი 3 კი იმ კოლბებს, რომლებიც ძალიან მგრძობიარეა წითელი ფერის მიმართ და უფრო ნაკლებად მწვანე ფერის მიმართ. სურათზე ასევე მოცემულია **მყარტანიანი გადამწოდის (მტგ)** მგრძობიარობის მრუდი, რომელიც, როგორც ვხედავთ, მოიცავს ადამიანის სამივე ტიპის კოლბის მგრძობიარობას.

ხილული სპექტრის ზემოქმედების შედეგად ადამიანის შეგრძნების ფორმირებისას ტვინი მხედველობაში იღებს რეაქციას რამდენიმე კოლბისგან შედგენილი არიდან. ცნობილია, რომ ამისთვის საკმარისია სულ სამი ტიპის კოლბა. ზოგიერთი ცხოველის თვალის ბადურა შეიცავს მხოლოდ ერთ ან ორი ტიპის რეცეპტორს. აქედან გამომდინარე, ცხადია, მათ მიერ სამყაროში არსებული ფერების აღქმა ადამიანთან შედარებით უფრო შეზღუდულია – ღარიბია.

ზემოთქმული დაედო საფუძვლად ტექნიკური ხედვის სისტემებში ფერების წარმოდგენას – შეიქმნა ფერთა სამფერიანი RGB (წითელი-მწვანე-ლურჯი) სისტემა. გამოსახულების პიქსელში ფერის აღწერისათვის გამოიყენება 3 ბაიტი, რაც გვაძლევს $(2^8)^3 \approx 16$ მლნ. სხვადასხვა ფერის კოდს. მაგრამ ეს არ ნიშნავს, რომ არსებობს 16 მლნ. ფერი, რადგანაც ადამიანს რეალურად არ შეუძლია ამ რაოდენობის ფერის ერთმანეთისგან გარჩევა. კომპიუტერს შეუძლია განასხვავოს ერთმანეთისგან ნებისმიერი ორი კოდი, მაგრამ ისინი შეიძლება არ წარმოადგენდეს იმ ფერებს, რომელთა ერთმანეთისგან გარჩევა ადამიანებს შეუძლიათ.

ფერთა RGB სისტემაში თითოეული ფერისთვის გამოყოფილია თითო ბაიტი. რაც შეეხება ამ კომპონენტების მექსიერებაში შენახვის თანამიმდევრობას, თეორიული განხილვისათვის, ცხადია, ამას მნიშვნელობა არ აქვს. მაგრამ იგი მნიშვნელოვანია პროგრამირებისას. რაც შეეხება ფერად დისპლეებს (მონიტორებს), რომელთა ფერების გარჩევისუნარიანობა ადამიანის მხედველობის სისტემის შესაბამისია, მათ **სრულფეროვან** (true color) ან სხვანაირად, **ფერთა რეალისტურ აღმწარმოებელ დისპლეებს** უწოდებენ. ამასთან ფერთა წარმოდგენისთვის საჭიროა არაუმცირეს 16 ბიტისა, ხოლო 15 ბიტიანი RGB კოდირებისას თითოეული R , G , B ფერის წარმოდგენისთვის გამოიყენება 5 ბიტი. ქვემოთ ნაჩვენებია ფერების წარმოდგენა სამ სხვადასხვა RGB, CMY (ფერთა სუბტრაქტული სისტემა, ცისფერი-მეწამული-ყვითელი, Subtractive Color System), და HSI (ტონი-გაჯერებულობა-ინტენსივობა, Hue-Saturation-Intensity) სამფერიან სისტემებში (ცხ. 9.1).

აღსანიშნავია, რომ ხშირად ახდენენ ხოლმე ინტენსივობის მნიშვნელობების მასშტაბირებას, იმისთვის, რომ ისინი მოხვდეს არა $[0;255]$, არამედ $[0;1]$ დიაპაზონში, რაც ამარტივებს ფერადი გამოსახულებების დამუშავების ალგორითმებს. რაც შეეხება HSI სისტემას, მისი კომპონენტების მიღება შესაძლებელია RGB სისტემიდან სპეციალური ალგორითმის გამოყენებით, რომელიც კომპონენტებისთვის გვაძლევს შემდეგ დიაპაზონებს:

$$H \in [0.0; 2\pi], S \in [0.0; 1.0] \text{ და } I \in [0; 255].$$

RGB სისტემას **ფერთა ადიტიურ სისტემას** უწოდებენ, რადგანაც ფერები აქ წარმოიდგინება ძირითადი ფერების შავ ფერთან $(0,0,0)$ აჯამვის საშუალებით. ასეთი სისტემა კარგად

უდგება RGB დისპლეებს, რომლებშიც სინათლის გამოსხივება ხდება ლუმინოფორის 3 ტიპის მიდამოდან. ლუმინოფორის სამი მეზობელი ელემენტი ქმნის ერთ პიქსელს. ამ ელემენტებს ეცემა ელექტრონების კონები ინტენსივობით c_1, c_2, c_3 . ადამიანის თვალი კი ერთმანეთთან ახლოს განლაგებული ელემენტებიდან წამოსულ გამოსხივებას აღიქვამს როგორც ერთ (c_1, c_2, c_3) ფერს. ეს ნიშნავს, რომ სამი განსხვავებული სიგრძის მქონე სინათლის ტალღა, წამოსული ელექტრონულ-სხივური მილაკის ეკრანის მცირე მიდამოდან, ერთმანეთს ერევა და ადამიანის მიერ აღიქმება როგორც ერთი.

ცხ. 9.1

	RGB	CMY	HSI
წითელი	(255, 0, 0)	(0, 255, 255)	(0.0, 1.0, 255)
ყვითელი	(255, 255, 0)	(0, 0, 255)	(1.05, 1.0, 255)
	(100, 100, 50)	(155, 155, 205)	1.05, 0.5, 100)
მწვანე	(0, 255, 0)	(255, 0, 255)	(2.09, 1.0, 255)
ლურჯი	(0, 0, 255)	(255, 255, 0)	(4.19, 1.0, 255)
თეთრი	(255, 255, 255)	(0, 0, 0)	(-1.0, 0.0, 255)
რუხი	(192, 192, 192)	(63, 63, 63)	(-1.0, 0.0, 192)
	(127, 127, 127)	(128, 128, 128)	(-1.0, 0.0, 127)
	(63, 63, 63)	(192, 192, 192)	(-1.0, 0.0, 63)
შავი	(0, 0, 0)	(255, 255, 255)	(-1.0, 0.0, 0)

სშირად, გამომდინარე სხვადასხვა მიზნებიდან, მაგალითად, ფერების მონაცემების ინტერპრეტაციისთვის ადამიანის ან კომპიუტერული პროგრამების მიერ, ან სხვა ფერთა სისტემაში გარდაქმნისთვის ახდენენ გამოსახულების ფერების მონაცემების ნორმირებას. ვთქვათ გვაქვს სცენა, რომლის გამოსახულება არაერთგვაროვნადაა განათებული. მაგალითად, ობიექტის ზედაპირის წერტილები განლაგებულია სინათლის წყაროებიდან სხვადასხვა მანძილზე, ან იმყოფება ზოგიერთი წყაროს ჩრდილში. ასეთ შემთხვევაში რომელიმე ფერის მოძიების ალგორითმმა შეიძლება ცუდი შედეგები მოგვცეს, თუ წინასწარ არ მოხდება პიქსელთა ინტენსივობის მნიშვნელობების ნორმირება. ამისთვის შეიძლება გამოვიყენოთ, მაგალითად, ასეთი ტიპის ნორმირება:

$$\text{ინტენსივობა } I = (R + G + B)/3;$$

$$\text{ნორმირებული წითელი } r = R/(R + G + B);$$

$$\text{ნორმირებული მწვანე } g = G/(R + G + B);$$

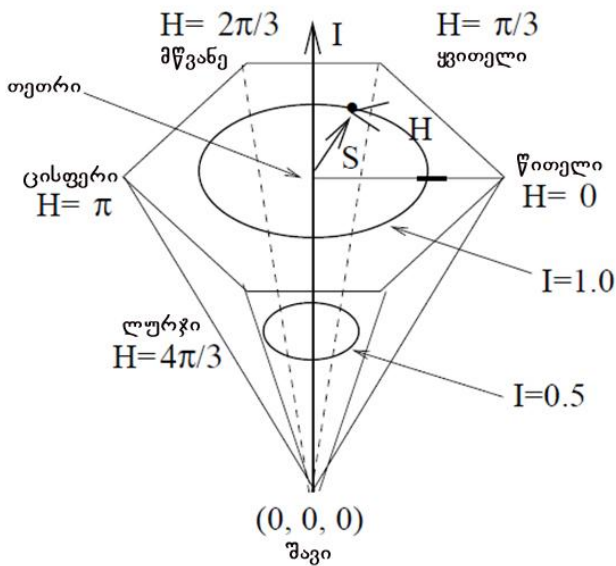
$$\text{ნორმირებული ლურჯი } b = B/(R + G + B).$$

ასეთი ნორმირების შემდეგ პიქსელის ნორმირებულ ფერთა მდგენელების ჯამი 1-ის ტოლი იქნება. არსებობს ნორმირების სხვა ხერხებიც.

ფერთა CMY სისტემას ეწოდება **სუბტრაქტული** იმის გამო, რომ კოდების მნიშვნელობები ახასიათებს შთანთქმული გამოსხივების ინტენსივობას. ამ სისტემაში სამფერიანი კოდების მაგალითებს წარმოადგენენ: თეთრი – (0,0,0), რადგანაც თეთრი ფერის სინათლე არ უნდა შთაინთქას, შავი – (255,255,255), რადგანაც თეთრი ფერის სინათლის ყველა კომპონენტი უნდა იყოს შთანთქმული და ყვითელი – (0,0,255), რადგანაც დაცემული თეთრი ფერის სინათლის ლურჯი

კომპონენტი უნდა იყოს შთანთქმული, ხოლო წითელი და მწვანე ფერები უნდა აირეკლოს და წარმოშვას ყვითელი ფერის აღქმის შეგრძნება.

რაც შეეხება ფერთა HSI სისტემას, მასში ფერადი პიქსელის კოდი ასეთია: პირველი მნიშვნელობა წარმოადგენს სინათლის ინტენსივობის (I) საერთო მნიშვნელობას, ხოლო დანარჩენი ორი კი ფერადობის ორ სიდიდეს – ტონი (H) და გაჯერებულობა (S). 9.3 სურათზე ნახვენებია HSI სისტემის ფერთა სივრცის სტრუქტურა.



სურ. 9.3

ვერტიკალური ღერძი შეესაბამება ინტენსივობას (I), ტონი (H) განისაზღვრება როგორც კუთხე 0-დან 2π-მდე წითელი ფერის

დერძის მიმართ. სუფთა წითელ ფერს შეესაბამება 0-ის ტოლი კუთხე, სუფთა მწვანე ფერს – $2\pi/3$, ხოლო სუფთა ლურჯს – $4\pi/3$. ფერების სივრცეში წერტილის წარმოსადგენად მესამე კოორდინატს წარმოადგენს გაჯერებულობა (S). ის ახასიათებს ფერის (ფერის ტონის) სისუფთავეს და იცვლება 0-დან 1-მდე. თუ $S = 1$, მაშინ გვაქვს იდეალურად სუფთა (გაჯერებული) ფერი, ხოლო თუ $S = 0$, მაშინ გვაქვს არაგაჯერებული ფერი ან, სხვა სიტყვებით, რუხი ფერის რაღაც ელფერი.

HSI სისტემას ასევე HSV სისტემასაც ეძახიან. ამ სისტემაში ინტენსივობის ნაცვლად გამოიყენება ტერმინი **სინათლოვნება** (color value). HSI სისტემას ხშირად იყენებენ გრაფიკის დიზაინერები, რადგანაც იგი საშუალებას იძლევა მოვახდინოთ სიკაშკაშისა და ტონის უშუალო კონტროლი.

9.3. ფერადი ჰისტოგრამები

ფერადი გამოსახულების ჰისტოგრამა შეიძლება გამოყენებულ იქნეს მონაცემთა ბაზებში ძიებისთვის, ან ობიექტების ამოცნობისთვის. ჰისტოგრამის ასაგებად საჭიროა გამოსახულებაზე დათვლილ იქნეს გარკვეული ტიპის პიქსელების რაოდენობა. ამისთვის გამოსახულებაზე გავლისას თითოეული პიქსელი იკითხება ერთჯერადად და ჰისტოგრამის შესაბამის ჯიბეში ხდება მათი რაოდენობის თანამიმდევრული აჯამება.

ფერადი გამოსახულების ჰისტოგრამის აგების მარტივი მიახლოებითი მეთოდი გულისხმობს ფერის (R, G, B) სამბაიტის მნიშვნელობაში ყოველი ფერითი მდგენელისგან ორი მაღალი თანრიგის ბიტის გაერთიანებას. ასეთნაირად მიღებული 6 ბიტიაანი კოდის საფუძველზე ხდება ჰისტოგრამის აგება, რომელიც შეიცავს $2^6 = 64$ თანრიგს (ჯიბეს). ასევე შეიძლება აიგოს ჰისტოგრამები თითოეული ფერისთვის ცალ-ცალკე და

შემდეგ მოხდეს მათი გაერთიანება. მაგალითად, ცალკეული 16-თანრიგიანი RGB ჰისტოგრამების გაერთიანებით შეიძლება 48-თანრიგიანი ჰისტოგრამის მიღება.

ვთქვათ, გვაქვს რაიმე I გამოსახულების $h(I)$ ჰისტოგრამა და ასევე მოდელური (ეტილონიური) M გამოსახულების $h(M)$ ჰისტოგრამა. ამ ორი **ჰისტოგრამის გადაკვეთა** ეწოდება ამ ჰისტოგრამების შესაბამისი ჯიბეების (ფერების მიხედვით ქვესიმრავლეების) ყველა K წყვილში არსებული მინიმალური მნიშვნელობების ჯამს. ფორმალურად ეს ასე ჩაიწერება

$$\text{გადაკვეთა } (h(I), h(M)) = \sum_{j=1}^K \min\{h(I)[j], h(M)[j]\}.$$

თუ მოვასდგნთ ამ სიდიდის ნორმირებას, რისთვისაც მისი მნიშვნელობა უნდა გაიყოს მოდელურ გამოსახულებაში არსებული პიქსელების რაოდენობაზე, მივიღებთ სიდიდეს, რომელსაც **მსგავსებას** (match value) უწოდებენ. ეს მახასიათებელი გამოიყენება მოცემული I და მოდელური M გამოსახულებების ფერთი შემაღვენლობის მსგავსების რაოდენობრივ ზომად და იგი ასეთნაირად გამოითვლება

$$\text{მსგავსება } (h(I), h(M)) = \frac{\sum_{j=1}^K \min\{h(I)[j], h(M)[j]\}}{\sum_{j=1}^K h(M)[j]}.$$

უნდა აღინიშნოს, რომ თუ მოდელურ M გამოსახულებაში არ არის მოცემული I გამოსახულების ფონური პიქსელები, ამით მსგავსების მნიშვნელობა არ მცირდება.

გარდა აღნიშნული მეთოდისა, არსებობს სხვა მეთოდებიც. მაგალითად, ჰისტოგრამების ნორმირება მათი სიხშირეების ჰისტოგრამებად გარდაქმნის გზით, რისთვისაც ყოველ ჯიბეში არსებული პიქსელების რაოდენობა (პიქსელების მრიც-

ხველის ჩვენება) იყოფა პიქსელების საერთო რაოდენობაზე ამის შემდეგ შესაძლებელია ეს ორი გამოსახულება ერთმანეთს შედარდეს მათი მსგავსების დასადგენად ევკლიდურ მეტრიკაზე დამყარებული კრიტერიუმის მიხედვით. არსებობს ასევე გამოსახულებების შედარების ალტერნატიული მეთოდიც, რომელშიც სისწირების ჰისტოგრამების ნაცვლად გამოიყენება განაწილების ფუნქციების შეპირისპირება.

9.4. ფერადი გამოსახულებების სეგმენტაცია

განვიხილოთ ფერად გამოსახულებაზე ადამიანის პირისახის გამოსახულების სახის ძიების ამოცანა ზოგადად. გამოსახულების მისაღებად გამოყენებულია კომპიუტერთან მიერთებული მცირეგაბარიტიანი კამერა. მისთვის გამოვიყენოთ გამოსახულებაზე იმ ძირითადი არის ძიების ალგორითმი, რომელიც შეესაბამება ადამიანის პირისახის გამოსახულების სახეს. დასაწყისში, იმისთვის, რომ სწავლების პროცესის ეტაპზე მოვასდინოთ ადამიანის პირისახის სახის შესაბამისი პიქსელების თვისებების განსაზღვრა, ხდება სხვადასხვა ადამიანის პირისახის გამოსახულებიდან მიღებული სახეების გაანალიზება. ასევე, ახალ გამოსახულებაზე ადამიანის პირისახის პიქსელების მოძებნა ხდება ფერთი (r, g) მნიშვნელობების საშუალებით, რომლებიც უნდა შეთანხმდეს ეტალონურ მონაცემებთან.

ოპერაციები, რომლებიც გამოყენებულია გამოსახულებაზე ადამიანის სახის შესაბამისი არის ძიებისთვის, დაყოფილია სამ ძირითად ეტაპად. პირველ ეტაპზე შესასვლელ მონაცემებს წარმოადგენს მარკირებული გამოსახულება, სადაც პიქსელებს ჭდეები ენიჭება კლასიფიკაციის ალგორითმის საშუალებით, რომელიც ამისთვის ითვალისწინებს სწავლების პროცესის დროს დამახსოვრებულ ეტალონურ გამოსახულებებს. მარკირე-

ბულ გამოსახულებაზე ფონისა და პირისახეების სახეთა პიქსელების უმეტესობა კორექტულად არის მარკირებული, მაგრამ ასევე შეიძლება აღმოჩნდეს პიქსელები, რომელთაც ჭდეები არაკორექტულად აქვს მინიჭებული. მარკირებული გამოსახულების ბმული კომპონენტების დამუშავებისას ასეთი კომპონენტები, გამომდინარე მათი ზომებიდან და სახის მიმართ მათი მდებარეობიდან, ან ერთიანდება, ანდა მოცილებულ უნდა იქნეს გამოსახულებიდან. ამავე დროს ბმული კომპონენტების დამუშავებისას მხედველობაში მიიღება მხოლოდ გამოსახულების წინა პლანზე არსებული პიქსელები.

დამუშავების მეორე ეტაპზე ადამიანის პირისახის სახედ აირჩევა ყველაზე უფრო შესაფერისი კომპონენტი და ხდება იმ კომპონენტების მოშორება, რომლებიც თავისი ზომით ძალიან პატარაა ან დიდია. თვით ზომის შესაფასებლად კი გამოყენებულია დიდი რაოდენობის ტესტური გამოსახულებების ანალიზის შედეგად მიღებული ემპირიული გამოცდილება. მეორე ეტაპის ოპერაციების რეალიზაციის შედეგად გამოსახულებაზე კომპონენტების რაოდენობა მცირდება. მესამე ეტაპზე კი უკვე ხდება ამ დარჩენილი კომპონენტების მოცილება ან გაერთიანება გამოსახულებაზე წინასწარ არჩეულ სახის არესთან. ამისთვის ალგორითმი იყენებს რამდენიმე ევრისტიკულ წესს პირისახის სახის სტრუქტურის მიმართ. კერძოდ, მხედველობაში მიიღება ის, რომ გამოსახულებაზე შეიძლება იმყოფებოდეს მხოლოდ ერთი ადამიანის პირისახე. წარმოდგენილი ალგორითმის პროგრამა 1 წმ.-ში ამუშავებს 30 გამოსახულებას. ამავე დროს დამატებითი ალგორითმების საშუალებით გამოსახულებაზე ხდება თვალებისა და ცხვირის შესაბამისი არეების ადგილმდებარეობის განსაზღვრა. აღნიშნული ალგორითმის ანალიზი გვიჩვენებს, რომ მასში მთავარ მომენტს წარმოადგენს ათასობით ევრის კოდების კლასტერიზაცია, რაც გვაძლევს

გამოსახულების მარკირების საშუალებას, რომელშიც გვექნება მხოლოდ რამდენიმე ერთმანეთისგან განსხვავებული ჭდე. აღსანიშნავია, რომ ამ ალგორითმში სწავლების პროცესში ადამიანთა პირისახეების სახეთა აღმოჩენა ხდება ხელით ანუ ინტერაქტიულად.

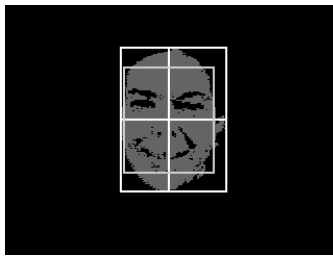
ადამიანის პირისახის სახის აღმოჩენის მაგალითი ნაჩვენებია 9.4 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, (ბ) მარკირებული და (გ) ადამიანის პირისახის სახეზე გამოყოფილი ცალკეული არეების საზღვრები.



ა



ბ



გ

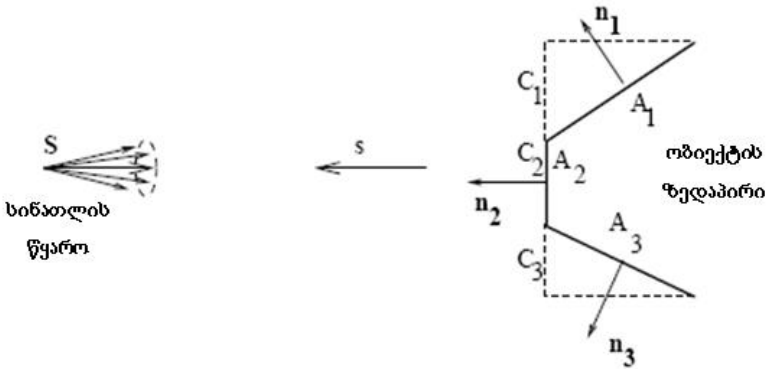
სურ. 9.4

9.5. განათებულობა

სინათლის ტალღის არეკვლის თვალსაზრისით ობიექტების ზედაპირების თვისებები ერთმანეთისგან განსხვავებულია. სხვა სიტყვებით, მათ განსხვავებული სარკული თვისებები გააჩნია. ამასთან, სინათლის დაცემული ტალღა სხეულის ზედაპირიდან აირეკლება გეომეტრიულ ოპტიკაში ცნობილი არეკვლის კანონის შესაბამისად. მაგალითად, მქრქალი ზედაპირებიდან სინათლე თანაბრად აირეკლება ყველა მიმართულებით. ამავე დროს ზედაპირზე დაცემული სინათლის ტალღის ენერგია (ინტენსივობა) პირდაპირ პროპორციულადაა დამოკიდებული მანძილზე სინათლის წყაროდან. ზემოთქმულიდან გამომდინარე, გამოსახულებაზე ინტენსივობის მნიშვნელობების განაწილება პიქსელებში იქნება არათანაბარი გადამწოდსა და ობიექტის ზედაპირის წერტილებს შორის სინათლის სხივების გასწვრივ მანძილთა განსხვავებულობის გამო. ასევე მნიშვნელობა აქვს თვით ზედაპირის ორიენტაციას სინათლის წყაროს მიმართ.

9.5.1. სინათლის ერთეულოვანი წყაროს გამოსხიება

ვთქვათ, რაიმე ობიექტის A ბრტყელ ზედაპირს ეცემა სინათლის ტალღა (ნაკადი) რაღაც საკმაოდ შორს მოთავსებული ერთეულოვანი S წყაროდან (სურ. 9.5). ამის გამო, როგორც ობიექტის ზედაპირზე დაცემული, ასევე განათებული ზედაპირის ყოველი წერტილიდან არეკვლილი სხივები შეიძლება ერთმანეთის პარალელური ჩავთვალოთ და მათ ნაცვლად განვიხილოთ ერთი ერთეულოვანი s ვექტორი. სინათლის ენერგია ზედაპირის ფართის ერთეულზე, ანუ ინტენსივობა i , რომელიც ეცემა ზედაპირულ A_j ელემენტს, პროპორციულია ამ ელემენტის პროექციის ფართისა სიბრტყეზე, რომელიც დაცემული



სურ. 9.5

სინათლის \mathbf{s} სხივის მართობულია. ამ პროექციის ფართის გამოსათვლელად საჭიროა A_j ელემენტის ფართი გამრავლდეს A_j ელემენტიდან ზედაპირის სიბრტყის მიმართ აღმართული ნორმალის ერთეულოვან \mathbf{n} ვექტორსა და \mathbf{s} ვექტორს შორის არსებული θ_j კუთხის კოსინუსზე, ანუ

$$C_j = A_j \cos \theta_j.$$

მაგრამ კუთხის კოსინუსი უდრის სკალარულ ნამრავლს $\mathbf{n} \cdot \mathbf{s}$, სადაც \mathbf{n} ნორმალის ერთეულოვანი ვექტორია ბრტყელი ზედაპირული A_j ელემენტის მიმართ. აქედან გამომდინარე, ზედაპირულ ელემენტზე დაცემული სინათლის ინტენსივობის მნიშვნელობა მიახლოებით შეიძლება ასეთნაირად წარმოვიდგინოთ

$$i \approx \mathbf{n} \cdot \mathbf{s}$$

თავის მხრივ ზედაპირზე დაცემული სინათლის ენერგია დამოკიდებულია თვით წყაროს სიმძლავრეზე, რომელიც სინათლის

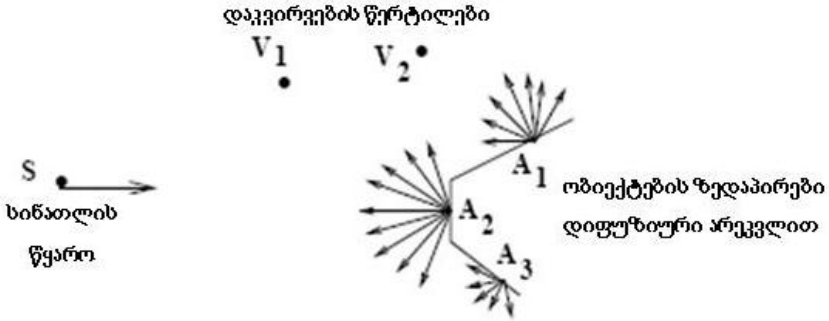
ენერჯიას შეიძლება ასხივებდეს ყველა მიმართულებით, ან რომელიმე კონკრეტული მიმართულებით მცირე კუთხის ფარგლებში. ჩვენ მიერ წარმოდგენილ მოდელში განხილულ იქნა ბრტყელი ზედაპირის განათება. მაგრამ თუ ზედაპირი არ არის ბრტყელი, ანუ მრუდწირულია, მაშინ ის შეიძლება წარმოვიდგინოთ როგორც მცირე ზომის ბრტყელი ელემენტების სიმრავლე. ზედაპირზე დაცემული სინათლის ტალღის იმ ნაწილს, რომელიც აირეკლება ზედაპირული ელემენტის მიერ, ამ ელემენტის ალბედოს უწოდებენ.

განსაზღვრება 9.2. *ზედაპირული ელემენტის ალბედო ტოლია არეკვლილი სინათლის ტალღის ენერჯიის ფარდობისა ზედაპირზე დაცემული სინათლის ტალღის ენერჯიასთან.*

მაგრამ ზოგიერთი ზედაპირისთვის არეკვლილი სინათლის წილი დამოკიდებულია სინათლის წყაროდან მიმართულებასა და ზედაპირის მიმართ აღმართულ ნორმალს შორის კუთხეზე. ამიტომ ალბედო ყოველთვის ვერ გამოდგება როგორც ზედაპირის მახასიათებელი.

9.5.2. დიფუზიური არეკვლა

აქამდე ჩვენ განვიხილეთ ობიექტების ზედაპირების განათებისა და არეკვლის მოდელი, რომელშიც არ იყო გათვალისწინებული დამკვირვებელი და შესაბამისად არ ხდებოდა ზედაპირული ელემენტების შეფასება ამ უკანასკნელთა მხრიდან. ახლა კი განვიხილოთ მოდელი, რომელშიც დამკვირვებელი უკვე გათვალისწინებულია (**V** წერტილები). ასეთი მოდელი, რომელიც განმარტავს თუ რა არის დიფუზიური არეკვლა ანუ არეკვლა ლამბერტის კანონის მიხედვით, მოცემულია 9.6 სურათზე.



სურ. 9.6

ასეთ დროს სინათლე, რომელიც ეცემა ზედაპირულ ელემენტს თანაბრად აირეკლება ყველა მიმართულებით იმ ნახევარსფეროს საზღვრებში, რომლის ცენტრს წარმოადგენს თვით ზედაპირული ელემენტი. აქედან გამომდინარე, დაკვირვების ყოველი წერტილიდან ეს ელემენტი გამოიხსნება როგორც მუდმივი სიკაშკაშის მქონე მიდამო (არე). საერთოდ, სინათლის დიფუზიურ არეკვლას ადგილი აქვს მაშინ, როცა ზედაპირზე არსებული ხორკლიანობის ზომები მნიშვნელოვნად აღემატება დაცემული სინათლის ტალღის სიგრძეს. არეკვლილი სინათლის ინტენსივობა პროპორციულია დაცემული სინათლის ინტენსივობისა, ხოლო თვით პროპორციულობის კოეფიციენტი ზედაპირის ალბედოს ტოლია. მუქი ზედაპირების ალბედო უფრო ნაკლებია, ვიდრე ღიასი. რაც შეეხება ზედაპირული A_j ელემენტიდან დიფუზიურად არეკვლილი სინათლის ინტენსივობის მიახლოებით მნიშვნელობას, ის ტოლია

$$i \approx n_j \cdot s$$

განსაზღვრება 9.3. დიფუზიური არეკვლის თვისების მქონე ზედაპირები სინათლეს აირეკლავს თანაბრად ყველა მიმარ-

თულებით. ამის გამო ასეთი ზედაპირები დაკვირვების ყველა წერტილიდან გამოიყურება როგორც ერთნაირად კაშკაშა.

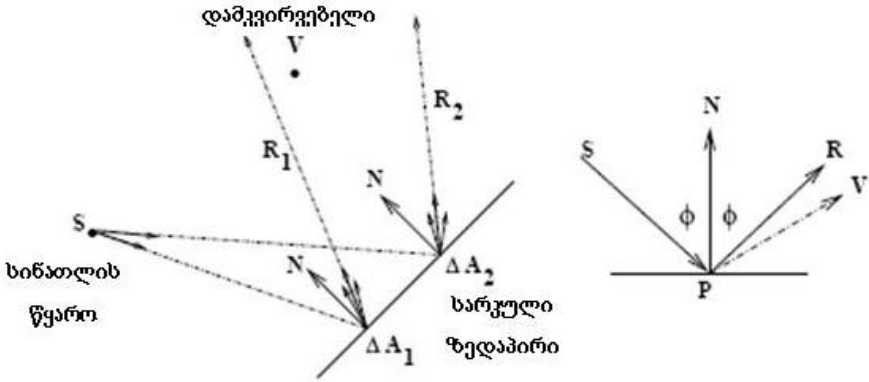
დიფუზიური არეკვლის თავისებურება ისაა, რომ ობიექტის ზედაპირული ელემენტი ერთნაირად კაშკაშა ჩანს ყველა იმ მიმართულებიდან, რომელიც ამ ელემენტის ნახევარსფეროზე გადის. 9.6 სურათზე მოცემულ შემთხვევაში ზედაპირული A_1 და A_2 ელემენტების სიკაშკაშეები ერთნაირი იქნება დაკვირვების V_1 და V_2 წერტილებიდან. თუ სამივე ზედაპირული ელემენტი ერთნაირი მასალისგანაა დამზადებული, მაშინ მათი ალბედოების სიდიდეებიც ერთნაირი იქნება. აქედან გამომდინარე, ელემენტი A_2 იქნება უფრო კაშკაშა, ვიდრე A_1 , ხოლო A_1 იქნება უფრო კაშკაშა, ვიდრე A_3 . განსხვავება სიკაშკაშეებში გამოწვეულია იმით, რომ განსხვავებულია კუთხეები ამ ელემენტებსა და განათებულობის მიმართულებას შორის. ზედაპირული ელემენტი A_3 დაკვირვების V_1 და V_2 წერტილებიდან საერთოდ არ ჩანს. ზოგადად კი, ელემენტი არ ჩანს, თუ

$$\mathbf{n} \cdot \mathbf{v} < 0,$$

სადაც \mathbf{v} ერთეულოვანი ვექტორია დამკვირვებლის ხედვის მიმართულებაზე.

9.5.3. სარკული არეკვლა

იმ შემთხვევაში, როცა სინათლის ტალღა ეცემა სარკული ზედაპირის მსგავს გლუვ ზედაპირს, მაშინ სარკული ზედაპირიდან სინათლის არეკვლა ხდება სინათლის არეკვლის კანონის შესაბამისად, ანუ დაცემისა და არეკვლის ϕ კუთხეები ერთმანეთის ტოლია (სურ. 9.7).



სურ. 9.7

ამასთან სარკულად არეკვლილი სხივი (\mathbf{R}), დაცემის წერტილში ზედაპირის მიმართ აღმართული ნორმალი (\mathbf{N}) და დაცემული სინათლის სხივი (\mathbf{S}) ერთ სიბრტყეში იმყოფება. იმ შემთხვევაში, თუ სარკე იდეალურია, მაშინ დაცემული სხივი (\mathbf{S}) სრულად აირეკლება სარკულად და ასევე ერთნაირი იქნება მათი სპექტრალური შემადგენლობა, ანუ არეკვლილი სხივი (\mathbf{R}) არ არის დამოკიდებული ობიექტის ზედაპირის ფერზე. მაგალითად, თუ თეთრი ფერის სინათლით გავანათებთ წითელი ფერის ობიექტს, მაშინ მასზე დავინახავთ თეთრ ათინათს. რაც შეეხება სარკულად არეკვლილი (\mathbf{R}) სხივის ინტენსივობას, მისი მიახლოებითი მნიშვნელობა განისაზღვრება ასეთნაირად:

$$i \sim (\mathbf{R} \cdot \mathbf{V})^\alpha$$

სადაც \mathbf{R} – სარკულად არეკვლილი სხივის მიმართულებაა, \mathbf{V} – მიმართულება ზედაპირული ელემენტიდან დაკვირვების წერტილისკენ, α – ზედაპირის ბზინვარების (shininess) პარამეტრია,

რომლის მნიშვნელობა ძალიან ბზინვარე ზედაპირებისთვის ტოლია ან მეტია 100-ის. α -ს ზრდასთან ერთად ხდება $\cos^{\alpha}\phi$ -ს შემცირება, ამასთან მით უფრო ძლიერად, რაც უფრო ϕ შორდება 0-ს.

განსაზღვრება 9.4. *სარკული არეკვლა ახასიათებს ზედაპირის სარკულობის თვისებას (სარკულობა).*

ობიექტის ზედაპირიდან სარკულად არეკვლილი სინათლის სხივი ვრცელდება მის ირგვლივ არსებულ ვიწრო კონუსში.

განსაზღვრება 9.5. *კაშკაშა ლაქას ობიექტის ზედაპირზე, რომელიც წარმოიქმნება სინათლის წყაროდან წამოსული სხივის სარკული არეკვლის შედეგად, ათინათი ეწოდება.*

ათინათი დამახასიათებელია მეტალის, შუშის, ცვილის და სხვა მასალისაგან დამზადებული ობიექტების ზედაპირებისთვის.

9.5.4. ინტენსივობაზე მოქმედი ფაქტორები

ცნობილია, რომ ობიექტის ზედაპირზე დაცემული სინათლის ინტენსივობა დამოკიდებულია სინათლის წყაროსა და ზედაპირს შორის არსებულ მანძილზე. კერძოდ, ზედაპირზე დაცემული სინათლის ინტენსივობა უკუპროპორციულია წყარომდე მანძილის კვადრატისა. ეს მოდელი გამოსადეგია აგრეთვე ობიექტიდან არეკვლილი სინათლისთვისაც. კერძოდ, რომელიმე წერტილში მოთავსებული დამკვირვებლისთვის ინტენსივობა აგრეთვე უკუპროპორციული იქნება მანძილისა დამკვირვებელსა და ზედაპირის ელემენტს შორის, ანუ მანძილის ზრდასთან

ერთად მოხდება ზედაპირზე დაცემული სინათლის ინტენსივობის შემცირება.

ზემოთქმულიდან გამომდინარე, ზედაპირების განათების მოდელის განხილვისას მხედველობაში უნდა მივიღოთ როგორც დიფუზიური, ასევე სარკული კომპონენტები. მაგალითად, თუ ჩვენ გავანათებთ ჩვეულებრივი ჯიბის ფარნით წითელ ვაშლს, მივიღებთ წითელ ობიექტს თეთრი ათინათით. წითელი ფერის სინათლე აირეკლება დიფუზიურად, ხოლო ათინათი წარმოიშობა სარკული არეკვლის შედეგად. თუ ჩვენ წითელი ვაშლის შემცველ სცენაზე მხედველობაში მივიღებდით მხოლოდ სარკულ არეკვლას, მაშინ ვაშლის ზედაპირის უმეტესი ნაწილი არ გამოჩნდებოდა საერთოდ. მაგრამ რთული სცენების შემთხვევაში, როცა სცენაზე გვაქვს რამდენიმე სხვადასხვა ზედაპირი და ამასთან სცენა ნათდება სინათლის რამდენიმე წყაროს საშუალებით, ასეთ შემთხვევაში ყოველი ზედაპირიდან არეკვლილი სინათლე რთულ სურათს ქმნის და პრაქტიკულად შეუძლებელი ხდება ზედაპირებიდან არეკვლილი სინათლის ენერჯის მხედველობაში მიღება. ასეთი პროცესის მოდელირებისთვის განათების მოდელს საჭიროა დაემატოს **გაბნეული სინათლის** კომპონენტი, რაც ხშირად გამოიყენება მანქანურ გრაფიკაში ზედაპირების შესაფერადებლად.

განსაზღვრება 9.6. გაბნეული სინათლე წარმოადგენს სინათლეს, რომელსაც სცენის ნებისმიერ არეში აქვს ინტენსივობის მუდმივი მნიშვნელობა. გაბნეული სინათლე მხედველობაში მიიღება სინათლის რამდენიმე წყაროს არსებობისა და სხვადასხვა ზედაპირებიდან მრავალჯერადი კვლავ არეკვლისას.

როგორც ვიცით, რეალურად სინათლის წყაროები ასხივებს ელექტრომაგნიტური გამოსხივების რაღაც გარკვეულ სპექტრში (გარდა ლაზერებისა, რომლებიც გამოსხივების მონოქრომატულ წყაროს წარმოადგენს). შესაბამისად ზედაპირიდან არეკვლილი ან შთანთქმული სინათლეც რაღაც გარკვეული სპექტრისაა. აქედან გამომდინარე, შესაძლებელია შესაბამისი გადაშოდების შემუშავება. მაგალითად, არსებობს მრავალსპექტრალური სკანერები, რომელთაც შეუძლია ერთეულოვანი ზედაპირული ელემენტიდან ინტენსივობის 200 მნიშვნელობის გენერირება. თუმცა ადამიანს სინათლის ხილული სპექტრის აღსაქმელად სულ სამი მნიშვნელობა (მაგალითად, RGB ან HSI) სჭირდება.

9.6. ფერის აღქმა ადამიანის მიერ

ადამიანის მიერ ფერების აღქმის მექანიზმების შესწავლა ძალზე მნიშვნელოვანია. ჯერ ერთი, ისინი შეიძლება გამოყენებულ იქნეს როგორც მოდელი შესაბამისი ტექნიკური სისტემების შესაქმნელად. მეორეც, მანქანური გრაფიკისა და გრაფიკული დისპლეების გამოყენების ძირითადი მიზანი სწორედ ადამიანსა და კომპიუტერს შორის ინტერფეისის ორგანიზაციაა. სინამდვილეში ადამიანების მიერ ფერების აღქმა სუბიექტურია. თითოეული ფერი ადამიანში სხვადასხვა შეგრძნებებს აღძრავს, მაგალითად, შენობების შიგნით ოთახების კედლებს პასტელურ ფერებში ღებავენ; ბევრ ადამიანზე ცისფერი ფერი დამაწყნარებლად მოქმედებს და ა.შ. ამიტომ ფერთა კომბინაციების შერჩევა ძალზე მნიშვნელოვანია ინტერფეისების პროექტირებისას.

ადამიანის თვალის ბადურაზე იმ რეცეპტორების რაოდენობა, რომლებიც წითელი და მწვანე ფერის მიმართ მგრძობი-

არეა, ბევრად სჭარბობს იმ რეცეპტორების რიცხვს, რომლებიც ლურჯი ფერის მიმართ არის მგრძობიარე. ეს განსაკუთრებით იგრძნობა ბადურის ფოვებაზე (ბადურის ცენტრალური უბანი), სადაც ლურჯი ფერის მგრძობიარე რეცეპტორების რიცხვი ძალიან მცირეა. ამის გამო ფერებთან დაკავშირებული ინფორმაციის დამუშავება ძირითადად მიმდინარეობს ბადურასთან დაკავშირებულ ნეირონულ ქსელში.

9.7. მრავალსპექტრალური გამოსახულებები

გადამწოდი, რომლის საშუალებითაც ყოველ პიქსელში სამივე ფერის (RGB) მნიშვნელობების გაზომვა ხდება, მიეკუთვნება **მრავალსპექტრალური გადამწოდების** კატეგორიას. მაგრამ ინტენსივობა შეიძლება გაიზომოს ხილული სპექტრის გარეთაც. მაგალითად, სპექტრის ინფრაწითელ დიაპაზონში. ასეთ დიაპაზონში თანამგზავრიდან გადაღებულ სურათებზე გახურებული ასფალტის გზები კაშკაშა ობიექტებად ჩანს, ხოლო ცივი წყლის მასივები კი მუქად გამოიყურება. ხშირად ერთ პიქსელზე ჩატარებულ რამდენიმე გაზომვას შეუძლია გაამარტივოს ზედაპირების კლასიფიკაციის ამოცანის ამოხსნა. რაც შეეხება მრავალსპექტრალურ მასკანირებულ სისტემას, მისი აპარატურული ნაწილი შეიძლება ძალიან ძვირადღირებული აღმოჩნდეს. მისი დამზადება მოითხოვს მაღალ სიზუსტეს, რადგანაც მან უნდა შეძლოს თითოეული ზედაპირული ელემენტიდან არეკვლილი რამდენიმე სპექტრალური დიაპაზონის გამოსხივების ერთად შეკრება.

X თავი. ტექსტურა

10.1. ტექსტურის არსი

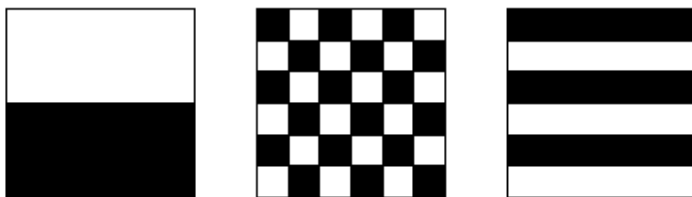
ბუნებრივი სცენების უმეტესობა ხასიათდება იმით, რომ მათი სივრცის უდიდესი ნაწილისთვის არ არის დამახასიათებელი რაიმე მნიშვნელოვანი (არსებითი) დეტალების არსებობა. ასეთ არეებზე სცენა შეიძლება დავახასიათოთ როგორც რაღაც განმეორებადი სტრუქტურის არსებობა, რომელიც მსგავსია ქსოვილის სტრუქტურის, ან აგურით აშენებული კედლის სურათის. ასეთი სტრუქტურის მქონე არეებს **ტექსტურებს** უწოდებენ. ტექსტურის ელემენტები და მათი სივრცითი ორგანიზაციის წესი შეიძლება ნებისმიერად იცვლებოდეს, მაგრამ უცვლელი უნდა იყოს სიკაშკაშის მნიშვნელობების ცვლილების განმეორებადობის მახასიათებლები. უფრო კონკრეტულად ტექსტურა მოიცავს გამოსახულების შემდეგ თვისებებს:

- ტექსტურაში შეიძლება მოიძებნოს ფრაგმენტი, რომლის “სურათი” რეგულარულად მეორდება იმ არის ფარგლებში, რომლის ზომები გაცილებით დიდია ფრაგმენტის ზომებთან შედარებით;
- ეს “სურათი” წარმოიქმნება ფრაგმენტის გარკვეული არაშემთხვევითი წესით განლაგებული ელემენტარული შემადგენელი ნაწილებისგან;
- ფრაგმენტის ელემენტარული ნაწილები წარმოადგენს ერთგვაროვან ერთეულებს, რომლებსაც მთელ ტექსტურულ არეში დაახლოებით ერთნაირი ფორმა აქვს.

გამოსახულებაზე ტექსტურის არსებობის დროს ხშირად საჭირო ხდება ხოლმე ასეთი არის ანუ ტექსტურის როგორც საზღვრების დადგენა, ასევე მისი უმცირესი სტრუქტურული ერთეულის, მარცვლის (ტექსელის) ზომების განსაზღვრა. ზემოთქმულიდან გამომდინარე, ტექსტურა შეიძლება განხილულ იქნეს როგორც ნიშან-თვისება და აქედან გამომდინარე აუცილებელი ხდება მისი რაოდენობრივი აღწერა.

ტექსტურა სხვა ნიშან-თვისებებთან ერთად წარმოადგენს კიდევ ერთ მახასიათებელ ნიშან-თვისებას, რომელიც გამოიყენება გამოსახულების არსობრივ არეებად სეგმენტაციისთვის მათი შემდგომი კლასიფიკაციის მიზნით.

ტექსტურა გამოსახულებაზე აღწერს ფერების ან ინტენსივობების მნიშვნელობების სივრცით განაწილებას. ვთქვათ რომელიმე არის ჰისტოგრამაზე 50% უჭირავს თეთრ პიქსელებს, ხოლო 50% – შავს. ასეთი შავი და თეთრი ფერების ერთნაირი განაწილების მქონე, მაგრამ ერთმანეთისგან განსხვავებული სამი ტექსტურა მოცემულია 10.1 სურათზე.

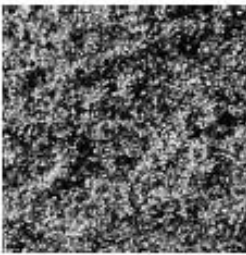


სურ. 10.1

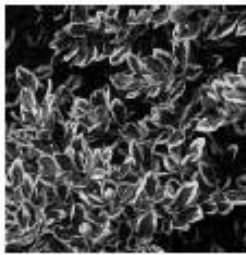
წარმოდგენილი გამოსახულებები შეიძლება განვიხილოთ როგორც ერთმანეთისგან განსხვავებული სამი ტექსტურის მაგალითი.

ტექსტურა შეიძლება დაიყოს ბუნებრივ და ხელოვნურ ტექსტურებად. 10.1 სურათზე წარმოდგენილი ხელოვნური ტექს-

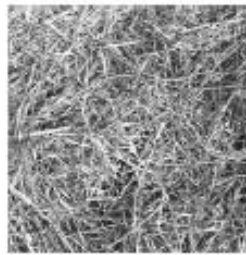
ტურა შედგება შავი და თეთრი ფერის არეებისგან. მაგალითად, ჭადრაკის ტიპის ტექსტურაზე არეები წარმოადგენს ფერების მონაცვლეობით განლაგებულ პატარა კვადრატებს ორგანოზომილებიან ბადეზე. ტექსტურის მაგალითს წარმოადგენს ნეიტრალურ ფონზე განლაგებული გრაფიკული ნიშნებისგან შემდგარი სტრუქტურა. ასეთი ნიშნების მაგალითებია წრფივი მონაკვეთები, წერტილები, ვარსკლავები ან ასოები და ციფრები. ბუნებრივ ტექსტურებს წარმოადგენს ბუნებრივი სცენების გამოსახულებების ობიექტები, რომლებიც ქმნის პერიოდულ სტრუქტურებს, რომლის მაგალითებია: ქვიშა, ქვები, ბალახი, მიწაზე ჩამოვარდნილი ფოთლები, სახლების სახურავები და ა.შ. (სურ. 10.2)



ფოთლები



ფოთლები



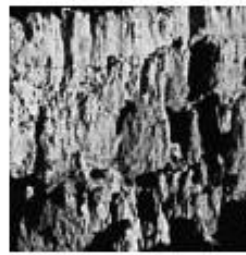
ბალახი



აგურები



აგურები
სურ. 10.2



ქვა

ტექსტურული ანალიზის ერთ-ერთ ამოცანას წარმოადგენს ტექსტურის ცნების ზუსტი განსაზღვრა. ცნობილია ორი ძირითადი მიდგომა:

- **სტრუქტურული:** ტექსტურა წარმოადგენს გარკვეული რეგულარობით (არაშემთხვევითად) ან რიგითობით განლაგებული პრიმიტიული ტექსელების სიმრავლეს.
- **სტატისტიკური:** ტექსტურა არის გამოსახულების ან მის რომელიმე არეში ინტენსივობის მნიშვნელობების განაწილების რაოდენობრივი მახასიათებელი.

პირველი მიდგომა თავისი ბუნებრივი მიმზიდველობით მართლაც გამოიყენება რეგულარულ ხელოვნურ გამოსახულებებზე. მაგრამ პრაქტიკაში ხშირად იყენებენ უფრო ზოგად მეორე მიდგომას, რომელიც გამოთვლების ჩატარებისთვის უფრო მარტივია.

10.2. ტექსტურის აღწერა ტექსელების საშუალებით

იმის გამო, რომ ტექსტურა წარმოადგენს სივრცეში რაიმე წესით განლაგებული ელემენტარული ე.წ. **ტექსელების** (texel) სიმრავლეს, ამიტომ მისი სტრუქტურული აღწერა შესაძლებელია შეიცავდეს როგორც თვით ტექსელების, ასევე მათი სივრცითი მიმართებების აღწერას. ცხადია, ტექსელები უნდა ექვემდებარებოდეს სეგმენტაციას და ასევე მათი სივრცითი ურთიერთკავშირის განსაზღვრისათვის უნდა არსებობდეს შესაბამისი გამოთვლითი მეთოდი.

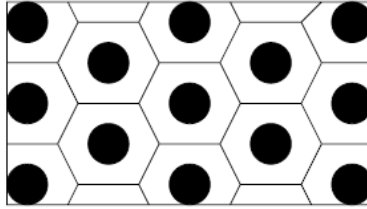
ტექსტურის სტრუქტურული გეომეტრიული აღწერის ერთ-ერთი მეთოდი შემდეგში მდგომარეობს: გამოსახულებაზე ტექსელებად აირჩევა ის არეები, რომელთა აღმოჩენა შესაძლებელია

ლია დამუშავების უბრალო, მაგალითად, ზღურბლური ბინარი-ზაციის მსგავსი ოპერაციების გამოყენებით. ტექსტების სივრცითი ურთიერთკავშირის მახასიათებლები განისაზღვრება მათთვის აგებული ვორონოის დიაგრამიდან.

ვთქვათ, გამოსახულებაზე აღმოჩნდა ტექსტების რაღაც სიმრავლე და ამასთან თითოეული მათგანი შეიძლება წარმოდგენილ იქნეს ერთი მახასიათებელი წერტილით, მაგალითად, ტექსტის სიმძიმის ცენტრით. ამ წერტილების სიმრავლე აღვნიშნოთ S . ამ სიმრავლის ნებისმიერი P და Q წყვილის შემართებელ მონაკვეთზე შეიძლება გავვლებულ იქნეს პერპენდიკულარი, რომელიც PQ მონაკვეთს შუაზე გაყოფს. ეს პერპენდიკულარი სიბრტყეს ყოფს ორ ნახევარსიბრტყედ. ამასთან, ერთი ნახევარსიბრტყის წერტილები უფრო ახლოსაა P წერტილთან, ხოლო მეორის წერტილები კი Q -თან. აღვნიშნოთ $H^Q(P)$ იმ წერტილების ნახევარსიბრტყე, რომლებიც უფრო ახლოსაა P წერტილთან PQ მონაკვეთის პერპენდიკულარის მიმართ. ასეთივე პროცედურა შეიძლება გამეორდეს S სიმრავლის ყოველი Q წერტილისთვის. P წერტილისთვის ვორონოის მრავალკუთხედი ეწოდება მრავალკუთხედს, რომელიც შეიცავს ყველა იმ წერტილს, რომლებიც P წერტილთან უფრო ახლოს იმყოფება, ვიდრე S სიმრავლის ნებისმიერ სხვა წერტილთან. ვორონოის მრავალკუთხედი განისაზღვრება გამოსახულებით

$$V(P) = \bigcap_{Q \in S, Q \neq P} H^Q(P).$$

ვორონოის მრავალკუთხედები, აგებული წრიული ტიპის ტექსტების სიმრავლისთვის, წარმოდგენილია 10.3 სურათზე. როგორც ვხედავთ, შიდა ტექსტებისთვის ვორონოის მრავალკუთხედები წარმოადგენს ექვსკუთხედებს, ხოლო სასაზღვრო ტექსტებისთვის კი განსხვავებული ფორმა აქვს.



სურ. 10.3

მას შემდეგ, რაც გამოსახულებაზე მოხდა ტექსტების აღმოჩენა და მათთვის აიგო ვორონოის მრავალკუთხედები, უკვე შეიძლება მრავალკუთხედების ფორმის ნიშან-თვისებების გამოთვლა. ამის შემდეგ ეს თვისებები შეიძლება გამოყენებულ იქნეს მრავალკუთხედების კლასტერებად დაჯგუფებისათვის. ეს კლასტერები გამოსახულებაზე მოგვცემს ტექსტების ერთგვაროვან არეებს. მაგალითად, 10.3 სურათის მსგავსი შაბლონი, თუ მას განვავრცობთ დიდი განზომილების გამოსახულებაზე, შეიძლება შეესაბამებოდეს ერთგვაროვანი ტექსტურის მქონე არეს, რომლის აღწერისათვის გამოსადეგი იქნება წესიერი ექვსკუთხედის ფორმის ნიშან-თვისებები.

10.3. ტექსტურის რაოდენობრივი მახასიათებლები

ტექსტურის ხარისხობრივი აღწერისთვის იყენებენ თვით ტექსტულის ზომას. სხვადასხვა ტექსტურებში მათ სხვადასხვა ზომები გააჩნია, რაც ლოკალური სტრუქტურის სივრცითი განმეორებადობის პერიოდთანაა დაკავშირებული. დიდი პერიოდი შეესაბამება მსხვილ სტრუქტურას, ხოლო მცირე – წვრილს. მართალია ტექსტულის ზომა არ არის საკმარისი იმისთვის, რომ მოვახერხოთ ტექსტურის რაოდენობრივი გაზომვა, მაგრამ ის გვაძლევს შესაძლებლობას მოვახდინოთ გარკვეული დასკვნის გაკეთება. კერძოდ, ნიშან-თვისებების მცირე რიცხვითი მნიშვნე-

ნელობები უნდა შეესაბამებოდეს წვრილ ტექსტურებს, ხოლო დიდი მნიშვნელობები – მსხვილს.

ტექსტების აღმოჩენა ბუნებრივ გამოსახულებებზე გაცილებით უფრო რთულია, ვიდრე ხელოვნურად შედგენილ გამოსახულებებზე. ამიტომ ნახევარტონურ (ან ფერად) გამოსახულებაზე პიქსელური მონაცემების საშუალებით ტექსტების აღმოჩენის სანაცვლოდ არჩევენ გამოითვალონ ტექსტურის რიცხვითი მახასიათებლები (მაგალითად, ტექსტურული სტატისტიკები, სივრცითი ავტოკორელაციური ფუნქცია, მოცემული წერტილის მიდამოში სიკაშკაშის ვარდნების რიცხვი და სხვ.). ასეთი მდგომის უპირატესობა იმაში მდგომარეობს, რომ რიცხვითი მახასიათებლები შესაძლებელია გამოყენებული იქნეს როგორც ტექსტურების სეგმენტაციისთვის, ასევე კლასიფიკაციისთვის.

10.3.1. კიდების სიმკვრივე და მიმართულება

კიდების გამოყოფა წარმოადგენს გამოყენების თვალსაზრისით კარგად ცნობილ და მარტივ გზას გამოსახულებაზე მახასიათებელი ნიშან-თვისებების აღმოჩენისთვის. ამიტომ ბუნებრივი იქნება, რომ ტექსტურული ანალიზისთვის გამოვიყენოთ კიდების აღმოჩენის დეტექტორი. გამოსახულებაზე ფიქსირებული ზომის მოცემულ არეში კიდების რაოდენობა წარმოადგენს ამ არის შევსებულობის გარკვეულ მახასიათებელს. კიდების გამოყოფისას ასევე ხდება მათი მიმართულებების გამოთვლა, რომლების გამოყენება ტექსტურის აღწერისათვის.

განვიხილოთ N პიქსელისგან შედგენილი არე. მოვახდინოთ ამ არის დამუშავება კიდის გრადიენტული ოპერატორით. ასეთი დამუშავების შედეგად არის ყოველი p პიქსელისთვის გვექნება შემდეგი ორი სიდიდე: 1) გრადიენტის მოდული

$Mag(p)$, და 2) გრადიენტის მიმართულება $Dir(p)$. ტექსტურის ერთ-ერთ უმარტივეს მახასიათებელ ნიშან-თვისებას წარმოადგენს კიდეების სიმკვრივე (კიდეების რიცხვი ფართობის ერთეულზე)

$$F = \frac{| \{p | Mag(p) \geq T \} |}{N}$$

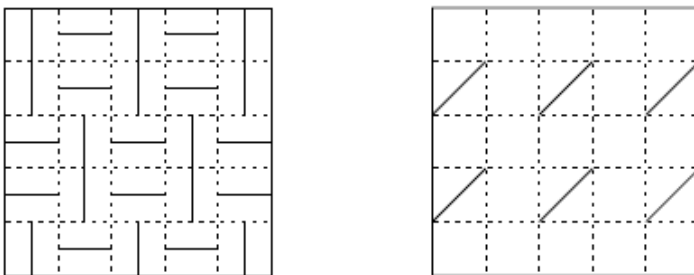
სადაც, T ზღურბლის სიდიდეა. კიდეების სიმკვრივე ახასიათებს ტექსტურის შეესებულობის ხარისხს, მაგრამ არა მის ორიენტაციას.

ტექსტურის რაოდენობრივი დახასიათებისთვის ზემოწარმოდგენილი მახასიათებელი შეიძლება გავაფართოოთ ტექსტურის ორიენტაციის მახასიათებლით. ამისთვის გამოვიყენოთ გრადიენტის მოდულისა და მიმართულების ჰისტოგრამები. აღვნიშნოთ $H_{mag}(R)$ -ით R არის პიქსელებში გრადიენტის მოდულის მნიშვნელობებით აგებული ნორმირებული ჰისტოგრამა, ხოლო H_{dir} -ით გრადიენტის მიმართულების ნორმირებული ჰისტოგრამა. ჩვეულებრივ, ორივე ჰისტოგრამა შეიცავს ფიქსირებული თანრიგების მცირე რაოდენობას (დაახლოებით 10), რომლებიც შეესაბამება გრადიენტის მოდულისა და მიმართულების მნიშვნელობათა დიაპაზონებს. აღნიშნული ჰისტოგრამების ნორმირება ხდება მათი გაყოფით R არეში პიქსელების N_R რაოდენობაზე. შედეგად, R არის ტექსტურის რაოდენობრივი აღწერა წარმოდგინება ჰისტოგრამების წყვილის საშუალებით:

$$F_{mag\ dir} = (H_{mag}(R), H_{dir}(R)).$$

განვიხილოთ 10.4 სურათზე წარმოდგენილი ორი გამოსახულება ზომით (5×5) -ზე. როგორც ვხედავთ, მარცხენა გამო-

სახელბა უფრო მეტად არის შევსებული, ვიდრე მარჯვენა. კერძოდ, მარცხენა გამოსახულებაზე კიდები არსებობს 25-ვე პიქსელში. შესაბამისად კიდების სიმკვრივე უდრის 1.0. მაშინ როდესაც, მარჯვენა გამოსახულებაში 25 პიქსელზე მოდის 6 კიდე, და აქედან გამომდინარე კიდების სიმკვრივე უდრის 0.24. ახლა ავაგოთ გრადიენტის მოდულისთვის ორთაწრიანი პისტოგრამები, თითო თანრიგი მუქი და ღია კიდებისთვის. რაც შეეხება გრადიენტის მიმართულების პისტოგრამებს, ავირჩიოთ სამი თანრიგი შესაბამისად ჰორიზონტალური, ვერტიკალური და დიაგონალური კიდებისთვის. მარცხენა გამოსახულება შეიცავს 6 მუქ და 19 ღია კიდეს.



სურ. 10.4

ამ გამოსახულებისთვის გრადიენტის მოდულის ნორმირებულ პისტოგრამას აქვს ასეთი სახე – (0.24; 0.76). ეს ნიშნავს, რომ კიდების 24% არის მუქი და 76% კი ღია. მარცხენა გამოსახულება შეიცავს 12 ჰორიზონტალურ და 13 ვერტიკალურ კიდეს, ხოლო დიაგონალურ კიდებს, არ შეიცავს. გრადიენტის მიმართულების ნორმირებულ პისტოგრამას ექნება შემდეგი სახე: (0.48; 0.52; 0.00), რაც ნიშნავს, რომ გამოსახულებაზე არის 48% ჰორიზონტალური კიდე, 52% ვერტიკალური და 0% დიაგონალური.

მარჯვენა გამოსახულება შეიცავს 6 ღია კიდეს, ხოლო მუქ კიდეს – არა. გრადიენტის მოდულის ნორმირებული ჰისტოგრამა ტოლია (0.00; 0.24). ამ გამოსახულებაზე არის მხოლოდ დიაგონალური კიდევები, რის გამოც გრადიენტის მიმართულების ჰისტოგრამას აქვს სახე (0.00; 0.00; 0.24). წარმოდგენილი ორი გამოსახულების ერთმანეთისგან განსასხვავებლად საკმარისი იქნებოდა კიდევების სიმკვრივის ნიშან-თვისების გამოყენება, მაგრამ ჰისტოგრამული ნიშან-თვისებები წარმოადგენს ტექსტურის აღწერისათვის გაცილებით უფრო მძლავრ მეთოდს.

ჰისტოგრამების ერთმანეთთან შესადარებლად გამოიყენება მათ შორის მანძილის ცნება. ორ n თანრიგიან H_1 და H_2 ჰისტოგრამას შორის მანძილი L_1 გამოითვლება ფორმულით:

$$L_1(H_1, H_2) = \sum_{i=1}^n |H_1[i] - H_2[i]|.$$

10.3.2. მომიჯნაობის მატრიცები

მომიჯნაობის მატრიცა (co-occurrence matrix) წარმოადგენს ორგანზომილებიან C მასივს (მატრიცას), რომელშიც სტრიქონებისა და სვეტების ინდექსები ქმნის გამოსახულებაზე პიქსელებისთვის დასაშვებ სიკაშკაშის მნიშვნელობათა (დონეთა) V სიმრავლეს. მაგალითად, ნახევარტონური გამოსახულებისთვის V შეიძლება იყოს ინტენსივობის დასაშვებ მნიშვნელობათა სიმრავლე, ხოლო ფერადი გამოსახულებისთვის კი შესაძლო ფერთა სიმრავლე. შესაბამისად, მომიჯნაობის მატრიცის განზომილება განისაზღვრება გამოსახულებაზე სიკაშკაშის დონეების რაოდენობით. C მატრიცის $C(i, j)$ ელემენტის რიცხვითი მნიშვნელობა მიუთითებს, თუ რამდენჯერ შეხვდა გამოსახულებაზე სიკაშკაშის i მნიშვნელობა რომელიმე მოცემულ სი-

ვრცით მიმართებაში (ჰორიზონტალური, ვერტიკალური ან დიაგონალური მიმართულებით) სიკაშკაშის j მნიშვნელობას. სივრცითი მიმართების მაგალითს შეიძლება წარმოადგენდეს ასეთი მიმართება: “ i -ს მნიშვნელობა j -ს მნიშვნელობისთვის წარმოადგენს მარჯვენა მეზობელს”.

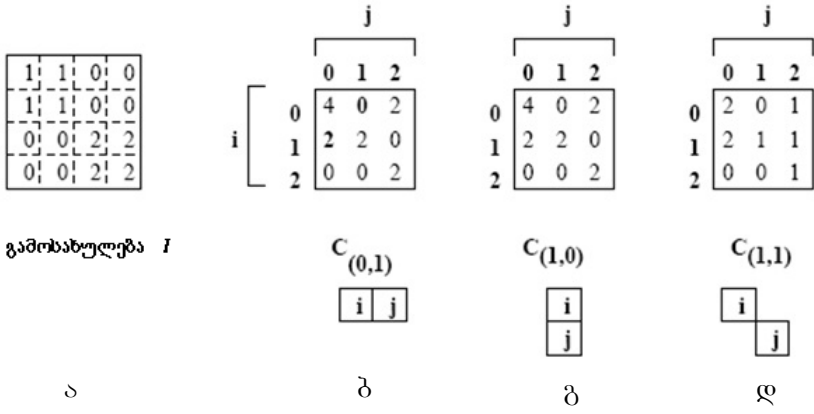
ზემოთქმულის ასახსნელად განვიხილოთ მაგალითი, რომელშიც V წარმოადგენს გამოსახულებაზე სიკაშკაშის მნიშვნელობათა სიმრავლეს, ხოლო სივრცითი მიმართება კი აღიწერება \mathbf{d} ვექტორის საშუალებით. ამ ვექტორით მოიცემა წანაცვლება i -ური და j -ური სიკაშკაშის მნიშვნელობის მქონე პიქსელებს შორის.

აღვნიშნოთ \mathbf{d} -თი წანაცვლების ვექტორი (dr, dc) , სადაც dr შეესაბამება გამოსახულებაზე გადაადგილებას სტრიქონების მიმართულებით (ზემოდან ქვემოთ), ხოლო dc კი გადაადგილებას გამოსახულების სვეტების მიმართულებით (მარცხნიდან მარჯვნივ). გამოსახულებაზე სიკაშკაშის მნიშვნელობათა სიმრავლე აღვნიშნოთ V -თი. მაშინ რაღაც I გამოსახულებისთვის ნახევარტონური მომიჯნაობის მატრიცა C_d განისაზღვრება შემდეგი გამოსახულებით:

$$C_d[i, j] = |\{[r, c] | I[r, c] = i, I[r + dr, c + dc] = j\}|,$$

სადაც $r = \overline{1, N}$, $c = \overline{1, M}$. 10.5,ა სურათზე მოცემულია (4×4) განზომილების I ნახევარტონური გამოსახულება ნახევარტონების სამი დონით, ანუ $V = \{0, 1, 2\}$. შესაბამისად, წარმოდგენილი გამოსახულებისთვის მომიჯნაობის მატრიცა (3×3) განზომილების იქნება. 10.5,ბ,გ,დ სურათზე მოცემულია ნახევარტონური მომიჯნაობის მატრიცები $C_{[0,1]}$, $C_{[1,0]}$ და $C_{[1,1]}$ სხვადასხვა წანაცვლების ვექტორებისთვის, კერძოდ, (ბ) $\mathbf{d} = (0, 1)$, (გ)

$d = (1,0)$, და (დ) $d = (1,1)$. $C_{[0,1]}$ მატრიცაში ელემენტი $[1,0]$ უდრის 2-ს. ეს ნიშნავს, რომ გამოსახულებაზე $j = 0$ მნიშვნელობა ორჯერ გვხვდება როგორც $i = 1$ -ის მარჯვენა მეზობელი. ელემენტი $[0,1]$ უდრის 0. ეს ნიშნავს, რომ მნიშვნელობა $j = 1$ ერთხელაც არ გვხვდება როგორც $i = 0$ პიქსელის მეზობელი. მომიჯნაობის ყველაზე მეტი მნიშვნელობა, რომელიც უდრის 4, იხსება $[0,0]$ ელემენტში. ამგვარად, მოცემულ გამოსახულებაზე ნულოვანი პიქსელი წარმოადგენს მეორე ნულოვანი პიქსელის მარჯვენა მეზობელს გამოსახულების ოთხ სხვადასხვა ადგილას.



სურ. 10.5

არსებობს სტანდარტული ნახევარტონური მომიჯნაობის მატრიცის ორი ძირითადი ნაირსახეობა. პირველს ეწოდება ნორმირებული ნახევარტონური მომიჯნაობის მატრიცა, რომელიც აღინიშნება N_d -თი და განისაზღვრება გამოსახულებით

$$N_d[i, j] = \frac{C_d[i, j]}{\sum_i \sum_j C_d[i, j]} \cdot$$

298

ამ გამოსახულებით გამოთვლილი მომიჯნაობის მნიშვნელობები მიეკუთვნება $[0,1]$ დიაპაზონს. ისინი შეიძლება განხილულ იქნეს როგორც ალბათობის მნიშვნელობები უფრო დიდი განზომილების მატრიცაში.

მეორე ნაირსახეობას ეწოდება **სიმეტრიული ნახევარტონური მომიჯნაობის მატრიცა** და აღინიშნება S_d . ამ მატრიცაში ხდება ერთმანეთთან სიმეტრიულ მიმართებაში მყოფი მომიჯნავე წყვილების დაჯგუფება:

$$S_d[i, j] = C_d[i, j] + C_{-d}[i, j].$$

მომიჯნაობის მატრიცების საშუალებით ხდება ტექსტურის თვისებების წარმოჩენა, მაგრამ მათი გამოყენება უშუალოდ გამოსახულების ანალიზისათვის მოუხერხებელია, მაგალითად, ორი ტექსტურის შედარების ამოცანაში. ამის ნაცვლად მომიჯნაობის მატრიცები გამოიყენება იმ ნიშან-თვისებების რიცხვითი მნიშვნელობების გამოთვლისთვის, რომლებიც გამოდგება ტექსტურის უფრო კომპაქტური წარმოდგენისთვის.

ნორმირებული მომიჯნაობის მატრიცის საფუძველზე შესაძლებელია გამოვითვალოთ ტექსტურის შემდეგი ნიშან-თვისებები:

$$\text{ენერჯია} = \sum_i \sum_j N_d^2[i, j];$$

$$\text{ენტროპია} = - \sum_i \sum_j N_d[i, j] \log_2 N_d[i, j];$$

$$\text{კონტრასტი} = \sum_i \sum_j (i - j)^2 N_d[i, j];$$

$$\text{ერთგვაროვნება} = \sum_i \sum_j \frac{N_d[i, j]}{1 + |i - j|};$$

$$\text{კორელაცია} = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d[i,j]}{\sigma_i \sigma_j},$$

სადაც, μ_i, μ_j – საშუალო მნიშვნელობები, ხოლო σ_i, σ_j – საშუალო კვადრატული გადახრებია გამოსახულების სტრიქონებისა და სვეტების მიმართულებით გამოთვლილი $N_d[i]$ და $N_d[j]$ ჯამებისთვის, რომლებიც თავის მხრივ გამოითვლება შემდეგი ფორმულების საშუალებით:

$$N_d[i] = \sum_j N_d[i,j];$$

$$N_d[j] = \sum_i N_d[i,j].$$

ტექსტურული ნიშან-თვისებების გამოთვლისას მნიშვნელოვან პრობლემას წარმოადგენს წანაცვლების \vec{d} ვექტორის შერჩევა. \vec{d} ვექტორის ისეთი მნიშვნელობების შესარჩევად, რომლებიც ტექსტურული სახის ყველაზე გამოხატულ სტრუქტურას შეესაბამება, გამოიყენება ჰიპოტეზების შემოწმების χ^2 სტატისტიკური კრიტერიუმი. \vec{d} ვექტორის ეს მნიშვნელობები შეესაბამება

$$\chi^2(d) = \left(\sum_i \sum_j \frac{N_d^2[i,j]}{N_d[i]N_d[j]} - 1 \right)$$

სიდიდის მაქსიმუმს.

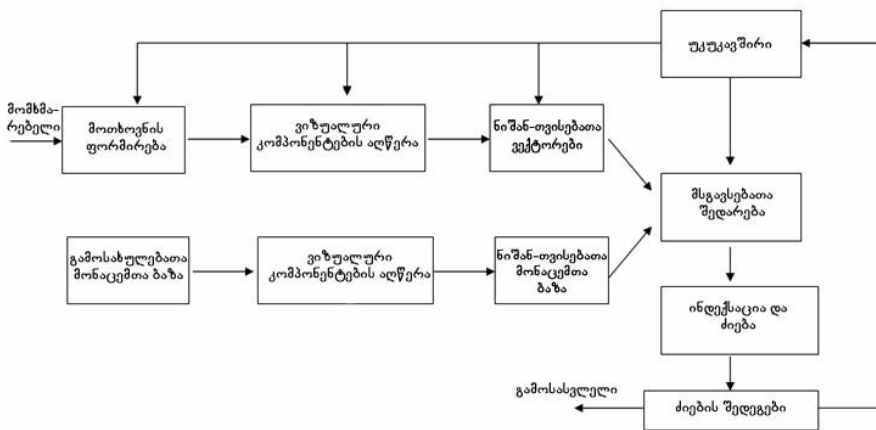
XI თავი. გამოსახულებების მოძიება მათი შინაარსის მიხედვით

11.1. გამოსახულებების მონაცემთა ბაზები

კომპიუტერების მექანიზმებისა და წარმადობის ზრდამ შესაძლებელი გახადა გამოსახულებების მონაცემთა ბაზების რეალური გამოყენება. ასეთი ბაზები არსებობს ხელოვნების ნაწარმოებთა სურათების, თანამგზავრებიდან მიღებული სურათების, სამედიცინო სურათების და ასევე საკოლექციო სურათების სახით. მხატვრული კოლექციების მომხმარებლებს შეიძლება გაუჩნდეთ რომელიმე მხატვრის სურათების ნახვის სურვილი, ან იმის გარკვევა, თუ რომელ მხატვარს ეკუთვნის მოცემული სურათი. სამედიცინო სურათების ბაზებით შეუძლიათ ისარგებლონ სტუდენტებმა, რომლებიც სწავლობენ ანატომიას, ან ექიმებმა, რომელთაც სჭირდებათ მოცემული დაავადების მქონე ავადმყოფების სურათები. ასეთი მონაცემთა ბაზებისადმი მიმართვები შეიძლება იყოს განსხვავებული და მრავალფეროვანი. უმრავლეს შემთხვევაში ასეთი ბაზები შეიძლება ინდუცირებული იყოს მხოლოდ გასაღები სიტყვების მიხედვით. ეს სიტყვები კომპიუტერში შეჰყავს ოპერატორს, რომელიც ახდენს გამოსახულებების კლასიფიკაციას კატეგორიების მიხედვით. მაგრამ ამავე დროს მონაცემთა ბაზებში გამოსახულებები შეიძლება მოძიებულ იქნას მათი ვიზუალური შინაარსის, კერძოდ, ფერთა განაწილების, არეების (ობიექტების) ტექსტურისა და ფორმების, არეების სივრცითი განლაგების მიხედვით გამოსახულების წარმოდგენისა და ინდექსაციისთვის.

შინაარსზე დამყარებული გამოსახულებების მოძიება წარმოადგენს მეთოდს, რომელიც დიდი მოცულობის მონაცემთა ბაზაში მომხმარებლის მოთხოვნიდან გამომდინარე გამოსა-

ხულების მოძიებისთვის იყენებს ვიზუალურ შინაარსს. შინაარსზე დამყარებულ ტიპიური გამოსახულებების მოძიების სისტემებში მონაცემთა ბაზაში არსებული გამოსახულებების ვიზუალური შინაარსი გამოიყოფა და აღიწერება მრავალგანზომილებიანი ნიშან-თვისებათა ვექტორების საშუალებით. ასეთი სისტემის ბლოკ-სქემა ნაჩვენებია 11.1 სურათზე.



სურ. 11.1

თავის მხრივ მონაცემთა ბაზაში გამოსახულებების ნიშან-თვისებათა ვექტორების საშუალებით ხდება ნიშან-თვისებათა მონაცემთა ბაზის ფორმირება. გამოსახულებების მოძიებისთვის მომხმარებელი უზრუნველყოფს სისტემას გამოსახულებები-მაგალითებით (ნიმუშებით), ან მოსაძებნი ობიექტების ესკიზებით. ამის შემდეგ სისტემა აღნიშნული ნიმუშებიდან ახდენს შესაბამისი ნიშან-თვისებათა ვექტორების ფორმირებას. შემდეგ ხდება მსგავსების, ანუ მანძილის გამოთვლა მოთხოვნაში მითითებულ გამოსახულებები-მაგალითების ნიშან-თვისებათა ვექტორებსა და მონაცემთა ბაზებში არსებულ გამოსახუ-

ლებებს შორის და ძიება ხორციელდება ინდექსაციის დახმარებით. ეს უკანასკნელი უზრუნველყოფს მონაცემთა ბაზაში ძიებისთვის ეფექტური გზის მოძებნას. ამისათვის არსებულ საძიებო სისტემებში ძიების პროცესის ვარიანტებისთვის გათვალისწინებულია უკუკავშირი, რაც არსობრივი შედეგების მიღების შესაძლებლობას იძლევა.

გამოსახულებების ძიების ასეთი ალგორითმები თავისი განვითარების საწყის ეტაპზეა. უკვე შემუშავებულია და გამოიყენება რამდენიმე კომერციული და კვლევითი სისტემა, სადაც გამოსახულებების მოძიება ხდება მათი შინაარსიდან გამომდინარე. ამავე დროს, მიმართვას ასეთ სისტემის მიმართ შეიძლება ჰქონდეს თავისი შეზღუდვებიც.

გამოსახულებების მონაცემთა ბაზასთან სამუშაოდ საჭიროა ძიების რაიმე ხერხის ქონა, რომელიც უფრო მოსახერხებელი და ეფექტური იქნება, ვიდრე ბაზაში ყველა გამოსახულების გადარჩევა.

ის კომპანიები, რომლებიც ქმნის გამოსახულებების მონაცემთა ბაზებს და უზრუნველყოფს მათთან კლიენტების წვდომას, ახორციელებს დამუშავების ორ ეტაპს: პირველ ეტაპზე ხდება კოლექციისათვის გამოსახულებების შერჩევა, ხოლო მეორეზე კი – მათი კლასიფიკაცია მათთვის საერთო კატეგორიებისა და გასაღები სიტყვების მინიჭებით. გამოსახულებების მოძიებისათვის შეიძლება გამოყენებულ იქნეს ისეთი ატრიბუტები, როგორცაა ზოგადი კატეგორია, მაგალითად, გამოსახულებაზე არსებული ობიექტების და ადამიანთა სახელები, შექმნის თარიღი და წყაროს დასახელება. რაიმე მიმართვის შესაბამისად ძიების პროცესის დაჩქარებისთვის ხდება გამოსახულებების ინდექსაცია ამ ატრიბუტების მიხედვით. რაც შეეხება ძიებას გასაღები სიტყვების მიხედვით, მას გააჩნია რიგი შეზღუდვებისა. თვით გასაღები სიტყვების შერჩევა რთულ ამოცანა

ნას წარმოადგენს და ამასთან არაცალსახაა. სხვადასხვა ადამიანებს გამოსახულების აღწერისათვის შეუძლიათ გამოიყენონ სხვადასხვა გასაღები სიტყვები, თუმცა მათი საშუალებით გამოსახულებების ძიება საკმაოდ არაეფექტურია.

ზოგიერთი გამოსახულებების მონაცემთა ბაზები შემუშავებული იქნა ძიების კონკრეტული სისტემის შრომისუნარიანობის სადემონსტრაციოდ. ამის მაგალითს წარმოადგენს IBM კომპანიის მიერ შემუშავებული სისტემა – **მოთხოვნა გამოსახულებების შინაარსის მიხედვით (QBIC, Query by Image Content)**. ამ სისტემაში გამოსახულებების ძიება ხორციელდება მათი ვიზუალური შინაარსიდან გამომდინარე, კერძოდ, მხედველობაში მიიღება ფერებს შორის პროცენტული თანაფარდობა, ფერების სივრცითი განაწილება და ტექსტურა. მეორე სისტემა – **საძიებო მანქანა Virage (Virage Search Engine)** შემუშავებული იქნა კომპანია Virage, Inc.-ის მიერ. ამ სისტემაში გამოსახულებების ძიება ხორციელდება ფერის, კომპოზიციის, ტექსტურისა და სტრუქტურის მიხედვით.

11.2. მოთხოვნების აგება ნიმუშის მიხედვით

მონაცემთა ბაზების ტერმინოლოგიით **მოთხოვნა ნიმუშის მიხედვით (query-by-example, QBE)** ეწოდება რაღაც ცხრილის სახით წარმოდგენილ მოთხოვნას, რომელიც შევსებულია მნიშვნელობებითა და შეზღუდვებით. ეს ცხრილი შემდეგ შეიძლება კონვერტირებული იყოს SQL ენის ფორმატში. ნიმუშის მიხედვით მოთხოვნების აგების პირველი სისტემა შემუშავებული იქნა კომპანია IBM-ის მიერ. დღეისდღეობით მსგავსი ტიპის სისტემის მაგალითს წარმოადგენს Microsoft Access.

განსხვავებით სტანდარტული ტიპის რელაციური ბაზებისგან, რომლებშიც ატრიბუტებს რიცხვები ან ტექსტური მო-

ნაცემები წარმოადგენს, ნიმუშის მიხედვით მოთხოვნების ფორმირების მექანიზმი დაიყვანება მომხმარებლისთვის მოხერხებულ ინტერფეისის მიცემაში ფორმის მოთხოვნის პარამეტრებით შესავსებად და არ შეიცავს სხვა რაიმე დამატებით საშუალებებს.

გამოსახულებების მონაცემთა ბაზებში მოთხოვნების ნიმუშის მიხედვით გამოყენება უფრო ფართო შესაძლებლობებს იძლევა. მიმართვის ატრიბუტების მნიშვნელობების დაბეჭდვის ნაცვლად, გამოსახულებების მონაცემთა ბაზის მომხმარებელმა სისტემას შეუძლია წარუდგინოს გამოსახულების ნიმუში, ან დახატოს იგი ეკრანზე ინტერაქტიულ რეჟიმში, ან სქემატურად გამოსახოს საძიებელი ობიექტის კონტური (ესკიზი). ამის შემდეგ ძიების სისტემამ მონაცემთა ბაზაში უნდა მოიძიოს წარმოდგენილი გამოსახულების მსგავსი გამოსახულებები, ან გამოსახულებები, რომლებზედაც ეს ობიექტებია განლაგებული. ამაში მდგომარეობს შინაარსის მიხედვით გამოსახულებების მოძიების ყველა სისტემის არსი. თითოეულ მათგანს გააჩნია თავისი საკუთარი საშუალებები როგორც მოთხოვნის წარსადგენად, ასევე მსგავსების განსაზღვრისათვის მოთხოვნასა და მონაცემთა ბაზის კონკრეტულ გამოსახულებას შორის და ძიების შედეგად მიღებული გამოსახულებებიდან შესაბამისი გამოსახულების არჩევისათვის.

მსჯელობის ზოგადობისთვის ჩავთვალოთ, რომ მოთხოვნა შედგება მომხმარებლის მიერ წარმოდგენილი გამოსახულებისგან და მასთან ერთად შეზღუდვების რაღაც კრებულისგან. გამოსახულების როლში შეიძლება გამოდიოდეს ციფრული ფოტო, მომხმარებლის მიერ უხეშად დახატული სურათი, ან კონტურული გამოსახულება. ასევე შეიძლება მოცემული იყოს ცარიელი გამოსახულებაც. მაგრამ ამ შემთხვევაში მოძიებული გამოსახულებების სიმრავლე უნდა აკმაყოფილებდეს მოთხოვნას.

ნაში მითითებული შეზღუდვების ერთობლიობას. შეზღუდვები შეიძლება მოცემული იყოს რაღაც გასაღები სიტყვების სახით, ან შეიძლება იძლეოდეს იმ ობიექტებს, რომლებიც უნდა იყოს გამოსახულებაზე, ან უნდა აღწერდეს ამ ობიექტებს შორის სივრცით მიმართებებს. ყველაზე უფრო ზოგად სახეში მოთხოვნა შეიცავს ციფრულ გამოსახულებას, რომელიც უნდა შედარდეს მონაცემთა ბაზაში არსებულ გამოსახულებებს, თანახმად გამოსახულებებს შორის მანძილის განსაზღვრისთვის გამოყენებული მეტრიკისა. თუ მსგავსების მნიშვნელობა უდრის 0-ს, ეს ნიშნავს, რომ გამოსახულება ზუსტად შეესაბამება მომხმარებლის მოთხოვნას. თუ მსგავსების მნიშვნელობა მეტია 0-ზე, ეს ნიშნავს, რომ გამოსახულებასა და მოთხოვნას შორის არსებობს მსგავსების სხვადასხვა ხარისხები. საძიებო სისტემები ჩვეულებრივად მონაცემთა ბაზებიდან იძლევა გამოსახულებათა კრებულს, დალაგებულს მათი მსგავსების სხვადასხვა ხარისხის მიხედვით.

11.3. გამოსახულებებს შორის მსგავსების განსაზღვრა

იმისათვის, რომ რაოდენობრივად განისაზღვროს მსგავსება მოთხოვნაში მითითებულ გამოსახულებასა და მონაცემთა ბაზაში არსებულ გამოსახულებებს შორის, იყენებენ მანძილის განსაზღვრის გარკვეულ მეტრიკას, ან მახასიათებელს. გამოსახულებათა მსგავსების მახასიათებლები შეიძლება დაიყოს ოთხ ჯგუფად:

- მსგავსება ფერების მიხედვით;
- ტექსტურული მსგავსება;
- ფორმის მსგავსება;

- ობიექტების და მათი მიმართებების მსგავსება.

11.3.1. ფერთა მსგავსების მახასიათებლები

ხშირად ასეთ მახასიათებლებად ირჩევენ ძალიან მარტივებს. ისინი საშუალებას იძლევა, რომ ერთი გამოსახულების ფერების შემადგენლობა შედარდეს მეორე გამოსახულების ფერების შემადგენლობას, ან მოთხოვნაში მოცემულ პარამეტრებს. მაგალითად, QBIC სისტემაში მომხმარებელს შეუძლია თავის მოთხოვნაში მიუთითოს საძებნ გამოსახულებებში ფერების პროცენტული შემადგენლობა. მომხმარებელს შეუძლია ფერების ცხრილიდან აირჩიოს 5 ფერამდე და მიუთითოს თითოეული მათგანის რაოდენობრივი შემადგენლობა პროცენტებში ყოველი საძიებელი გამოსახულებისთვის. ამის შემდეგ სისტემა იწყებს იმ გამოსახულებების მოძიებას, რომლებშიც ფერთა პროცენტული შემადგენლობა ახლოსაა მოთხოვნაში მითითებულ პარამეტრებთან. ამასთან მხედველობაში არ მიიღება გამოსახულებაზე სხვადასხვა ფერების პიქსელების სივრცითი განაწილება. გამოსახულებების მონაცემთა ბაზებში სასურველი გამოსახულების ძიების მსგავსი ხერხი ეფუძნება ფერთა ჰისტოგრამების შეპირისპირებას. ამისათვის შეიძლება გამოყენებულ იქნეს ფერთა ჰისტოგრამა. ამ დროს ძიების სისტემის მომხმარებელი მიუთითებს გამოსახულება-ნიმუშს და ირჩევს ბრძანებას იმ გამოსახულებების მოძიებისთვის, რომლებიც ახლოსაა გამოსახულება-ნიმუშთან მანძილის თვალსაზრისით ფერთა ჰისტოგრამის საფუძველზე. მანძილის ზომები ფერთა ჰისტოგრამის საფუძველზე უნდა ითვალისწინებდეს ორი სხვადასხვა ფერის მსგავსების რაოდენობრივი შეფასების შესაძლებლობას. მაგალითად, QBIC სისტემაში მანძილი ფერთა ჰისტოგრამის საფუძველზე შემდეგნაირად განისაზღვრება:

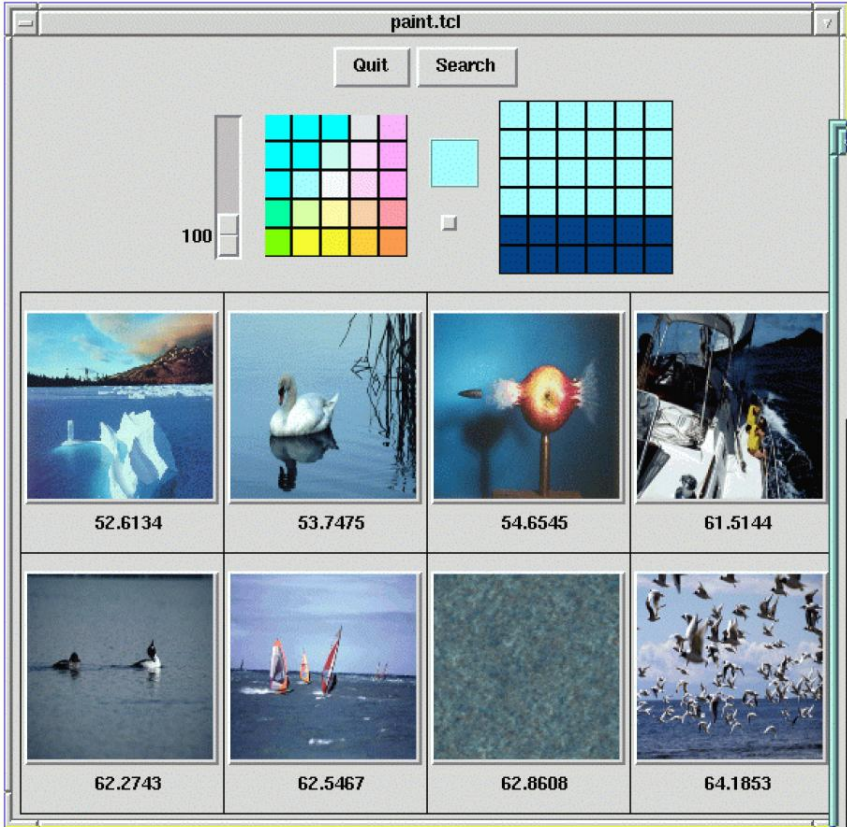
$$d_{hist}(I, Q) = (h(I) - h(Q))^T A(h(I) - h(Q)),$$

სადაც, $h(I)$ და $h(Q)$ წარმოადგენს შესაბამისად I და Q გამოსახულებების K ჯიბიან ჰისტოგრამებს, ხოლო A კი $K \times K$ განზომილების მსგავსების მატრიცაა. მსგავსების მატრიცაში ელემენტები, რომელთა მნიშვნელობები ახლოსაა 1-თან, შეესაბამება მსგავს ფერებს, ხოლო ელემენტები, რომლებიც ახლოსაა ან ტოლია 0-ის, შეესაბამება მკვეთრად განსხვავებულ ფერებს.

მანძილის განსაზღვრის კიდევ ერთ ზომას წარმოადგენს **ფერთა მაკეტი** (color layout). მოთხოვნის ფორმირებისას მომხმარებელს, ჩვეულებრივ მიეწოდება ცარიელი ბადე. მომხმარებელმა ამ ბადის თითოეული უჯრისთვის მნიშვნელობა შეიძლება აირჩიოს ფერების ცხრილიდან და მაკეტის ბადე გააფერადოს მათი საშუალებით.

საძიებო სისტემა ძიების შედეგად ახდენს ფერთა ასეთი პალიტრის მქონე გამოსახულებების ნაკრების შერჩევას. ფერთა მაკეტის მიხედვით ძიების ჩატარებისას გაფერადებული ბადის ნაცვლად მოთხოვნაში შეიძლება მითითებულ იქნეს გამოსახულება-ნიმუში. ასეთ შემთხვევაში სისტემა იწყებს იმ გამოსახულებების ძიებას, რომლებშიც ფერთა სივრცითი განაწილება ახლოსაა ფერთა განაწილებასთან გამოსახულება-ნიმუშში. ფერთა მაკეტის გამოყენებისას მომხმარებელს მიეწოდება ცარიელი ბადე. მომხმარებელმა ამ ბადის თითოეული უჯრის შესავსებად შეიძლება გამოიყენოს ცხრილიდან აღებული მისთვის სასურველი ფერი. 11.2 სურათზე მოცემულ შემთხვევაში მომხმარებელმა თავისი მოთხოვნის ფორმირებისთვის ფერთა ცხრილიდან (ზედა მარცხენა ბადე) გამოიყენა ორი ფერი და მათი საშუალებით გააფერადა (6×6) განზომილების ბადის მაკეტი (ზედა მარჯვენა ბადე).

გაფერადებული ბადის ფერთა მაკეტის საფუძველზე გამოსახულებების მსგავსების რაოდენობრივი შეფასებისთვის საჭიროა მათ შორის მანძილის განსაზღვრა, რომელიც მხედველობაში მიიღებდა ორი გაფერადებული ბადის შინაარს.



სურ. 11.2

მანძილის ამ ზომამ, d_{color} უნდა უზრუნველყოს მოთხოვნაში მითითებული ბადის თითოეული უჯრედის შედარება მონაცემთა ბაზის ნებისმიერი გამოსახულების გაფერადებული ბადის

შესაბამის უჯრედთან. ყველა უჯრედის შედარების შედეგები ერთიანდება გამოსახულებებს შორის **ბადეზე დამყარებული ფერთა მსგავსების** (gridded color distance) მნიშვნელობის

$$d_{gridded_color}(I, Q) = \sum_g \hat{d}_{color}(C^I(g), C^Q(g))$$

მისაღებად, სადაც, $C^I(g)$ არის მონაცემთა ბაზიდან I გამოსახულების ბადის g უჯრედის ფერი, ხოლო $C^Q(g)$ კი მოთხოვნაში მოცემული Q გამოსახულების შესაბამისი g უჯრედის ფერი.

ბადეში შემავალი უჯრედების ფერების წარმოსადგენად გამოიყენება სხვადასხვა ხერხები, კერძოდ:

- ბადის უჯრედის საზღვრებში ფერის საშუალო მნიშვნელობა;
- ფერის საშუალო მნიშვნელობა და საშუალო კვადრატული გადახრა;
- ფერების მრავალჯიბიანი ჰისტოგრამა.

ბადეების უჯრედებს შორის მანძილის ზომა d მხედველობაში უნდა იღებდეს ფერის წარმოდგენის მოცემულ ხერხს და უზრუნველყოფდეს ისეთი მანძილების მიღებას, რომელთა ინტერპრეტაციაც შესაძლებელია. მაგალითად, თუ ფერის საშუალო მნიშვნელობა წარმოდგენილია რიცხვების (R, G, B) სამეულის სახით, მაშინ ასეთ შემთხვევაში მსგავსების მისაღებ ზომად (მაგრამ არა საუკეთესოდ) შეიძლება გამოყენებული იყოს ასეთი სახის ზომა

$$\hat{d} = \|R^l, G^l, B^l - R^l, G^l, B^l\|^2.$$

ანალოგიური მიდგომა გამოიყენება მანძილის განსაზღვრისათვის ჰისტოგრამების საფუძველზე QBIC სისტემაში.

11.3.2. ტექსტურული მსგავსების მახასიათებლები

ტექსტურული მსგავსების მახასიათებლები უფრო რთულია, ვიდრე ფერების მიხედვით მსგავსების მახასიათებლები. გამოსახულებებს, რომლებიც ერთმანეთთან ახლოსაა ტექსტურული მსგავსების მახასიათებლებით, უნდა ჰქონდეს ფერების (ან სიკაშკაშის მნიშვნელობების) ერთნაირი სივრცითი განაწილება. თუმცა ამ გამოსახულებების ფერების (ან სიკაშკაშეების) მნიშვნელობები ერთმანეთს შეიძლება არ ემთხვეოდეს. ორი ტექსტურის მსგავსების შეფასებისთვის შეიძლება გამოყენებულ იქნეს X თავში განხილული ტექსტურული მახასიათებლები.

ტექსტურული მსგავსების მახასიათებლების განხილვისას საჭიროა დაზუსტდეს ორი ასპექტი:

- ტექსტურის წარმოდგენა;
- მსგავსების განსაზღვრა ტექსტურის არჩეული წარმოდგენის გათვალისწინებით.

ტექსტურის წარმოდგენის ერთ-ერთი ყველაზე გავრცელებული ხერხი ეფუძნება ტექსტურული აღწერის ვექტორის გამოყენებას (texture description vector). ეს არის ვექტორი, რომლის რიცხვითი კომპონენტები შეესაბამება მოცემული გამოსახულების ტექსტურის განზოგადებულ პარამეტრებს ან გამოსახულების გარკვეულ არეს. ტექსტურული აღწერის ვექტორების

მაგალითებს წარმოადგენს: პარალიკის ხუთკომპონენტიანი ვექტორი, რომელშიც ინახება შესასვლელი მატრიციდან აღებული ნიშან-თვისებები და ცხრაკომპონენტიანი ვექტორი, რომლის კომპონენტები შეესაბამება ლავსის ენერგეტიკულ ტექსტურულ მახასიათებლებს. ტექსტურული აღწერის ვექტორი შეიძლება გამოყენებულ იქნეს მთლიანი გამოსახულების ტექსტურისთვის, მაგრამ ასეთ შემთხვევაში კარგი შედეგი მიიღწევა მხოლოდ ერთგვაროვანი ტექსტურის მქონე გამოსახულებებისთვის. უფრო ზოგადი სახის გამოსახულებებისთვის ტექსტურული აღწერის ვექტორი გამოითვლება ყოველი პიქსელისთვის მის მცირე მიდამოში (15 × 15). ამის შემდეგ ხდება პიქსელების დაჯგუფება კლასტერირების გამოყენებით, რომელიც თავის უნიკალურ ჭდეს ანიჭებს ყოველ ახლად აღმოჩენილ ტექსტურულ კატეგორიას.

პიქსელებს შეიძლება შევუფარდოთ ტექსტურული აღწერის ვექტორი და ამის შემდეგ მოვახდინოთ პიქსელების მარკირება შესაბამის ტექსტურულ კატეგორიასთან კუთვნილების მიხედვით. ეს საშუალებას იძლევა განვსაზღვროთ მანძილის სხვადასხვა ტექსტურული ზომები. ასეთი ზომის უმარტივეს მაგალითს წარმოადგენს **მანძილი არჩეული ნიმუშიდან** (pick-and-click distance). ამ დროს მომხმარებელი ირჩევს ტექსტურას მოთხოვნაში მითითებული მოცემული გამოსახულების ტექსტურული არის შიგნით ნებისმიერ პიქსელზე თავის დაწკაპუნებით ან ირჩევს ტექსტურის ნიმუშს ტექსტურების რომელიმე მისაწვდომი კრებულიდან. ტექსტურის არჩეული ნიმუში წარმოიდგინება ტექსტურული აღწერის ვექტორის სახით, რომელიც ძიების პროცესში ედარება მონაცემთა ბაზაში მყოფი გამოსახულებების ტექსტურული აღწერის ვექტორებს. მანძილის ზომა განისაზღვრება შემდეგი გამოსახულებით:

$$d_{pic_and_click}(I, Q) = \min_{i \in I} \|T(i) - T(Q)\|^2,$$

სადაც, $T(i)$ წარმოადგენს I გამოსახულების i -ური პიქსელის ტექსტურული აღწერის ვექტორს, ხოლო $T(Q)$ – მომხმარებლის მიერ არჩეული იმ ტექსტურული ნიმუშის ტექსტურული აღწერის ვექტორს, რომელიც გამოყენებულია მოთხოვნაში. შეიძლება მოგვეჩვენოს, რომ ეს მეთოდი მოითხოვს დიდი მოცულობის გამოთვლებს, მაგრამ მოცულობა შეიძლება მნიშვნელოვნად შემცირდეს, თუ გამოსახულებების მონაცემთა ბაზაში შევინახავთ თითოეული გამოსახულების ტექსტურული კატეგორიების სიას. ტექსტურული კატეგორიები განისაზღვრება კლასტერიზაციის პროცედურების გამოყენებით. ძიების პროცესში მონაცემთა ბაზის ყოველი გამოსახულებისთვის აუცილებელია მოხდეს მოთხოვნაში მოცემული ტექსტურული აღწერის ვექტორის შედარება ამ გამოსახულებისთვის სიაში შენახული ტექსტურული აღწერის ვექტორებთან.

მანძილის განსაზღვრა ტექსტურის არჩეული ნიმუშიდან, როგორც ეს ზემოთ იყო გამოყენებული, მომხმარებლის მხრიდან მოითხოვს ტექსტურის ინტერაქტიულად მითითებას. ამის გამო გაძნელებულია მანძილის ასეთი ზომის ავტომატური გამოყენება მოთხოვნაში აღნიშნული გამოსახულების მიმართ. ტექსტურული მსგავსებისთვის უფრო ზოგად ზომას წარმოადგენს ზემოგანხილული ფერების მსგავსებისთვის გამოყენებული ზომის განზოგადება ტექსტურების შემთხვევისთვის. მოთხოვნაში მოცემულ გამოსახულებას ედება ბადე და მისი ყოველი უჯრედისთვის იანგარიშება ტექსტურული აღწერის ვექტორი. ანალოგიური პროცედურა ტარდება მონაცემთა ბაზის გამოსახულებებისთვისაც. **ბადეზე დამყარებული ტექსტურათა მსგავსება** (gridded texture distance) გამოითვლება შემდეგი ფორმულით:

$$d_{gridded_texture}(I, Q) = \sum_g d_{texture}(T^I(g), T^Q(g)),$$

სადაც, $d_{texture}$ შეიძლება იყოს ბადის უჯრედებს შორის ევკლიდური მანძილი ან მანძილის რომელიმე სხვა ზომა.

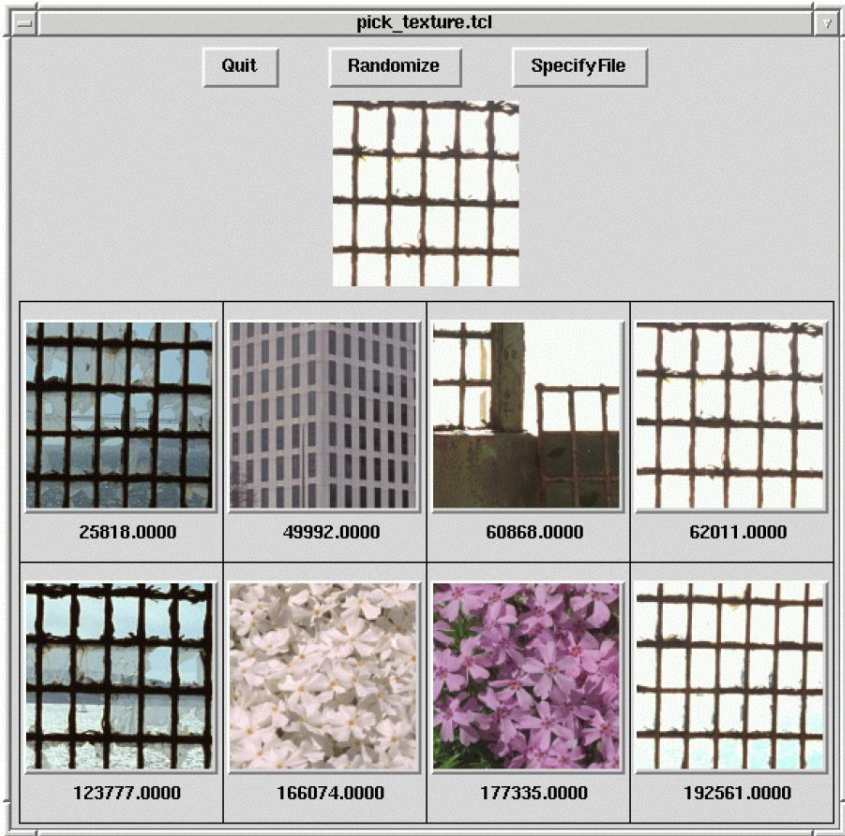
ფერთა ჰისტოგრამებს შორის მანძილის განსაზღვრის მსგავსად შეიძლება მანძილი განისაზღვროს ასევე ტექსტურულ ჰისტოგრამებს შორისაც. ამისთვის თითოეული ტექსტურული კატეგორიისთვის ჰისტოგრამაში წარმოდგენილია პიქსელთა რაოდენობა, რომელთა ტექსტურული აღწერის ვექტორი მიეკუთვნება მოცემულ კატეგორიას.

მომხმარებლის მოთხოვნის საფუძველზე მონაცემთა ბაზებში მსგავსი ტექსტურების მქონე გამოსახულებების ძიების შედეგები წარმოდგენილია 11.3 სურათზე.

11.3.3. ფორმის მსგავსების მახასიათებლები

ფერი და ტექსტურა შეიძლება გამოყენებულ იქნეს როგორც გამოსახულების გლობალური ატრიბუტები. ფერსა და ტექსტურაზე დამყარებული მანძილის ზომა განკუთვნილია იმის განსაზღვრავად, შეიცავს თუ არა რომელიმე გამოსახულება მოცემულ ფერს ან ტექსტურას და მოთავსებულია თუ არა ამ ფერის და ტექსტურის არეები გამოსახულების სივრცის იმავე არეებში, რაც მოთხოვნაში მითითებულ გამოსახულებაზეა. ფორმა არ წარმოადგენს გამოსახულების ატრიბუტს. აზრი არ აქვს შეკითხვას, თუ რა ფორმა აქვს გამოსახულებას. ფორმის ცნება შეიძლება გამოყენებულ იქნეს გამოსახულების რომელიმე არის მიმართ. ფორმის მახასიათებლების გამოყენებისას საჭირო ხდება დამატებითი დამუშავების ოპერაციების (განსხვავებით ფერის და ტექსტურის მახასიათებლებისგან) ჩატარება, რადგანაც ფორმის მსგავსების მახასიათებლების გამოთვლის

წინ საჭიროა არეების იდენტიფიკაცია. ხშირად ეს ხდება ხელით, თუმცა ზოგიერთ შემთხვევაში შესაძლებელია ავტომატური სეგმენტაციის გამოყენებაც.



სურ. 11.3

ობიექტების ამოცნობის მეთოდები მათ ორგანოზომილებიან სახეთა ფორმაზე დაყრდნობით მნიშვნელოვან როლს ასრულებს გამოსახულებათა ანალიზის ამოცანებში. ჩვენ მიერ V თავში უკვე შესწავლილი იქნა ბინარული გამოსახულების

არეების ზოგიერთი თვისება. შესწავლის შემდეგ ეს თვისებები ჩათვლილ იქნა როგორც ფორმის გლობალური მახასიათებლები (global shape properties), რადგანაც ისინი აღწერს ფორმას მთლიანობაში. გამოსახულების ორი არის ფორმის შედარება შეიძლება განხორციელებულ იქნეს ფორმის გლობალური თვისებების გამოყენებით სახეთა ამოცნობის სტატისტიკური მეთოდების საფუძველზე (VIII თავი). არეების ფორმების შეპირისპირებისთვის შეიძლება ასევე გამოყენებულ იქნეს სტრუქტურული მეთოდები. ამ დროს შესაძლებელია ფორმა აღიწეროს მისი პრიმიტიული ელემენტების ჩამონათვალით და მათ შორის სივრცითი მიმართებებით. იმის გამო, რომ ფორმის სტრუქტურული აღწერა წარმოადგენს რელაციურ გრაფს, ფორმის აღწერების შეპირისპირებისთვის შეიძლება გამოყენებულ იქნეს გრაფების შეპირისპირების მეთოდები. ასეთი მიდგომა წარმოადგენს ძალიან მძლავრ იარაღს, რადგან ამ შემთხვევაში ფორმის აღწერა თავის თავში მოიცავს სივრცით მიმართებებს, რომლებიც ინვარიანტულია ორგანოზომილებიანი გარდაქმნების უმეტესობის მიმართ. მაგრამ თვით გრაფების შეპირისპირების პროცესი შეიძლება მიმდინარეობდეს ძალიან ნელა, რადგან ელემენტების რიცხვის ზრდასთან ერთად დამუშავების დროც ექსპონენციალურად იზრდება.

შინაარსის მიხედვით გამოსახულებების ძიების სისტემებში სასურველია გამოყენებულ იქნეს მეთოდები, რომლებიც სწრაფად განსაზღვრავს თუ რამდენად ახლოსაა ერთმანეთთან გამოსახულების არის ფორმა და მოთხოვნაში მითითებული ფორმა. ხშირად გვაქვს ისეთი მოთხოვნა, რომ ფორმების შეპირისპირების მეთოდები ინვარიანტული იყოს გადატანისა და ზომის მიმართ, ან ბრუნვის მიმართ, რათა მოცემული ფორმის არე გამოსახულებაზე მოძებნილ იქნეს სივრცეში მისი ორიენტაციისაგან დამოუკიდებლად. თუმცა, მონაცემთა ბაზაში

გამოსახულებების მოძიებისთვის ინვარიანტულობა ბრუნვის მიმართ ყოველთვის არ მოითხოვება. არსებობს ბევრი გამოსახულება, რომლებზედაც ობიექტებს ერთი და იგივე ორიენტაცია აქვს. მაგალითად, ხეები, შენობები, მანქანები და ა.შ.

არსებობს ფორმის მრავალი მახასიათებელი. მათ შორის უხეში გლობალური მახასიათებლები, რომლებიც ამარტივებს, მაგრამ არ უზრუნველყოფს ობიექტების ამოცნობას, ასევე დეტალიზებული მახასიათებლებიც, რომელთა საშუალებითაც ხდება ძალიან სპეციფიკური ფორმის ობიექტების ამოცნობა. მარტივი მახასიათებლის მაგალითს წარმოადგენს **ფორმის ჰისტოგრამა**. ის საშუალებას იძლევა, რომ დამუშავების პროცესიდან ამოღებულ იქნეს ის არე, რომელიც არ შეიძლება წარმატებულად შევუპირისპიროთ მოთხოვნაში მითითებულ ფორმას. ამავე დროს, ფორმის ჰისტოგრამის გამოყენებისას შეიძლება ადგილი ჰქონდეს მრავალ მცდარ მოქმედებას. ამ მხრივ უფრო ეფექტურია საზღვრების შეპირისპირების მეთოდები. ისინი ოპერირებენ არის საზღვრის სახით წარმოდგენილ ფორმასთან. ეს მეთოდები საშუალებას იძლევა, რომ ვიპოვოთ არეები, რომელთა ფორმები აღიწერება მსგავსი საზღვრებით. ესკიზურ სახეში წარმოდგენილი ფორმების შეპირისპირება შეიძლება კიდევ უფრო ზუსტი აღმოჩნდეს, რადგან ამ დროს განიხილება არა მარტო ერთი ობიექტის საზღვარი, არამედ გამოსახულებაზე არსებული კონტურების სიმრავლე, რომლებიც შეესაბამება ერთ ან მეტი რაოდენობის ობიექტებს. კონტურების ეს სიმრავლე (ესკიზი) შეეპირისპირება მოთხოვნაში მოცემულ კონტურულ გამოსახულებას, რომელიც დახატული ან არჩეული იყო მომხმარებლის მიერ. ფორმის ჰისტოგრამული და ესკიზური მახასიათებლები განხილული იქნება ქვემოთ.

11.3.4. ფორმის ჰისტოგრამები

ზემოთ ჩვენ განვიხილეთ ჰისტოგრამების გამოყენება გამოსახულებათა შეპირისპირებისთვის მათი ფერისა და ტექსტურის მიხედვით მსგავსების დადგენის მიზნით. ჰისტოგრამების გამოყენება მიმზიდველია იმით, რომ მათ შორის მანძილის განსაზღვრა მარტივია და ამის გამო ადვილად გამოთვლადი. აქედან გამომდინარე, სრულიად ბუნებრივია ჰისტოგრამების გამოყენება ფორმების აღწერის შეპირისპირებისთვის. აქ მთავარი პრობლემა მდგომარეობს იმ ცვლადის განსაზღვრაში, რომლითაც უნდა აიგოს ჰისტოგრამა.

განვიხილოთ ნებისმიერი ფორმის მქონე ბმული არე ბინარულ გამოსახულებაზე. დავუშვათ, რომ ამ არის პიქსელების მნიშვნელობები 1-ის ტოლია, ხოლო გამოსახულების ყველა სხვა დანარჩენი პიქსელის კი 0-ის. ჰისტოგრამების შეპირისპირების ერთ-ერთი ხერხი მდგომარეობს **პროექციულ შეპირისპირებაში** (projection matching), რაც გულისხმობს აღნიშნული არის გამოსახულების სტრიქონებსა და სვეტებზე ჰორიზონტალური და ვერტიკალური პროექციების გამოყენებას. ვთქვათ, არე მოიცავს n სტრიქონსა და m სვეტს. ავავთ ჰისტოგრამა, სადაც ყოველ სტრიქონსა და სვეტს შევუსაბამოთ ჰისტოგრამის ერთი თანრიგი. ჰისტოგრამის თანრიგის მნიშვნელობა ტოლია შესაბამის სტრიქონში ან სვეტში მოთავსებული ერთის ტოლი პიქსელების რაოდენობისა. ასეთნაირად შესაძლებელია $n + m$ თანრიგიანი ჰისტოგრამის აგება, რომელიც გამოდგება ერთნაირი ზომის მქონე არეების შეპირისპირებისათვის.

იმისათვის, რომ პროექციული შეპირისპირება არ იყოს დამოკიდებული მასშტაბზე, შეიძლება მოვახდინოთ სტრიქონებისა და სვეტების შესაბამისი თანრიგების დაფიქსირება. თუ თანრიგების მიმდევრობას ისეთნაირად განვსაზღვრავთ, რომ

ისინი შეესაბამებოდეს გამოსახულებაზე არის დათვალიერების მიმართულებას ზედა მარცხენა კუთხიდან ქვედა მარჯვენა კუთხისკენ, მაშინ შესაძლებელია, რომ უზრუნველყოფილ იქნეს გადატანის მიმართ ინვარიანტულობა. პროექციული შეპირისპირება არ არის ინვარიანტული ბრუნვის მიმართ, მაგრამ მისი საშუალებით შესაძლებელია გამოსახულებების შედარება მობრუნების მცირე კუთხეების და სხვა უმნიშვნელო გეომეტრიული დამახინჯებების შემთხვევაში.

11.3.5. ესკიზური შეპირისპირება

ესკიზური შეპირისპირების სისტემებში მომხმარებელს აქვს შესაძლებლობა დახატოს გამოსახულების ესკიზური სურათი, რომელიც შეიცავს გამოსახულების ძირითად კონტურებს. ამის შემდეგ მონაცემთა ბაზაში ხდება მსგავსი კონტურების შემცველი გამოსახულებების მოძიება. ერთ-ერთ ასეთ სისტემაში, ART MUSEUM, ინახება ცნობილი სურათების ფერადი გამოსახულებები. გამოყენებამდე ხდება მათი წინასწარი დამუშავება შუალედურ მდგომარეობამდე, ე.წ. **აბსტრაქტულ გამოსახულებამდე** (abstract image) მიყვანის მიზნით. ოპერაციები, რომლებიც სრულდება წინასწარი დამუშავების ეტაპზე, შემდეგია:

- აფინური გარდაქმნის საშუალებით ხდება გამოსახულების შემცირება (მაგალითად, განზომილებამდე (64×64) პიქსელი). ხმაურის მოსაცილებლად გამოიყენება მედიანური ფილტრი. მიღებულ შედეგს **ნორმირებულ გამოსახულებას** უწოდებენ.
- სიკაშკაშის ვარდნების აღმოსაჩენად იყენებენ გრადიენტულ ოპერატორებს. ამ დროს სრულდება 2 ძირითადი მოქმედება: ჯერ გამოსახულებაზე გლობალური ზღურბ-

ლის საშუალებით, რომელიც შეირჩევა გრადიენტის საშუალოსა და დისპერსიის მიხედვით, ხდება გლობალური ვარდნების აღმოჩენა. ამის შემდეგ გლობალური ვარდნიდან ლოკალურად გამოთვლილი ზღურბლების გათვალისწინებით აირჩევა ლოკალური ვარდნები. ასეთი დამუშავების შედეგად მიიღება ხმაურისგან გაწმენდილი კონტურული გამოსახულება (refined edge image).

- მიღებულ კონტურულ გამოსახულებაზე სორციელდება კონტურების დავიწროება და ზედმეტი კონტურების მოშორება. შედეგად მიღებულ გამოსახულებას აბსტრაქტული გამოსახულება ეწოდება. ის წარმოადგენს ხმაურისგან თითქმის გაწმენდილ გამოსახულებას, რომელზედაც წარმოდგენილია საწყისი გამოსახულების ძირითადი კონტურები.

მას შემდეგ, რაც მომხმარებელი თავის მოთხოვნაში მიუთითებს გამოსახულების უხეშ ესკიზს, ამის შემდეგ ხდება მისი ზემოაღწერილი წინასწარი დამუშავება, კერძოდ, ზომების ნორმირება, ბინარიაზაცია, დავიწროება და ზედმეტი კონტურების მოშორება. შედეგად მიიღება გამოსახულება, რომელსაც ხაზოვანი ესკიზი ეწოდება (linear sketch).

ხაზოვანი ესკიზის მიღების შემდეგ უკვე შეიძლება მისი შეპირისპირება აბსტრაქტულ გამოსახულებებთან. შეპირისპირების ალგორითმს კორელაციური ხასიათი აქვს. ორივე გამოსახულება ბადეების საშუალებით ნაწევრდება უჯრედებად. მოთხოვნაში მოცემული გამოსახულების თითოეული უჯრედისთვის იანგარიშება ლოკალური კორელაცია მონაცემთა ბაზიდან აღებული გამოსახულების შესაბამის უჯრედთან. ამასთან, გამოთვლების საიმედოობის გაზრდის მიზნით ლოკალური კორელაცია გამოითვლება რამდენჯერმე მონაცემთა ბაზიდან

აღებული გამოსახულების უჯრედის მდებარეობის წანაცვლების რამდენიმე მნიშვნელობისთვის. ამ რამდენიმე წანაცვლებისთვის გამოთვლილი კორელაციებიდან აირჩევა მაქსიმალური. სწორედ ეს მნიშვნელობა აიღება როგორც შეპირისპირების შედეგი მოთხოვნაში მოცემული გამოსახულების და მონაცემთა ბაზიდან აღებული გამოსახულების ორ უჯრედს შორის. ყველა უჯრედზე ჩატარებული მსგავსი გამოთვლების შედეგად მიღებული ლოკალური კონტრასტის მაქსიმალური მნიშვნელობების ჯამი იძლევა გამოსახულებების მსგავსების მახასიათებლის მნიშვნელობას. გამოსახულებებს შორის მსგავსება ანუ მანძილის ზომა მახასიათებლის ამ მნიშვნელობის უკუპროპორციულია. ამგვარად, ზემოთქმულიდან გამომდინარე მანძილის ზომა ფორმალურად ასეთნაირად შეიძლება ჩაიწეროს

$$d_{sketch}(I, Q) = \frac{1}{\sum_g \max_n [d_{correlation}(shift_n(A^I(g), L^Q(g)))]}$$

სადაც, $A^I(g)$ აღნიშნავს აბსტრაქტული გამოსახულების ბადის g უჯრედს, გამოთვლილს მონაცემთა ბაზიდან აღებული I გამოსახულებისთვის, $shift(A^I(g))$ აღნიშნავს იგივე g უჯრედს იმავე აბსტრაქტულ გამოსახულებაზე წანაცვლების შემდეგ, ხოლო $L^Q(g)$ კი აღნიშნავს მოთხოვნაში მითითებული Q გამოსახულების ხაზოვანი ესკიზის ბადის g უჯრედს.

XII თავი. მოძრაობის აღმოჩენა დინამიკურ სცენებზე

ბუნებაში გარდა სტატიკური სცენებისა, ხშირად გვხვდება დინამიკური სცენებიც. მაგალითად, ვიდეოზე გადაღებული გამოცდილი სამთო მოთხილამურის მთიდან დაშვება შეიძლება გამოვიყენოთ როგორც სასწავლი ფილმი ახალბედა სამთო მოთხილამურის აღსაზრდელად. ასევე, რომელიმე დინამიკურ პროცესზე, მაგალითად, მცენარის ზრდის ან სხვა რომელიმე მოვლენის შესახებ დროის გარკვეულ ინტერვალებში გადაღებული გამოსახულებების მიმდევრობა საშუალებას იძლევა დააკვირდეთ ამ მოვლენას დინამიკაში. დინამიკური სცენის მაგალითს წარმოადგენს ტრანსპორტის მოძრაობა, რომელზედაც დაკვირვება შეიძლება განხორციელდეს უძრავი ან მოძრავი დამკვირვებლის მიერ. ამასთან, მიღებული გამოსახულებების მიმდევრობაზე არსებული ცვლილებები შეიძლება გამოყენებულ იქნეს როგორც მახასიათებელი ნიშან-თვისებები მოძრავი ობიექტების აღმოსაჩენად, ან მათი მოძრაობის ტრაექტორიების განსასაზღვრავად. მაგალითად, გამოსახულებაზე პიქსელების სიკაშკაშის ან ფერის ცვლილება, როგორც მახასიათებელი ნიშან-თვისება, შეიძლება სასარგებლო აღმოჩნდეს გამოსახულებაზე მოძრავი ობიექტების აღმოჩენისა და ამოცნობისთვის. ობიექტის მოძრაობის ანალიზი შესაძლებლობას იძლევა განისაზღვროს მისი ფორმა, სიჩქარე ან სხვა რომელიმე პარამეტრი. ასეთი ამოცანების ამოხსნა კი მოითხოვს გამოსახულებების მიმდევრობის (კადრების მიმდევრობის) სეგმენტაციის ავტომატურ საშუალებებს მათზე მნიშვნელოვანი მომენტების, მოვლენების ან სცენების გამოყოფის მიზნით. ჩვენ ძირითად ყურადღებას დაუთმობთ საკითხების შესწავლას, დაკავშირე-

ბუღს მოძრაობის აღმოჩენასთან ორგანზომილებიან გამოსახულებებსა და ვიდეომიმდევრობებზე, და ასეთი გამოსახულებების დამუშავებას მახასიათებელი ნიშან-თვისებების გამოყოფის მიზნით.

დინამიკურ სცენაზე ობიექტების მოძრაობის ანალიზის ამოცანა იყოფა ორი ეტაპად:

- მოძრაობის გაზომვა;
- მოძრაობის შესახებ მონაცემების გამოყენება სცენის სეგმენტაციისთვის და სამგანზომილებიანი ინფორმაციის გამოყოფა ობიექტების ფორმისა და მოძრაობის შესახებ.

12.1. მოძრაობის არსი გამოსახულებაზე

გამოსახულებების მიმდევრობაზე შეიძლება დავაკვირდეთ მრავალნაირ მოძრაობას. რეალურად თითოეულ მათგანთან დაკავშირებულია სხვადასხვა გამოყენებითი ამოცანა, რომელთა სპექტრი საკმაოდ ფართოა, დაწყებული უბრალოდ მოძრავი ობიექტის აღმოჩენიდან და დამთავრებული რამდენიმე მოძრავი ობიექტის ფარდობითი მოძრაობისა და ფორმის განსაზღვრით.

ზოგადად გამოყოფენ მოძრაობის ოთხ შესაძლო შემთხვევას (იგულისხმება, რომ ტერმინები **კამერა** და **დამკვირვებელი** სინონიმებია):

- უძრავი კამერა, ერთი მოძრავი ობიექტი, მუდმივი ფონი;
- უძრავი კამერა, რამდენიმე მოძრავი ობიექტი, მუდმივი ფონი;
- მოძრავი კამერა, შედარებით მუდმივი ფონი;
- მოძრავი კამერა, რამდენიმე მოძრავი ობიექტი.

უმარტივეს შემთხვევაში უძრავი კამერა აკვირდება შედარებით მუდმივ ფონს. ასეთ ფონზე ობიექტების მოძრაობა იწვევს გამოსახულებების შესაბამის არეებში პიქსელების სიკაშკაშის მნიშვნელობების და ფერის ცვლილებას. აღნიშნული პიქსელების საშუალებით შესაძლებელია მოძრავი ობიექტის ფორმის, მოძრაობის სიჩქარის და ტრაექტორიის განსაზღვრა. ასეთი ტიპის ვიდეოგადაღებები ფართოდ გამოიყენება დაცვით და გამშვებ სისტემებში; წარმოებაში, სადაც საჭიროა მუშა სივრცეში რომელიმე დეტალის არსებობის შესახებ ინფორმაცია; სატრანსპორტო ნაკადის კონტროლის საქმეში მოძრავი ავტომანქანების აღმოსაჩენად; და ა.შ. მოძრავი ობიექტების მიღევენება გარკვეული დროის განმავლობაში საშუალებას იძლევა განვსაზღვროთ მათი მოძრაობის ტრაექტორია და ასევე მივიღოთ ინფორმაცია მათი ქცევის შესახებ. ამის მაგალითს წარმოადგენს ადამიანთა თავშეყრის ადგილის კონტროლი ვიდეოკამერის საშუალებით. რამდენიმე კამერის შემთხვევაში კი ასეთი თვალთვალი საშუალებას იძლევა განვსაზღვროთ ობიექტის ტრაექტორია 3-განზომილებიან სივრცეში.

იმ შემთხვევაში, როცა ვიყენებთ მოძრავ კამერას, ის თავის მხრივ გამოსახულებაზე იწვევს გარკვეულ ცვლილებებს მაშინაც კი, თუ სცენა უძრავია. მოძრავი კამერის გამოყენებას აქვს გარკვეული მიზანი. კერძოდ:

- **იძლევა გარემომცველი სივრცის შესახებ უფრო მეტ ინფორმაციას;**
მაგალითად, პანორამული კამერის საშუალებით შეიძლება სცენის პანორამული-ფართოკუთხიანი ხედის მიღება.

- მიღებული მონაცემებით შესაძლებელია სცენაზე ობიექტთა განლაგების ფარდობითი სიღრმის გამოთვლა; ეს უკანასკნელი ემყარება იმას, რომ კამერასთან უფრო ახლოს მყოფი ობიექტების გამოსახულებები უფრო სწრაფად იცვლება, ვიდრე შორს მდგომი ობიექტების.
- რამდენიმე წერტილიდან დაკვირვება (ტრიანგულაციური გაზომვები) საშუალებას იძლევა მივიღოთ კამერასთან ახლოს მყოფი ობიექტების 3-განზომილებიანი ფორმის აღქმისა და გამოთვლისთვის საჭირო მონაცემები; ეს პროცესი მსგავსია ბინოკულარული სტერეოსისტემის გამოყენებით 3-განზომილებიანი ობიექტების ფორმის აღქმის პროცესის.

გამოსახულებების მიმდევრობის ანალიზი მოიცავს 3 ძირითად ნაბიჯს:

- მოძრავი ობიექტების აღმოჩენა;
- ასეთი ობიექტების მოძრაობის თვალთვალი კადრიდან კადრზე;
- ობიექტების ტრაექტორიების ანალიზი მათი ქცევის ამოცნობის მიზნით.

გამოსახულებების მიმდევრობების ანალიზი დინამიკური სცენების ამოცნობისა და გაანალიზების შესაძლებლობას იძლევა. კერძოდ, ისეთი სამეცნიერო და საინჟინრო ამოცანების შესწავლის შესაძლებლობას, როგორცაა სატრანსპორტო ნაკადები, ბიოლოგიური ზრდის პროცესები, დაწყებული მოლეკულური დონიდან და დამთავრებული ეკოსისტემის დონით,

სამრეწველო პროცესები, ავტონომიური სისტემები, რობოტები და სხვ.

12.2. გამოსახულებების გამოკლება

როგორც უკვე აღვნიშნეთ (I თავი) გამოსახულებების გამოკლება გამოიყენება როგორც ერთ-ერთი ხერხი მუდმივ ფონზე მოძრავი ობიექტების აღმოსაჩენად. ვთქვათ, 30 კადრი/წამში სიხშირის მქონე ვიდეოკამერის საშუალებით საწარმოში ვაკვირდებით კონვეიერის მოძრაობას, რომელიც გამოიყურება როგორც ერთგვაროვანი მუქი ფონი. თუ ამ დროს კამერის ხედვის არეში ხდება ღია ფერის ობიექტების გადაადგილება, მაშინ ამ ობიექტების წინა და უკანა კიდეები ერთი კადრის განმავლობაში სულ რამდენიმე პიქსელით გადაადგილდება. თუ ახლა დროის $t - 1$ მომენტში არსებულ I_{t-1} გამოსახულებას გამოვაკლებთ დროის t მომენტში არსებულ I_t გამოსახულებას, მაშინ მიღებულ გამოსახულებაზე მხოლოდ ამ პიქსელების მნიშვნელობები იქნება ნულისგან განსხვავებული და, შესაბამისად, კარგად შესამჩნევი. გამოსასვლელი პიქსელის მნიშვნელობის გამოთვლა შეიძლება ფორმულით

$$Q[i, j] = I_{t-1}[i, j] - I_t[i, j];$$

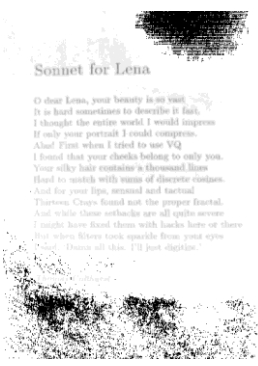
ან შესასვლელი პიქსელების მნიშვნელობათა სხვაობის აბსოლუტური სიდიდით

$$Q[i, j] = |I_{t-1}[i, j] - I_t[i, j]|.$$

იმ შემთხვევაში, თუ გვსურს გამოსახულებისთვის რაიმე C მუდმივას გამოკლება, მაშინ გამოსახველი პიქსელის მნიშვნელობა გამოითვლება ფორმულით

$$Q[i, j] = I_t[i, j] - C.$$

გამოკლების ოპერატორის მოქმედება იცვლება, თუ პიქსელის მნიშვნელობა უარყოფითი ხდება. ასეთ შემთხვევაში გამოიყენება სპეციალური გრაფიკული ფორმატები, რომლებშიც უარყოფითი მნიშვნელობების მქონე პიქსელების ასახვისათვის იყენებენ ე.წ. ფერების რუკას (colormap). იმ შემთხვევაში, თუ გამოსახულების ფორმატს არ შეუძლია ასეთი პიქსელების ასახვა, მაშინ მათ მნიშვნელობად იღებენ 0-ს (ანუ გამოსახულებაზე ისინი შავი ფერის ხდება), ან პირიქით, ხდება უარყოფითი მნიშვნელობის შებრუნება. მაგალითად, სიკაშკაშის მნიშვნელობა -30 ასეთი შებრუნების შემდეგ ტოლი ხდება 225-ის (255-30). ასეთი შემთხვევა ნაჩვენებია 12.1 სურათზე, სადაც (ა) საწყისი გამოსახულებაა, ხოლო (ბ) სიკაშკაშეების შებრუნების შემდეგ მიღებული გამოსახულება.



ა



ბ

სურ. 12.1
327

აღნიშნული პრობლემა არ წარმოიშობა იმ შემთხვევაში, თუ ოპერატორი იყენებს სხვაობის აბსოლუტურ მნიშვნელობას. ამის გამო ხშირად სწორედ ასეთ ოპერატორს იყენებენ. მაგალითისთვის განვიხილოთ ასეთი შემთხვევა: ვთქვათ, გვაქვს რომელიმე სცენის გამოსახულებების მიმდევრობა. ორ კადრს შორის სხვაობის აბსოლუტური მნიშვნელობა, როცა სცენაზე არაფერი მოძრაობს, გვაძლევს გამოსახულებას, რომელზედაც პიქსელების მნიშვნელობები ძირითადად ნულის ტოლი იქნება, მაგრამ თუ სცენაზე მოძრაობას ექნება ადგილი, მაშინ პიქსელები გამოსახულების სივრცის იმ ადგილებში, სადაც სიკაშკაშეების ცვლილებას კჰონდა ადგილი, მოგვცემს სხვაობების დიდ მნიშვნელობებს. მაგალითისათვის განვიხილოთ გამოსახულება (სურ. 12.2,ა), რომელზედაც წარმოდგენილია სხვადასხვა ხელსაწყოები. მსგავსი სცენის გამოსახულება, მხოლოდ მცირეოდენი განსხვავებებით წარმოდგენილია 12.2,ბ სურათზე. თუ გამოვთვლით ამ ორ კადრს შორის სხვაობის აბსოლუტურ მნიშვნელობას, მაშინ ის არეები, სადაც სიკაშკაშეების ცვლილება მოხდა, უფრო ღია ფერის გახდება (სურ. 12.2,გ).



ა



ბ

სურ. 12.2



ბ
სურ. 12.2

ამგვარად, გამოკლების შედეგად მიღებული გამოსახულება მუქია იმ ადგილებში, სადაც ნახევარტონების დონეების სივრცითი ცვლილება მცირეა. შესაბამისად, გამოსახულებაზე მოძრაობის დაფიქსირება შეგვიძლია გამოსახულების მხოლოდ იმ ადგილებში, სადაც ადგილი ჰქონდა სიკაშკაშის დონეების ცვლილებას. მიღებული შედეგი მიუთითებს სიკაშკაშის დონეების სივრცითი ცვლილების გადამწყვეტ როლზე მოძრაობის აღმოჩენის განსაზღვრაში.

12.3. სცენაზე წერტილთა გადაადგილების ვექტორების გამოთვლა

3D სცენაზე წერტილების გადაადგილება შესაბამისად იწვევს მათი გამოსახულებების გადაადგილებას 2D გამოსახულებაზე. 12.3 სურათზე მოცემულია წერტილების გამოსახულებების მასშტაბის ცვლილების და პანორამირების სამი შემთხვევა.

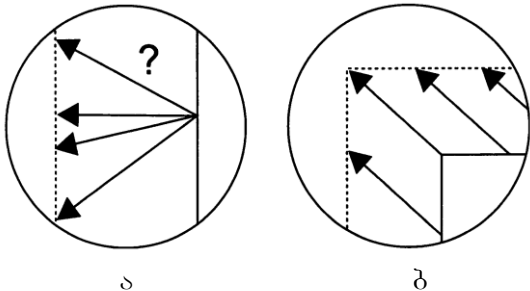


სურ. 12.3

სცენაზე არსებული წერტილების მასშტაბის შემცირებას (zoom out) ადგილი აქვს უძრავი კამერის ფოკუსური მანძილის შემცირებისას ან ფიქსირებული ფოკუსური მანძილის მქონე კამერის სცენის ობიექტებიდან დაშორების ზრდისას. ამ დროს გამოსახულებაზე ოპტიკურ ღერძზე არსებული წერტილის გამოსახულება არ გადაადგილდება. გამოსახულების ამ წერტილს **შეკრების ფოკუსი** (focus of contraction) ეწოდება. ობიექტების მასშტაბის გაზრდა (zoom in) ხდება უძრავი კამერის ფოკუსური მანძილის ზრდისას ან კამერის გადაადგილებისას სცენის იმ განსაკუთრებული წერტილის – ე.წ. **გაფართოების ფოკუსისკენ** (focus of expansion), რომელშიც გამოსახულების მასშტაბი არ იცვლება. ხოლო **კამერის პანორამირებას** (ჰგავს სცენის მიმართ დამკვირვებლის თავის მობრუნების ეფექტს) მიეყვაროთ სამგანზომილებიანი წერტილების გამოსახულებების გადაადგილებისკენ.

როგორც აღვნიშნეთ, მოძრაობა $2D$ გამოსახულებებზე მჭიდროდაა დაკავშირებული სიკაშკაშის დონეების სივრცით და დროით ცვლილებებთან. თითოეული ამ სიდიდის გამოთვლა შესაძლებელია სივრცითი და დროითი წარმოებულების გამოსათვლელი ლოკალური ოპერატორების საშუალებით. ასეთი

ოპერატორი “ხედავს” დაკვირვების ქვეშ მყოფი ობიექტის მხოლოდ იმ მცირე სეგმენტს, რომელიც სიდიდით მისი ნიღბის ზომის ტოლია. 12.4 სურათზე წარმოდგენილია ასეთი ნიღბის ანუ აპერტურის გამოსახულებაზე დადების ეფექტი.



სურ. 12.4

12.4,ა სურათზე ნაჩვენებია კონტური, რომელიც გამოსახულებების მიმდევრობის პირველი გამოსახულებიდან (უწყვეტი წირი) გადაადგილდა და მეორე გამოსახულებაზე აღმოჩნდა სხვა ადგილას (წყვეტილი წირი).

აღნიშნულ შემთხვევაში გადაადგილება პირველი გამოსახულებიდან მეორე გამოსახულებაზე შეიძლება აღიწეროს გადაადგილების ვექტორის საშუალებით. მაგრამ ამ შემთხვევაში არ შეგვიძლია, რომ გადაადგილება ცალსახად განვსაზღვროთ. საქმე ისაა, რომ გადაადგილების ვექტორს შეეძლო პირველი გამოსახულების კონტურზე არსებული რომელიმე წერტილი შეეერთებინა მეორე გამოსახულების კონტურის ნებისმიერ წერტილთან. ამ შემთხვევაში შეგვიძლია განვსაზღვროთ მხოლოდ გადაადგილების ვექტორის ის კომპონენტი, რომელიც კონტურის ნორმალის გასწვრივაა მიმართული, მაშინ როცა, კონტურის პარალელურად მიმართული კომპონენტი უცნობი

რჩება. ეს განუსაზღვრელობა ცნობილია როგორც **აპერტურის პრობლემა**.

გადაადგილების ვექტორის ცალსახა განსაზღვრა მაშინ არის შესაძლებელი, როცა ობიექტის კუთხე ოპერატორის ნიღბის შიგნით მდებარეობს (სურ. 12.4,ბ). აღნიშნული მიუთითებს, რომ ლოკალური ოპერატორების საშუალებით მოძრაობის შესახებ მხოლოდ არასრული ინფორმაციის მიღება შეიძლება.

განსაზღვრება 12.1. *გამოსახულებაზე 2D გადაადგილების ვექტორების 2D მახივს, რომელიც წარმოადგენს 3D სცენის წერტილების მოძრაობას (მოცემულს, როგორც ეს ნახვენება 12.3 სურათზე) ეწოდება მოძრაობის ველი.*

ამგვარად, გამოსახულებაზე გადაადგილების ვექტორები აღწერს სცენის 3D წერტილების გამოსახულებების გადაადგილებას. გადაადგილების ვექტორის აგება ასეთნაირად ხდება: გადაადგილების ვექტორის დასაწყისი იმყოფება 3D წერტილის დროის t მომენტში მიღებულ გამოსახულებაზე, ხოლო ბოლო – $(t + \delta)$ მომენტში მიღებულ იმავე წერტილის გამოსახულებაზე. ამგვარად, გადაადგილების თითოეული ვექტორი შეესაბამება წერტილის მყისიერ სინქარეს დროის t მომენტში.

განსაზღვრება 12.2. *გაფართოების ფოკუსი ეწოდება გამოსახულების იმ წერტილს, საიდანაც გამოდის მოძრაობის ველის ყველა გადაადგილების ვექტორი.*

ჩვეულებრივად, მას შეესაბამება 3D სცენის იმ წერტილის გამოსახულება, რომლისკენაც მოძრაობს კამერა.

განსაზღვრება 12.3. *გამოსახულების იმ წერტილს, საითკენაც იკრიბება მოძრაობის ველის ყველა გადაადგილების ვექტორი, ეწოდება შეკრების ფოკუსი.*

ეს უკანასკნელი წარმოადგენს 3D სცენის იმ წერტილის გამოსახულებას, რომელსაც კამერა შორდება.

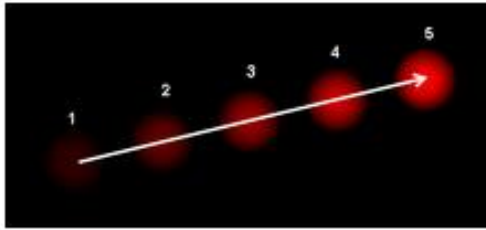
მოძრაობის ველის გამოყენება შეიძლება როგორც ობიექტების ამოცნობისთვის, ასევე მათი მოძრაობის ანალიზისთვის. უნდა აღინიშნოს, რომ გადაადგილების ვექტორების გამოთვლისას ორი შეზღუდვიდან აიღება ერთ-ერთი. კერძოდ,

- ან მიიღება დაშვება, რომ 3D სცენის P წერტილის და მისი რაღაც მიდამოს გამოსახულებაზე ინტენსივობა დროის იმ $[t_1, t_2]$ ინტერვალში, რომლის განმავლობაში ხდება ამ წერტილის გადაადგილების შეფასება, თითქმის მუდმივია;
- ან მიიღება დაშვება, რომ დროის $[t_1, t_2]$ ინტერვალის განმავლობაში ობიექტების კიდეების გამოსახულებების ინტენსივობებს შორის სხვაობა თითქმის მუდმივია.

12.4. ოპტიკური ნაკადი

როგორც ზემოთ აღვნიშნეთ, დროში ცვლადი გამოსახულებებიდან (გამოსახულებების მოწესრიგებული მიმდევრობიდან) შეიძლება ბევრი სასარგებლო ინფორმაციის მიღება. აღმოჩნდა, რომ ზოგიერთი ამოცანის ამოხსნა უფრო ადვილი ხდება, თუ გვაქვს გამოსახულებების დროითი მიმდევრობა. გამოსახულებების გამოკლების შესწავლისას აღვნიშნეთ, რომ გამოსახულებების მიმდევრობაზე ობიექტთა გადაადგილებასთან ერთად ადვილი აქვს ინტენსივობის სურათების გადაადგილებასაც.

სცენის გამოსახულებაზე ინტენსივობის სურათის მოხვედრით მოძრაობას, რომელსაც ადგილი აქვს სცენის ობიექტების მიმართ კამერის მოძრაობისას, **ოპტიკურ** ან **გამოსახულების ნაკადს** უწოდებენ. ოპტიკური ნაკადი (ან გამოსახულების სინქარე) ეს არის სინქარეთა ველი, რომელიც გამოსახულებების მიმდევრობის ერთ გამოსახულებას გარდაქმნის მეორე გამოსახულებაში. თუმცა, როგორც ასეთი, ის არ განისაზღვრება ცალსახად. გამოსახულებების სიბრტყეზე ის აღიქმება როგორც აშკარა მოძრაობა ვიზუალური აღქმის საფუძველზე და აქვს სინქარის განზომილება. თუ ოპტიკური ნაკადი განისაზღვრება ორი თანმიმდევრობითი გამოსახულებით, მაშინ ის წარმოინდება როგორც გადაადგილების ვექტორი ერთი გამოსახულების ნიშან-თვისებებიდან (ვთქვათ, ობიექტის კუთხე, სიმძიმის ცენტრი) მეორე გამოსახულების იგივე ნიშან-თვისებებამდე. 12.5 სურათზე წარმოდგენილია მოძრავი ობიექტის ოპტიკური ნაკადის ვექტორი ვიდეომიმდევრობაზე.



სურ. 12.5

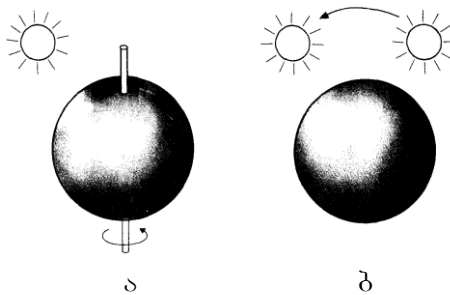
რაც შეეხება მოძრაობის ველს, ის სუფთა გეომეტრიული ცნებაა და ყოველგვარი ორპროექციის გარეშე წარმოადგენს 3D ფიზიკური მოძრაობის ველის (გადაადგილების ვექტორების) პროექციას 2D გამოსახულებაზე. სხვა სიტყვებით, მოძრაობის ველი არის ობიექტის რეალური მოძრაობა პროექცირებუ-

ლი გამოსახულების სიბრტყეზე. ამასთან, მნიშვნელოვანია აღინიშნოს, რომ ოპტიკური ნაკადი წარმოადგენს უწყვეტი სივრცისთვის დამახასიათებელ ცნებას, მაშინ, როდესაც მოძრაობის ველი მის დისკრეტულ ანალოგს წარმოადგენს.

განსაზღვრება 12.4. *ოპტიკური ნაკადი არის მოძრაობის ველი, გამოთვლილი იმ დაშვებით, რომ გამოსახულების ინტენსივობა შესაბამისი წერტილების სიახლოვეში შედარებით მუდმივია.*

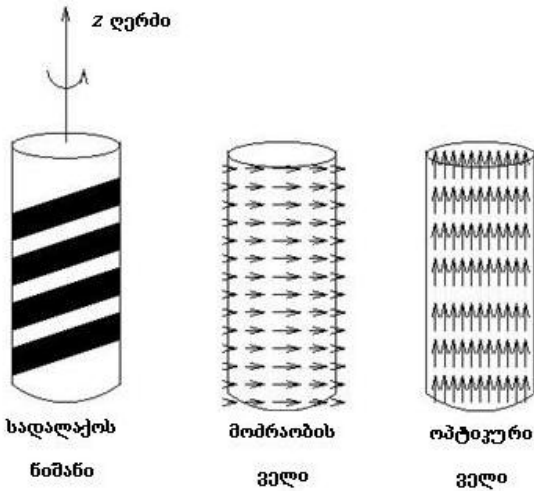
იდეალურ შემთხვევაში არსებობს შესაბამისობა ოპტიკურ ნაკადსა და მოძრაობის ველს შორის, თუმცა ეს ყოველთვის ასე არ არის. განვიხილოთ მაგალითი.

ვთქვათ, გვაქვს უძრავი ობიექტის წინ მბრუნავი იდეალური ერთგვაროვანი სფერო (სურ. 12.6,ა). მისი ზედაპირის მრუდწირულობის გამო მასზე ადგილი აქვს ინტენსივობის სურათის ცვლილებას. თუმცა, ამ დროს ინტენსივობის სურათი ზედაპირთან ერთად არ მოძრაობს და არ ხდება გამოსახულების დროში ცვლილება. აღნიშნულ შემთხვევაში ოპტიკური ნაკადი ყველგან ნულის ტოლია, ხოლო მოძრაობის ველი კი ნულისგან განსხვავებულია.



სურ. 12.6

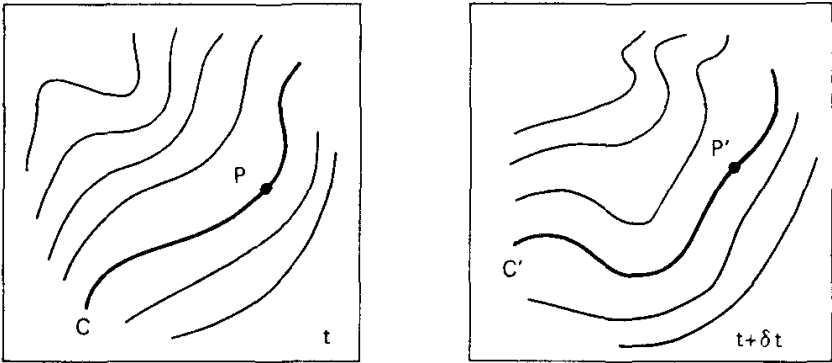
ახლა განვიხილოთ იგივე სფერო, მხოლოდ უძრავი, რომელიც ნათდება სინათლის მოძრავი წყაროს საშუალებით (სურ. 12.6,ბ). შევამჩნევთ, რომ სინათლის წყაროს მოძრაობასთან ერთად იცვლება სფეროს ზედაპირის ინტენსივობის სურათიც, ანუ სფეროს ზედაპირზე ხდება სიკაშკაშის მნიშვნელობების ცვლილებები. ასეთ შემთხვევაში, ოპტიკური ნაკადი ნულისგან განსხვავებული უნდა იყოს, თუმცა მოძრაობის ველი ყველგან ნულის ტოლია. 12.7 სურათზე წარმოდგენილია კიდევ ერთი მაგალითი, როცა ვერტიკალური z ღერძის მიმართ მბრუნავი ობიექტის მოძრაობის ველი და ოპტიკური ნაკადი ერთმანეთს არ ემთხვევა. ოპტიკური ნაკადი და მოძრაობის ველი ასევე არ ემთხვევა ერთმანეთს, როცა გვაქვს მაგალითები სცენებისა სარკული არეკვლებით და ჩრდილებით.



სურ. 12.7

ახლა ვნახოთ, თუ რა იგულისხმება ინტენსივობის სურათის მოჩვენებითი მოძრაობის ქვეშ. განვიხილოთ გამოსახულებები

ბაზე P წერტილი, რომელშიც ინტენსივობის მნიშვნელობა დროის t მომენტში E -ს ტოლია (სურ. 12.8).



სურ. 12.8

დავსვათ კითხვა: გამოსახულების რომელი P' წერტილი იქნება მოცემული წერტილის შესაბამისი დროის $(t + \delta)$ მომენტის შესაბამის მეორე გამოსახულებაზე, ან სხვანაირად, როგორ მოძრაობს დროის ამ ინტერვალში ინტენსივობის სურათი. ჩვეულებრივ, P წერტილის მიდამოში უნდა იყოს E ინტენსივობის მნიშვნელობის მქონე წერტილების საკმაო რაოდენობა. თუ გამოსახულების ამ ნაწილში ინტენსივობის ცვლილება უწყვეტია, მაშინ P წერტილი მოთავსებული იქნება ინტენსივობის ერთნაირი მნიშვნელობების მქონე C წირზე. დროის $(t + \delta)$ მომენტში კი მისი შესაბამისი იქნება ინტენსივობის ერთნაირი, E მნიშვნელობის მქონე C' წირი. აქ საინტერესო საკითხს წარმოადგენს იმის განსაზღვრა, თუ რა შესაბამისობა არსებობს C და C' წირების წერტილებს შორის. მაგრამ ამ საკითხზე პასუხის გაცემა რთულია, რადგანაც მოძრაობის დროს წირის ფორმა იცვლება.

ამგვარად, ცვლად გამოსახულებაზე ოპტიკური ნაკადის ცვლილების ცალსახად განსაზღვრა ლოკალურ ინფორმაციაზე დაყრდნობით შეუძლებელია.

12.5. ოპტიკური ნაკადის განტოლება

ცნობილია ოპტიკური ნაკადის გამოთვლის მეთოდები გამოსახულების არა მარტო ინტერესის წერტილებში (პიქსელებში), არამედ გამოსახულების ყველა პიქსელში. ჩვენ განვიხილავთ კლასიკურ მეთოდს, რომელიც ერთობლივად ითვალისწინებს როგორც სივრცით, ასევე დროით გრადიენტებს, გამოთვლილს არანაკლებ ორი თანმიმდევრობითი კადრისთვის. ამ მეთოდებს **დიფერენციალური მეთოდები** ეწოდება. 12.9 სურათზე წარმოდგენილია კამერის ხედვის არეში მოძრავი ობიექტის გამოსახულებები (ფონის ინტენსივობის მნიშვნელობა არის 3, ხოლო ობიექტის – 9).

3333333333	3333333333
3333333333	3333333333
3333333333	3373333333
3373333333	3397533333
3397533333	3399753333
3399753333	3399975333
3399975333	3333333333
3333333333	3333333333

ა

ბ

სურ. 12.9

12.9,ა სურათზე, რომელიც შეესაბამება დროის t_1 მომენტს, მოცემულია გამოსახულების ქვედა მარცხენა კუთხეში მოთავსებული სამკუთხა ობიექტი, ხოლო 12.9,ბ სურათზე კი – მეორე

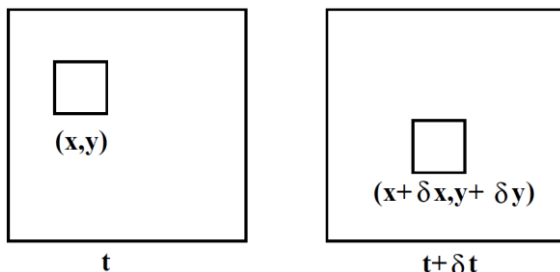
გამოსახულება დროის t_2 მომენტში, მხოლოდ ზემოთ, ერთი პიქსელით წანაცვლებული იგივე ობიექტი. წარმოდგენილი მაგალითი ხსნის ზოგიერთ იმ დაშვებას, რომლებიც მხედველობაში უნდა იქნეს მიღებული ოპტიკური ნაკადის მათემატიკური მოდელის აგებისას. კერძოდ, დაიშვება, რომ:

- ობიექტის არეკვლის თვისებები და მისი განათებულობა დროის $[t_1, t_2]$ ინტერვალში უცვლელია;
- დროის აღნიშნულ ინტერვალში მანძილის ცვლილება ობიექტიდან კამერამდე ან სინათლის წყაროებამდე უმნიშვნელოა;
- გამოსახულების (დროის t_1 მომენტის შესაბამისი) ნებისმიერი მცირე $N_{x,y}$ მიდამო შეიძლება ნაპოვნი იქნეს მეორე გამოსახულებაზე გარკვეულ წანაცვლებულ $N_{x+\delta x, y+\delta y}$ მდებარეობაში დროის t_2 მომენტში.

მართალია, რეალურად ეს დაშვებები არ სრულდება, მაგრამ ზოგჯერ ისინი ოპტიკური ნაკადის ვექტორების გამოთვლის შესაძლებლობას იძლევა.

გამოვიყენოთ ახლა მიღებული დაშვებები ოპტიკური ნაკადის განტოლების მისაღებად და ვუჩვენოთ, თუ როგორ შეიძლება მისი საშუალებით ოპტიკური ნაკადის ვექტორების გამოთვლა.

ვთქვათ, გამოსახულების $n \times n$ მიდამოს ცენტრალური პიქსელი, რომელიც აღიწერება უწყვეტი ინტენსივობის $f(x, y, t)$ ფუნქციით დროის δt ინტერვალში გადაადგილდა $(\delta x, \delta y)$ მანძილით და აღმოჩნდა წერტილში $(x + \delta x, y + \delta y)$ ინტენსივობით $f(x + \delta x, y + \delta y, t + \delta t)$ (სურ. 12.10).



სურ. 12.10

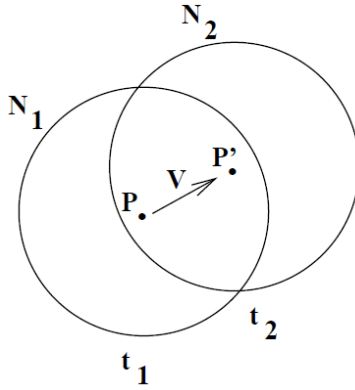
მაშინ, ზემომოყვანილი დაშვებების გათვალისწინებით, მოცემულ წერტილში ინტენსივობის მნიშვნელობებისთვის შეიძლება დაიწეროს

$$f(x, y, t) = f(x + \delta x, y + \delta y, t + \delta t).$$

ინტენსივობის ფუნქციის ტეილორის მწკრივად გაშლით და იმ დაშვების გათვალისწინებით, რომ პიქსელის წანაცვლების სიდიდე მცირეა, გვექნება

$$f(x + \delta x, y + \delta y, t + \delta t) \approx f(x, y, t) + \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial t} \delta t \quad (12.1).$$

(x, y, t) წერტილის $n \times n$ მიდამოს სიმცირის გამო (12.1) გამოსახულებაში უგულებელყოფილია მაღალი რიგის წევრები (მათი სიმცირის გამო) და განიხილება მხოლოდ წრფივი წევრები. აღსანიშნავია, რომ მოცემული გამოსახულება წარმოადგენს ერთი ცვლადის ფუნქციის შემთხვევის – $f(x + \delta x) \approx f(x) + f'(x)\delta x$ განზოგადებას რამდენიმე ცვლადის შემთხვევაზე. ოპტიკური ნაკადის განტოლების მიღების შემდეგი ნაბიჯისთვის განვიხილოთ 12.11 სურათზე წარმოდგენილი შემთხვევა.



სურ. 12.11

ობტიკური ნაკადის ვექტორი $\mathbf{V} = [\delta x, \delta y]$ წარმოადგენს ვექტორს, რომლისთვისაც უნდა მოიძებნოს დროის t_1 მომენტში N_1 გამოსახულების (x, y) მიდამოს წანაცვლება N_2 გამოსახულების ისეთივე ინტენსივობის მქონე $(x + \delta x, y + \delta y)$ მიდამოში დროის t_2 მომენტში. ეს ნიშნავს, რომ

$$f(x + \delta x, y + \delta y, t + \delta t) = f(x, y, t) \quad (12.2),$$

ანუ \mathbf{V} მიმართულებით მოძრაობისას N_2 გამოსახულების მიდამოს ინტენსივობა დროის t_2 მომენტში ტოლია N_1 გამოსახულების მიდამოს ინტენსივობისა დროის t_1 მომენტში.

თუ მოვახდენთ (12.1) და (12.2) ტოლობების გაერთიანებას, გვექნება

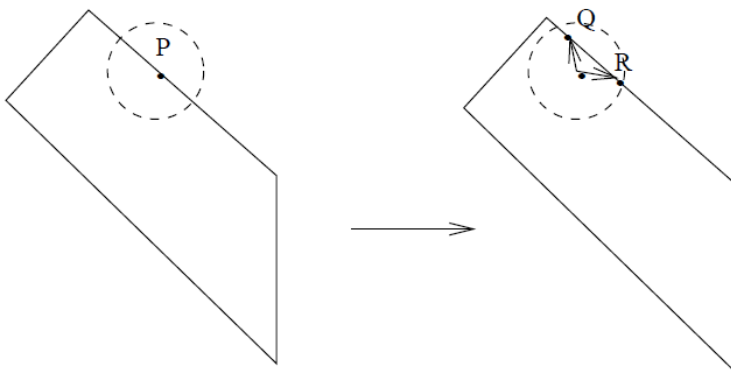
$$\frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial t} \delta t = 0.$$

იმ დაშვების გათვალისწინებით, რომ (x, y) მიდამოში ინტენსივობის ცვლილება არ ხდება, აღნიშნული განტოლებიდან მივიღებთ ოპტიკური ნაკადის განტოლებას

$$-\frac{\partial f}{\partial t} \delta t = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \circ [\delta x, \delta y] = \nabla f \circ [\delta x, \delta y],$$

რომელიც გვეუბნება, რომ გამოსახულებების მიმდევრობის შემთხვევაში ინტენსივობის ცვლილება გამოსახულების მოცემულ (x, y) წერტილში დროის t მომენტში ტოლია ამ წერტილში ინტენსივობის სივრცითი ∇f გრადიენტის სკალარული ნამრავლისა ოპტიკური ნაკადის \mathbf{V} ვექტორზე.

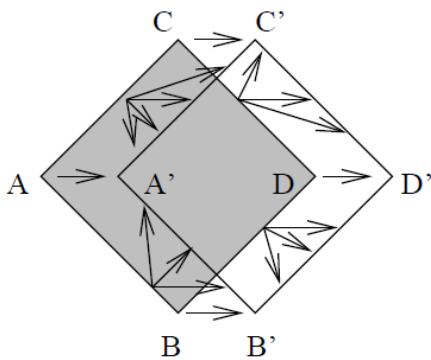
ამ განტოლების ამონახსნი არ იძლევა ნაკადის \mathbf{V} ვექტორის ზუსტად (ცალსახად) განსაზღვრის საშუალებას, მაგრამ მას ადებს წრფივ შეზღუდვას. გამოსახულებაზე შეიძლება იყოს ბევრი ისეთი N_2 მიდამო, რომელშიც ინტენსივობა ემთხვევა N_1 მიდამოს ინტენსივობას. 12.12 სურათზე წარმოდგენილ შემთხვევაში სწორხაზოვანი კიდეების მქონე ობიექტი დროის $[t_1, t_2]$ ინტერვალში წაინაცვლებს მარჯვნივ.



სურ. 12.12

სურათიდან ჩანს, რომ მცირე მიდამოს (აპერტურის) ცენტრით (x, y) წერტილში, შემთხვევაში შეიძლება არსებობდეს ნაკადის რამდენიმე შესაძლო ვექტორი. თუ განვიხილავთ მცირე აპერტურას, ცენტრით P წერტილში, მაშინ P წერტილის ახალ მდებარეობად შესაძლებელია განვიხილოდ კიდის QR მონაკვეთის გასწვრივ განლაგებული წერტილები R, Q ან რომელიმე სხვა.

12.13 სურათზე ნახვენებია მარჯვნივ მოძრავი კვადრატული ობიექტი განსხვავებული ოთხი კილით.



სურ. 12.13

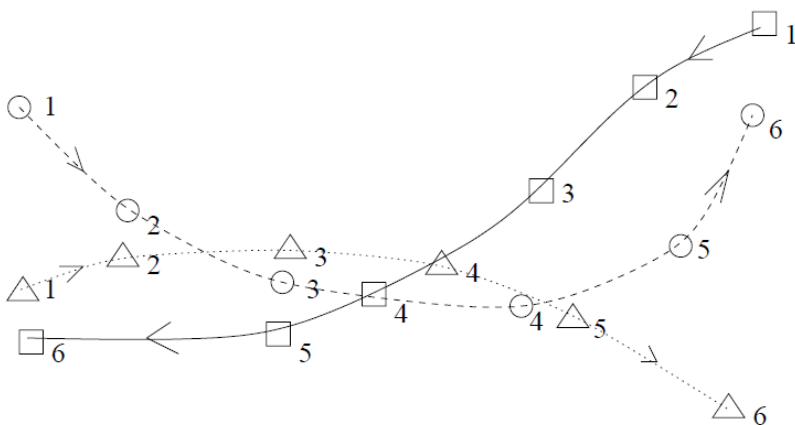
გადაადგილების ვექტორებსა და კიდეებზე მათ საწყის წერტილებზე დროის t_1 მომენტში დადებულია წრფივი შეზღუდვა, რომლის მიხედვითაც ვექტორების ბოლოები დროის t_2 მომენტში ასევე კიდეებზე უნდა იმყოფებოდეს. ერთნაირი შეზღუდვები მოცემული ობიექტის ყველა წვეროს მიმართ გვაძლევს უფლებას ვთქვათ მისი მარჯვნივ წანაცვლების შესახებ. იმის გამო, რომ ობიექტი მყარია, ასეთი ქცევა დამახასიათებელი იქნება ობიექტის საზღვრის გასწვრივ მოთავსებული ყველა წერტილისთვის.

ზოგად შემთხვევაში, შეიძლება ობიექტს არც ჰქონდეს მკვეთრად გამოხატული სწორხაზოვანი კიდები, მაგრამ მიუხედავად ამისა, 12.11 სურათი ისეთივე წარმატებით შეიძლება იქნეს გამოყენებული თანაბარი ინტენსივობის მქონე მრუდწირული კონტურებისთვისაც. ასეთ შემთხვევაში, სურათზე ნახვენები კიდები უნდა განვიხილოთ როგორც მხებები იმ კონტურის მიმართ, რომლის ლოკალური აპროქსიმაციისთვისაც ისინი გამოდგება.

რაც შეეხება 12.12 სურათს, ის შეიძლება შემდეგნაირად იყოს ინტერპრეტირებული: გამოსახულების P წერტილში ინტენსივობის ცვლილება შეიძლება ასე ჩაიწეროს $-\frac{\partial f}{\partial t} dt$. ოპტიკური ნაკადის განტოლების მიხედვით ინტენსივობის ეს ცვლილება ტოლია ინტენსივობის სივრცითი ∇f გრადიენტის და ოპტიკური ნაკადის \mathbf{V} ვექტორის სკალარული ნამრავლისა. ოპტიკური ნაკადის ვექტორის მოდული $|\mathbf{V}|$ სიდიდით შეიძლება ტოლი იყოს ნორმალის გასწვრივ ინტენსივობის ვარდნის ახალ მდებარეობამდე მანძილისა, ან შეიძლება აღმოჩნდეს გაცილებით უფრო დიდი, თუ ნაკადის ვექტორის მიმართულება მკვეთრად განსხვავდება სივრცითი გრადიენტის მიმართულებისგან. ამ უკანასკნელის მაგალითად შეიძლება მოვიყვანოთ თოკი, რომელსაც ძალიან სწრაფად ეწევიან ვერტიკალურად ზემოთ და რომელიც ამავე დროს ცოტათი ვიბრირებს ჰორიზონტალური მიმართულებითაც. მაშინ თოკის ბოლოები გამოსახულებაზე უმნიშვნელოდ წაინაცვლებს.

12.6. მოძრავი ობიექტების თვალთვალი ვიდეომიმდევრობებზე

ობიექტის თვალთვალი (object tracking) წარმოადგენს დინამიკურ სცენაზე მოძრავი ობიექტის (ან ობიექტების) აღმოჩენის (ლოკალიზაციის) პროცესს დროის გარკვეული პერიოდის განმავლობაში. ეს პროცესი გულისხმობს დროის სხვადასხვა მომენტებში სივრცეში მოძრავი ობიექტის მდებარეობის წერტილების დაგროვებას და შემდეგ მათ საფუძველზე მოძრაობის ტრაექტორიის აგებას. 12.14 სურათზე ნაჩვენებია სამი ობიექტის გლუვი ტრაექტორიები, აგებული დროის სხვადასხვა მომენტის შესაბამისი 6 წერტილის საშუალებით.



სურ. 12.14

ზოგადი დაშვებები, რომლებიც გათვალისწინებული უნდა იქონოს ვიზიკური ობიექტების 3D სივრცეში მოძრაობისას, შემდეგია:

1. მოძრავი ფიზიკური ობიექტის ადგილმდებარეობა დროის განმავლობაში იცვლება გლუვად;
2. მოძრავი ფიზიკური ობიექტის სინქარე დროის განმავლობაში იცვლება გლუვად როგორც სიდიდით, ასევე მიმართულებით;
3. დროის ყოველ მომენტში ობიექტი შეიძლება იმყოფებოდეს სივრცის მხოლოდ ერთ არეში;
4. ორი სხვადასხვა ობიექტი დროის ერთსა და იმავე მომენტში არ შეიძლება იმყოფებოდეს სივრცის ერთსა და იმავე არეში.

პირველი სამი დაშვება სრულდება 3D სივრცის 2D პროექციებისთვისაც. საქმე ისაა, რომ ობიექტების 3D სივრცეში უწყვეტი მოძრაობისას გამოსახულებებზე (გამოსახულებების მიმდევრობაზე) ჩანს მათი გამოსახულებების უწყვეტი ტრაექტორიები. რაც შეეხება მეოთხე დაშვებას, ის შეიძლება დაირღვეს იმის გამო, რომ კამერის წინ ობიექტები შეიძლება ერთმანეთს ფარავდეს. ასეთი პრობლემა წარმოიშობა მაშინ, როცა კამერა ერთ ადგილზეა ფიქსირებული. ადამიანების მონაწილეობით ჩატარებულმა ექსპერიმენტებმა აჩვენა, რომ მათ შეუძლიათ ამოიცნონ ობიექტები და მოახდინონ მოძრაობის ანალიზი კადრების მიმდევრობაზე დაყრდნობით იმ შემთხვევაში, თუ ეს უკანასკნელები შეიცავს მოძრავი ობიექტების მხოლოდ ცალკეულ წერტილებს.

განვიხილოთ ახლა ალგორითმი, რომელშიც ზემოჩამოთვლილი დაშვებები გამოიყენება კადრების მიმდევრობაზე გამოყოფილ წერტილებზე გამავალი გლუვი ტრაექტორიების სიმრავლის გამოსათვლელად. ჯერ შემოვიტანოთ გლუვი ტრაექტორიის მათემატიკური განსაზღვრება, ხოლო შემდეგ კი m ტრაექტორიების გლუვი სიმრავლის, როგორც ისეთი ცალსახა

ტრაექტორიების სიმრავლის განსაზღვრა, რომელთათვისაც “გლუვობის” m მნიშვნელობების ჯამი ოპტიმალურია.

განსაზღვრება 12.5. თუ i -ური ობიექტის თვალთვალის ხორციელდება დროის $t = 1, 2, \dots, n$ მომენტებში გამოსახულებების $p_{i,1}, p_{i,2}, \dots, p_{i,t}, \dots, p_{i,n}$ წერტილებში, მაშინ ამ წერტილების $T_i = (p_{i,1}, p_{i,2}, \dots, p_{i,t}, \dots, p_{i,n})$ მიმდევრობას i -ური ობიექტის ტრაექტორია ეწოდება.

აქ $p_{i,n}$ წარმოადგენს T_i -ური ტრაექტორიის წერტილს მე- n -ე კადრზე.

ტრაექტორიის ნებისმიერი ორი წერტილისთვის შეიძლება განისაზღვროს სხვაობის ვექტორი

$$V_{i,t} = p_{i,t+1} - p_{i,t}.$$

მოდრავი ობიექტის ტრაექტორიის $p_{i,t}$ წერტილში გლუვობის სიდიდე $S_{i,t}$ გამოითვლება ტრაექტორიის ამ წერტილში შემავალი და გამომავალი ვექტორების სხვაობის საფუძველზე. რაც შეეხება გლუვობის მიმართულებას, ის გამოითვლება ამ ვექტორების სკალარული ნამრავლით, ხოლო სიჩქარის გლუვობა კი გამოითვლება ამ ვექტორების სიგრძეების გეომეტრიული და არითმეტიკული საშუალოების შედარების საფუძველზე. ამგვარად, ტრაექტორიის $p_{i,t}$ წერტილში გლუვობის სიდიდე ტოლია

$$S_{i,t} = w \left(\frac{V_{i,t-1} \circ V_{i,t}}{|V_{i,t-1}| |V_{i,t}|} \right) + (1 - w) \left(\frac{2 \sqrt{|V_{i,t-1}| |V_{i,t}|}}{|V_{i,t-1}| + |V_{i,t}|} \right) \quad (12.3),$$

სადაც, w წონითი კოეფიციენტია და $0 \leq w \leq 1$, შესაბამისად გლუვობის სიდიდე $0 \leq S_{i,t} \leq 1$. აღნიშნული გამოსახულების პირველი წევრი შეესაბამება გლუვობის მიმართულებას, ხოლო მეორე – სიჩქარის გლუვობას. თუ ტრაექტორია წრფივია და წერტილები ერთმანეთისგან თანაბარი მანძილებითაა დაშორებული, მაშინ (12.3) განტოლების მიხედვით ყველა სხვაობის ვექტორი ერთმანეთის ტოლია და შესაბამისად $S_{i,t} = 1.0$, რაც წერტილში გლუვობის ოპტიმალურ მნიშვნელობას წარმოადგენს. სიჩქარის ან მიმართულების ცვლილება ამცირებს $S_{i,t}$ სიდიდეს.

დავუშვათ, გვაქვს n კადრისგან შედგენილი გამოსახულებების მიმდევრობა და თითოეულ მათგანზე m რაოდენობის მოძრავი წერტილი. მოვნიშნოთ ეს წერტილები პირველ კადრზე ჭდეებით $i = 1, 2, \dots, m$. ამოცანა მდგომარეობს შემდეგში: ვიპოვოთ მოძრავი m წერტილების T_i ტრაექტორიები კადრების n მიმდევრობაზე მინიმალური ჯამური გლუვობის მნიშვნელობით. ცხადია, ამ შემთხვევაში გვექნება m ტრაექტორია, რომელთაგან თითოეულზე იქნება n რაოდენობის წერტილი.

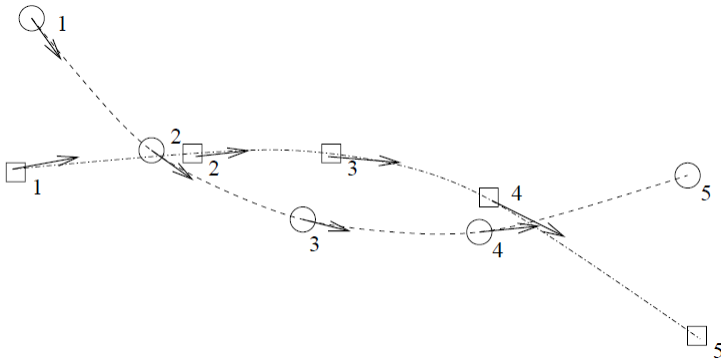
ჯამური გლუვობა T_s განისაზღვრება როგორც ყველა m ტრაექტორიაზე გამოყოფილ წერტილებში გლუვობის სიდიდეების მთლიანი ჯამი

$$T_s = \sum_{i=1}^m \sum_{t=2}^{n-1} S_{i,t} .$$

ზემოთქმულიდან გამომდინარე, ტრაექტორიების მთლიანი შესაძლო რაოდენობა n კადრზე ტოლია m^n , საიდანაც არჩეული უნდა იქნეს მხოლოდ m რაოდენობის ტრაექტორია. შესაბამისად, უნდა არსებობდეს ამ ტრაექტორიების გონივრულად არჩევის მეთოდი. იმის გამო, რომ თითოეული ტრაექტორია

გლუვია, აზრიანი იქნებოდა მიდგომა, რომლის მიხედვითაც ყველა შესაძლო ტრაექტორიების სიმრავლიდან მოხდებოდა იმ ტრაექტორიების არჩევა, რომლებიც მოახდენდა ტრაექტორიების გლუვობის, ანუ ჯამური T_s გლუვობის მაქსიმიზაციას. აქედან გამომდინარე, ტრაექტორიების პოვნის მეთოდი შეიძლება ფორმულირებულ იქნეს როგორც ოპტიმიზაციის ამოცანა. ამავე დროს, გამოკვლევებით დადგენილია, რომ ორი კადრის შემთხვევაში გლუვობის გამოყენებას მოძრაობის დაფიქსირებისთვის აზრი არ აქვს – მინიმუმ სამი კადრი მაინცაა საჭირო ობიექტის მდებარეობის ცვლილებების გამოსათვლელად. ცხადია, მოცემული ამოცანის გადასაწყვეტად, ანუ ზუსტი ტრაექტორიების საპოვნელად შეიძლება განხილული იქნეს ტრაექტორიების ყველა შესაძლო კომბინაცია. მართალია, იმის გამო, რომ $t = 1$ კადრზე საწყისი წერტილების არჩევა ხდება ნებისმიერად, შესაბამისად, ტრაექტორიების რაოდენობა m^{n-1} -ის ტოლი იქნება. მაგრამ ასეთი გამოთვლების ფასი ძალიან მაღალია. ამიტომ ალგორითმში ახალი წერტილების მიკუთვნება შესაბამისი ტრაექტორიებისადმი ხორციელდება იტერაციული ოპტიმიზაციის ალგორითმის გამოყენებით. კერძოდ, ეს ალგორითმი აგრძელებს ტრაექტორიებს $(t + 1)$ კადრამდე (ანუ უმატებს მათ $(t + 1)$ კადრის წერტილებს) იმ დაშვებით, რომ ტრაექტორიები t კადრამდე უკვე მიღებულია, ხოლო $(t + 1)$ კადრის წერტილების მიკუთვნება უკვე არსებული ტრაექტორიებისადმი ხორციელდება უახლოესი მეზობლის ძიების კრიტერიუმის გამოყენებით. ამასთან, ყოველ მომდევნო $(t + 1)$ კადრზე ტრაექტორიებისთვის ახლადმიერთებული წერტილების შესაბამისობა მოწმდება ჯამური გლუვობის მაქსიმიზაციის კრიტერიუმით. აღნიშნულ თვალთვალის ალგორითმს **ხარბი გაცვლის ალგორითმი** (greedy exchange algorithm) ეწოდება.

განვიხილოთ ამ ალგორითმის მოქმედება 12.15 სურათზე მოცემული მაგალითისთვის, სადაც წარმოდგენილია ორი, \bigcirc და \square ობიექტის ტრაექტორია და მათთან ერთად ოპტიკური ნაკადის ვექტორები პირველ ხუთ წერტილში.



სურ. 12.15

12.1 ცხრილში კი მოცემულია წერტილების კოორდინატები გამოსახულების სიბრტყეზე და გლუვობის სიდიდეები ტრაექტორიების ზოგიერთი წერტილისთვის.

ცხ. 12.1

t=1	t=2	t=3	t=4	t=5	smoothness
$\bigcirc_1(112\ 262)$	$\square_2(206\ 185)$	$\bigcirc_3(250\ 137)$			0.97
$\square_1(106\ 175)$	$\bigcirc_2(180\ 188)$	$\square_3(280\ 185)$			0.98
$\bigcirc_1(112\ 262)$	$\bigcirc_2(180\ 188)$	$\bigcirc_3(250\ 137)$			0.99
$\square_1(106\ 175)$	$\square_2(206\ 185)$	$\square_3(280\ 185)$			0.99
$\bigcirc_1(112\ 262)$	$\bigcirc_2(180\ 188)$	$\bigcirc_3(250\ 137)$	$\bigcirc_4(360\ 137)$		1.89
$\square_1(106\ 175)$	$\square_2(206\ 185)$	$\square_3(280\ 185)$	$\square_4(365\ 156)$		1.96
$\bigcirc_1(112\ 262)$	$\bigcirc_2(180\ 188)$	$\bigcirc_3(250\ 137)$	$\bigcirc_4(360\ 137)$	$\square_5(482\ 80)$	2.84
$\square_1(106\ 175)$	$\square_2(206\ 185)$	$\square_3(280\ 185)$	$\square_4(365\ 156)$	$\bigcirc_5(478\ 170)$	2.91
$\bigcirc_1(112\ 262)$	$\bigcirc_2(180\ 188)$	$\bigcirc_3(250\ 137)$	$\bigcirc_4(360\ 137)$	$\bigcirc_5(478\ 170)$	2.89
$\square_1(106\ 175)$	$\square_2(206\ 185)$	$\square_3(280\ 185)$	$\square_4(365\ 156)$	$\square_5(482\ 80)$	2.94

პირველ კადრზე წერტილებს ჭდეები შეიძლება მიენიჭოს ნებისმიერად, მაგალითად, ობიექტს $\square_1 \equiv 1 = T[1,1]$ და ობიექტს $O_1 \equiv 2 = T[2,1]$. ამის შემდეგ ტრაექტორიები შეიძლება გაგრძელდეს მომდევნო კადრების უახლოესი წერტილისკენ: უახლოეს წერტილს მიენიჭება ჭდე $T[1,2] = O_2$ და დარჩენილს კი – $T[2,2] = \square_2$. მაგრამ ჭდეების მინიჭებისას შევცდით ნამდვილი ტრაექტორიების არჩევაში. მართლაც, გამოვთვალოთ ჯამური გლუვობა ორივე ტრაექტორიისთვის მას შემდეგ, რაც ავირჩევთ უახლოეს მეზობელს დროის $t = 3$ მომენტში (მესამე კადრზე). ცხრილის 12.1. პირველი ორი სტრიქონიდან ჩანს, რომ აღნიშნული ჭდეების მინიჭების შედეგად ტრაექტორიების ჯამური გლუვობა ტოლია $0.97 + 0.98 = 1.95$. მაგრამ თუ ახლა წინა კადრზე ($t = 2$) მინიჭებულ ჭდეებს ($T[1,2] = O_2$ და $T[2,2] = \square_2$) ადგილებს შევუცვლით, მაშინ ჯამური გლუვობისთვის მივიღებთ უკეთეს მნიშვნელობას, $0.99 + 0.99 = 1.98$. ასეთი ცვლილების შედეგად დროის $t = 2$ მომენტში ტრაექტორიებს ექნება სახე, (\square_1, \square_2) და (O_1, O_2) . შემდეგი უახლოესი მეზობლის არჩევა დროის $t = 3$ და $t = 4$ მომენტებისთვის მოგვცემს ჯამური გლუვობის საუკეთესო მნიშვნელობას, ამიტომ ჭდეებისთვის ადგილების შეცვლა საჭირო აღარ იქნება. მაგრამ, დროის $t = 5$ მომენტში უახლოესი მეზობლის არჩევა მოგვცემს ტრაექტორიებს $(\square_1, \square_2, \square_3, \square_4, O_5)$ და $(O_1, O_2, O_3, O_4, \square_5)$. ამ შემთხვევაში ჯამური გლუვობა ტოლი იქნება $2.84 + 2.91 = 5.75$. თუ ახლა ტრაექტორიების ბოლო წერტილების ჭდეებს ადგილებს შევუცვლით და ისევ გამოვთვალოთ ჯამურ გლუვობას, ვნახავთ, რომ ჯამური გლუვობის მნიშვნელობა გაუმჯობესდება და ტოლი იქნება, $2.89 + 2.94 = 5.83$. ამგვარად, საბოლოო შედეგი იქნება გამოსახულებაზე O და \square ობიექტების მოძრაობის თვალთვალის შედეგად მიღებული ნამდვილი ტრაექტორიები, კერძოდ, 12.15 სურათზე მოცემული ჭდეების შესაბამისი.

12.7. ვიდეომიმდევრობების სეგმენტაცია

ვიდეომიმდევრობების სეგმენტაცია წარმოადგენს ვიდეომონაცემების დამუშავების ერთ-ერთ ამოცანას, რომლის მიზანია ვიდეოკადრების გრძელი მიმდევრობის დანაწევრება ცალკეული კინოსცენების ამსახველი კადრების ქვემიმდევრობებად. მაგალითად, რამდენიმე წუთიანი საინფორმაციო ტელევიზიური ყოველთვის შეიცავს 10-15 წამიან ფრაგმენტებს. ჩვეულებრივ, ასეთ ფრაგმენტებში კამერა გვიჩვენებს სტუდიაში მაგიდასთან მჯდომ წამყვანს. ფონი ამ დროს ყოველთვის მუდმივია, მაგრამ შეიძლება იცვლებოდეს სცენის მასშტაბი. ასეთი ფრაგმენტის შემდეგ ხდება უშუალოდ გადასვლა კონკრეტული მოვლენის (მაგალითად, პოლიტიკურის, სპორტულის და ა.შ.) ამსახველ ვიდეომიმდევრობაზე, სადაც თვით მოვლენა ნაჩვენებია სხვადასხვა პლანით და აგრეთვე მათ შორის გადასვლებით. ეს გადასვლები შეიძლება დაედოს საფუძვლად ვიდეომიმდევრობების სეგმენტაციას. კერძოდ, მათი აღმოჩენა შესაძლებელია გამოსახულებებზე დროის გარკვეული ინტერვალის განმავლობაში მახასიათებელი ნიშან-თვისებების მნიშვნელოვანი ცვლილებების შესახებ ინფორმაციაზე დაყრდნობით.

ვთქვათ, გვაქვს ვიდეომიმდევრობის I_t და $I_{t+\delta}$ კადრები. მაშინ მათ შორის სხვაობის რაოდენობრივი შეფასებისთვის შეიძლება გამოვიყენოთ შესაბამის პიქსელებს შორის საშუალო სხვაობა. დროის δ ინტერვალის იმისდა მიხედვით, თუ როგორია კამერის ეფექტი, შეიძლება მოიცავდეს ერთ ან რამდენიმე ინტერვალს. ამგვარად, სხვაობა I_t და $I_{t+\delta}$ კადრებს შორის არის

$$d_{pixel}(I_t, I_{t+\delta}) = \frac{\sum_{i=1}^N \sum_{j=1}^M |I_t[i,j] - I_{t+\delta}[i,j]|}{N \times M} \quad (12.4).$$

იმის გამო, რომ სხვაობა შეიძლება დიდი აღმოჩნდეს კამერის მცირე პანორამული წანაცვლების დროსაც კი, ან რაღაც ობიექტის მოძრაობისას თითქმის უძრავ სცენაზე, (12.4) გამოსახულებამ შეიძლება მოგვცეს ხარისხობრივად არასწორი შედეგები. კადრებს შორის განსხვავების შეფასების უფრო უკეთესი ვარიანტი მდგომარეობს გამოსახულების დაყოფაში დიდ ბლოკებად და შემდეგ იმის შემოწმებაში, არის თუ არა ამ ბლოკების უმეტესობა ორივე გამოსახულებაზე ერთმანეთის შესაბამისი. ამ მიდგომის მიხედვით, შესაბამის ბლოკებს შორის სიკამკაშის მნიშვნელოვანი ცვლილების ნიშან-თვისებად გამოიყენება სტატისტიკური მახასიათებელი, დასაჯერობის ფარდობა (12.5)

$$r = \frac{\left[\frac{v_1+v_2}{2} + \left(\frac{u_1-u_2}{2} \right)^2 \right]^2}{v_1 v_2} \quad (12.5),$$

$$d_{block}(B_1, B_2) = \begin{cases} 1, & \text{თუ } r > \tau_r; \\ 0, & \text{თუ } r \leq \tau_r; \end{cases}$$

$$d(I_1, I_2) = \sum_{B_{1i} \in I_1; B_{2i} \in I_2} d_{block}(B_{1i}, B_{2i}),$$

სადაც: u_1 და v_1 შესაბამისად ინტენსივობის მათემატიკური ლოდინი და დისპერსიაა I_1 გამოსახულების B_1 ბლოკში, ხოლო u_2 და v_2 კი შესაბამისად I_2 გამოსახულების B_2 ბლოკში; $d_{block}(B_1, B_2)$ – შესაბამისობა ბლოკებს შორის; τ_r - ზღურბლის მნიშვნელობა; $d(I_1, I_2)$ - სხვაობა I_1 და I_2 გამოსახულებებს შორის.

თუ ბლოკების უმეტესობას შორის ნულოვანი განსხვავებაა, მაშინ ითვლება, რომ ორივე გამოსახულება ერთ პლანს მიეკუთვნება. როცა ორივე გამოსახულებას აქვს მკვეთრად გა-

მოსატული ტექსტურა, მაშინ (12.5) უფრო სტაბილურია, ვიდრე (12.4). მაგრამ სამწუხაროდ ისიც არასტაბილური ხდება კამერის მცირე წანაცვლებებით გამოწვეული კადრთაშორისი ცვლილებების შემთხვევაში.

XIII თავი. გამოსახულებების სეგმენტაცია

13.1. სეგმენტაციის არსი და მიზნები

სცენის გამოსახულებაზე ობიექტებისა და მათი ნაწილების პოვნა ხორციელდება სეგმენტაციის, ან ნიშან-თვისებათა გამოყოფის მეთოდების საშუალებით. ნიშან-თვისებების გამოყოფის ქვეშ იგულისხმება გამოსახულებაზე განსაკუთრებული ტიპის ლოკალური სახასიათო ნიშნების გამოყოფა, კერძოდ, სიკაშკაშის დონეების მკვეთრად გამოხატული ვარდნების – კიდეების, წრფეების, მრუდების, ლაქების, კუთხეების და ა.შ. სეგმენტაციისას კი ხდება დისკრეტული გამოსახულების გამარტივება მისი ცალკეულ ბმულ არეებად – სეგმენტებად დანაწევრების გზით. სეგმენტები გამოსახულებაზე წარმოადგენს ერთნაირი ნახევარტონების დონეების მქონე მომიჯნავე ბმული ელემენტების ერთობლიობას. სეგმენტაცია იძლევა შესაძლებლობას მოვახდინოთ გადასვლა გამოსახულების უმარტივესი ნიშან-თვისებებიდან მისი აღწერისთვის საჭირო ისეთი საშუალებების კრებულზე, რომლებითაც შესაძლებელი იქნება სცენის გამოსახულების სიმბოლური აღწერა. ასეთი აღწერისას პიქსელებს ენიჭება არა სიკაშკაშის დონეების მნიშვნელობები, არამედ იმ კლასების აღმნიშვნელი ჭდეები, რომელთაც ისინი მიეკუთვნება. ასეთი სიმბოლოების მაგალითებს წარმოადგენს ობიექტის სწორხაზოვანი ან მრუდწირული საზღვრების (მომვლები კონტური, სილუეტი) შემაღლენელი წერტილებისგან შედგენილი ჯაჭვები, მუდმივი სიკაშკაშის მქონე ბმული კომპონენტები, ფერები ან ტექსტურები, ელემენტარული გეომეტრიული ფიგურები (მართკუთხედი, სამკუთხედი, წრეწირი). ამგვარად, გამოსახულების ანალიზი შეიძლება განვიხილოთ როგორც გამოსახულების ნაწილების ისეთი ერთობლიობების პოვნის ამოცანა,

რომელთა თვისებები და ურთიერთმიმართებები აკმაყოფილებს მოცემულ შეზღუდვებს და შეიძლება აღწერილ იქნეს მონიშნული გრაფის საშუალებით.

გამოსახულებათა ანალიზის მრავალ გამოყენებით ამოცანაში დანაწევრება ხდება ისეთ არეებად, რომლებთანაც დაკავშირებულია მოცემული ამოცანისთვის საჭირო არსობრივი ინფორმაცია. მაგალითად, დედამიწის ზონდირებისას თანამგზავრიდან მიღებული გამოსახულების დამუშავებისას არსობრივ არეებს წარმოადგენს ტყის მასივები, მდინარეები, სასოფლოსამეურნეო დანიშნულების ფართობები და ა.შ. სხვა ამოცანებში ინტერესის არეებად შეიძლება განიხილებოდეს სასაზღვრო პიქსელები, რომლებიც დაჯგუფებულია მონაკვეთებად ან წრეწირის რკალებად. ასევე ინტერესის არეებს შეიძლება წარმოადგენდეს პიქსელთა რაიმე ფორმის მქონე ერთობლიობები (წრეწირი, ელიფსი, მრავალკუთხედები და ა.შ.). გამოსახულების სეგმენტაციას აქვს ორი ძირითადი მიზანი. ერთი მდგომარეობს იმაში, რომ გამოსახულება დაიყოს მისი შემდგომი ანალიზისთვის მოსახერხებელ ნაწილებად. ხოლო მეორე მიზანი კი თვით გამოსახულების აღწერის ფორმის შეცვლაა. აქ იგულისხმება, რომ სეგმენტაციის შედეგად გამოსახულება უკვე აღიწერება არა ინტენსივობის ფუნქციის, არამედ უფრო მაღალი დონის სტრუქტურების საშუალებით.

ფორმალურად გამოსახულების სეგმენტაციის პროცესი ასე შეიძლება იქნეს წარმოდგენილი:

აღვნიშნოთ \mathbf{X} -ით ციფრულ გამოსახულებაზე არსებული ელემენტების (პიქსელების) სიმრავლე. ამავე დროს, ვთქვათ, ერთი მხრივ მოცემულია ის პირობები, რომლებითაც ხდება \mathbf{X} სიმრავლის ნებისმიერი ბმული \mathbf{A} კომპონენტის (არის) ერთგვაროვნების დადგენა და მეორე მხრივ, შესაბამისი ლოგიკური ფუნქცია $\mathbf{F}(\mathbf{A})$, რომელიც უდრის 1, თუ მოცემული პირობები

სრულდება, და უდრის 0, წინააღმდეგ შემთხვევაში. მაშინ სეგმენტაციის მიზანია ვიპოვოთ X სიმრავლის ისეთი დანაწევრება $A_i, i \in \overline{1, n}$ ბმულ კომპონენტებად (არეებად), რომ შესრულდეს შემდეგი პირობები:

1. $X = \cup_{i=1}^n A_i$ და $A_i \cap A_j = \emptyset$, როცა $i \neq j$;
2. ყოველი A_i წარმოადგენს ბმულ კომპონენტს;
3. ნებისმიერი A_i -თვის, $F(A_i) = 1$;
4. $F(B) = 0$ ყოველი ისეთი B ბმული კომპონენტისთვის, რომელსაც აქვს სახე
 $B = A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}, k > 1$. ეს პირობა მიუთითებს A_i ბმული კომპონენტების – სეგმენტების “მაქსიმალურობაზე”.

წარმოდგენილი პირობები მიუთითებენ იმ მთავარ ფორმალურ მოთხოვნებზე, რომლებსაც უნდა აკმაყოფილებდეს სეგმენტირებულ გამოსახულებაზე არსებული ერთგვაროვანი არეები (სეგმენტები). ამასთან სეგმენტაციის შედეგად მიღებული არეები უნდა აკმაყოფილებდეს შემდეგ მოთხოვნებს:

- სეგმენტირებული გამოსახულების არეები (სეგმენტები) უნდა იყოს სრული და ერთგვაროვანი რაიმე ნიშან-თვისების მიმართ (სიკაშკაშის, ფერის ან ტექსტურის);
- სეგმენტირებული არეების შიდა ნაწილებს უნდა ჰქონდეს მარტივი ფორმა და არ უნდა შეიცავდეს დიდი რაოდენობის წვრილ-წვრილ ნახვრეტებს;
- მომიჯნავე სეგმენტირებული არეები მნიშვნელოვნად უნდა განსხვავდებოდეს ერთმანეთისგან იმ მახასიათებე-

ლი ნიშნის მნიშვნელობის მიხედვით, რომლის მიმართაც ისინი ერთგვაროვანია;

- ყოველი სეგმენტირებული არის საზღვარი უნდა იყოს გლუვი, ყოველგვარი დაკბილვებისა და წყვეტების გარეშე.

რეალურად ზემოაღნიშნული მოთხოვნების შესრულება რთულია. ამის გამო საჭირო ხდება სცენის შესახებ დამატებით აპრიორული ინფორმაციის გამოყენება.

სეგმენტაციის ამოცანის გადაწყვეტა ემყარება ორ ურთიერთსაწინააღმდეგო მიდგომას. პირველში, გამოსახულების ელემენტების ლოკალური მახასიათებლების მსგავსების საფუძველზე ხდება მათი გაერთიანება შესაბამის სეგმენტში, რომელსაც შემდგომ ენიჭება სახელი ანუ არსობრივი ჭდე. ასეთ მიდგომას მიეკუთვნება მეთოდები: ზღურბლური დამუშავების და კლასტერიზაციის; არეების დაგროვების და დანაწევრების. მეორე მიდგომაში ხდება გამოსახულების იმ არეების გამოყოფა, რომლებიც შეესაბამება გამოსახულების ელემენტების ლოკალური მახასიათებლების მკვეთრ წყვეტებს ან განსხვავებებს. შესაბამისად, აქ სეგმენტაციის ამოცანა დაიყვანება საზღვრების გამოყოფის ამოცანაზე. შესაბამის მეთოდებს წარმოადგენს: სივრცითი დიფერენცირების მეთოდები; მაღალსიხშირული ფილტრაცია; ფუნქციონალური აპროქსიმაცია. თითოეულ ამ ჩამოთვლილ მეთოდში ნაჩვენებია გამოსახულების დამუშავებისადმი მიდგომის თავისებურება და შესაბამისად შემოთავაზებულია ერთგვაროვანი არეების გამოყოფის საკითხის დასმისა და გადაწყვეტის (წინასწარ მოცემული გარკვეული კრიტერიუმით) შესაბამისი ხერხი. სეგმენტები გამოიყოფა გამოსახულების წერტილების გარკვეული ლოკალური მახასია-

თებელი ნიშან-თვისებების მიმართ, როგორცაა მაგალითად, სიკაშკაშის დონე, ფერი ან ტექსტურა.

თითოეული ზემოაღნიშნული მეთოდის გამოყენების შედეგი მნიშვნელოვნად არის დამოკიდებული თვით გამოსახულების სირთულეზე. სეგმენტაციის შედეგად მიღებული ერთგვაროვანი არე წარმოადგენს ლოკალურ-კონტრასტულ წარმონაქმნს, რომლის არაერთგვაროვნება მის გარემომცველ გამოსახულების ნაწილთან შედარებით გაცილებით მცირეა, ვიდრე განსხვავება. იმის გამო, რომ გამოსახულებაზე ადგილი აქვს სეგმენტების მრავალნაირობას, ხოლო მათ შიგნით კი მახასიათებლების ვარიაციას, მათი შესაბამისი კლასტერები ნიშან-თვისებათა სივრცეში ერთმანეთს ფარავს. ამის გამო კლასტერიზაცია გამოიყენება უფრო მარტივი გამოსახულებების – პისტოლოგიური გადანაჭრების, სისხლის სინჯების და ა.შ. სეგმენტაციისათვის. ასეთ გამოსახულებებში ერთგვაროვანი არეების რაოდენობა მცირეა, ხოლო მათი ზომები საკმაოდ დიდი. დანაწევრების მეთოდები სეგმენტების არაერთგვაროვნების განსაზღვრისათვის ითხოვს გლობალური კონტრასტის არსებობას, რის გამოც მათი გამოყენებისას ჯერ ხდება გამოსახულების ძლიერი საწყისი დანაწევრება შემდგომი დაგროვებით. გამოსახულებაზე ლოკალურ-კონტრასტული წარმონაქმნების გამოსაყოფად უფრო გამოსადეგია დაგროვებისა და საზღვრების ანალიზის მეთოდები. ამის მიზეზი კი ის არის რომ თავისი გარემოცვის მიმართ უფრო კონტრასტულ არეს საზღვრის ცალკეულ უბნებზე შეიძლება ჰქონდეს უფრო ნაკლები კონტრასტი, ვიდრე თავის შიგნით. ასეთი უბნების არსებობას დაგროვების მეთოდების გამოყენებისას მივყავართ მეზობელი ელემენტების შერწყმისკენ.

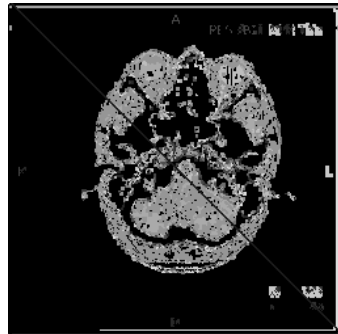
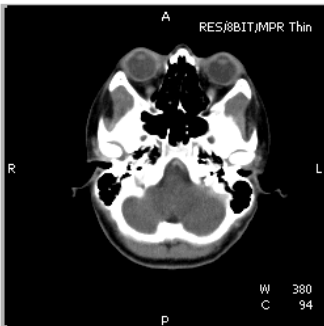
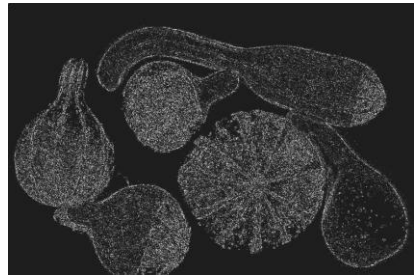
რთული გამოსახულებების სეგმენტაციისთვის უფრო გამოსადეგი აღმოჩნდა საზღვრების ანალიზის მეთოდები იმ მი-

ზეხით, რომ არეების საზღვრის გარკვეულ ნაწილში ერთგვაროვნების დარღვევისას მისი დანარჩენი ნაწილი შეიძლება აღდგენილ იქნეს ერთგვაროვნების მოცემული კრიტერიუმის თანახმად, ხოლო დაბალკონტრასტული უბნების (საზღვრების წყვეტების) აღდგენა ხდება შეპირისპირების მეთოდების გამოყენებით. რთული გამოსახულებების სეგმენტაცია მოითხოვს კომპლექსურ მიდგომას. აქ საჭიროა მხედველობაში მიღებულ იქნეს როგორც ერთგვაროვნების სხვადასხვა პირობები, ასევე თვით გამოსახულების სხვადასხვა მახასიათებლები, რაც გულისხმობს იმას, რომ, ერთი მხრივ, სეგმენტაციისთვის შეიძლება ერთობლივად იქნეს გამოყენებული როგორც კლასტერიზაციის, ასევე დაგროვების და საზღვრების ანალიზის მეთოდები. მეორე მხრივ, მხედველობაში უნდა იქნეს მიღებული როგორც გამოსახულების ელემენტების სხვადასხვა მახასიათებლები (სიკაშკაშე, ფერი), ასევე გამოსახულების სხვადასხვა ტიპის ფრაგმენტების (ელემენტი, გარემო, სეგმენტი) მახასიათებლები. ამავე დროს სცენის გამოსახულების სეგმენტაციისთვის შესაბამისი მეთოდის შერჩევისათვის ასევე საჭიროა მხედველობაში მივიღოთ გამოსახულების შესახებ არსებული კონტექსტური ინფორმაცია. ციფრული გამოსახულების სეგმენტაციის მაგალითები, სადაც მოცემულია როგორც საწყისი, ასევე სეგმენტირებული გამოსახულებები, ნაჩვენებია 13.1 სურათზე.

გარდა ზემოაღნიშნული მეთოდებისა არსებობს ასევე სეგმენტაციის სხვა მეთოდებიც. კერძოდ:

- **მოდელზე დამყარებული (model-based) სეგმენტაცია:** გამოსახულების სეგმენტაცია ხდება სეგმენტების მოცემული ფორმის და სხვა თვისებების (მოდელის) საფუძველზე;

- კიდზე დამყარებული (edge-based) სეგმენტაცია: ობიექტის კონტურები იძებნება კიდის წერტილების მიდევნების გზით;
- ტექსტურაზე დამყარებული (texture-based) სეგმენტაცია: სეგმენტების გამოსაყოფად გამოიყენება არეების ტექსტურული თვისებები;
- მოძრაობაზე დამყარებული (motion-based) სეგმენტაცია: ობიექტების ლოკალიზაცია ხდება მათ გადაადგილებაზე დაყრდნობით.



სურ. 13.1

ახლა განვიხილოთ სეგმენტაციის მეთოდები და მათი შესაბამისი ზოგიერთი ალგორითმი უფრო დეტალურად.

13.2. კლასტერიზაციის მეთოდები

მიჩნეულია, რომ გამოსახულებაზე ერთგვაროვანი ობიექტები გვევლინება როგორც კლასტერები (მსგავსი თვისებების მქონე ობიექტების არაცარიელი ქვესიმრავლეები) ნიშან-თვისებათა სივრცეში. აქედან გამომდინარე, სეგმენტების გამოყოფის ამოცანა შეიძლება განხილულ იქნეს როგორც გამოსახულების სივრცეში განხორციელებული კლასტერიზაცია (clustering).

გამოსახულების სეგმენტაციისას კლასტერიზაციის მეთოდის გამოყენებით მონაცემთა დაჯგუფება ხდება გამოსახულების სივრცეში, ანუ ამ დროს მხედველობაში მიიღება გამოსახულების ელემენტების არა მარტო სპექტრალური მახასიათებლები, არამედ მათ შორის სივრცითი ურთიერთობებიც, ხოლო კლასტერიზაცია კი ხორციელდება ნიშან-თვისებათა სივრცეში. რეალურად ეს იმას ნიშნავს, რომ გამოსახულების ყოველი პიქსელისთვის განისაზღვრება მახასიათებელ ნიშან-თვისებათა კრებული, წარმოდგენილი ვექტორის სახით. ამის შემდეგ ნიშან-თვისებათა სივრცეში გამოიყოფა წერტილთა (ვექტორთა) გროვები – კლასტერები, რომელთაც გამოსახულების სივრცეში შეესაბამება თვისებებით მსგავსი ბმული ელემენტების არეები – სეგმენტები. თავის მხრივ ნიშან-თვისებები შეიძლება ასახავდეს:

- პიქსელის მახასიათებელს, წარმოდგენილს სკალარული (სიკაშკაშე) ან ვექტორული (ფერი, სიკაშკაშე მრავალ-ზონალურ გამოსახულებაში, გეოფიზიკური და გეოქიმიური ველების მნიშვნელობები) სახით;

- პიქსელის მიდამოს მახასიათებლებს სკალარის (საშუალო სიკაშკაშე, სიკაშკაშის გრადიენტი) ან ვექტორის (მეზობელი ელემენტების სიკაშკაშეები, მათი სიკაშკაშის დისპერსია და ა.შ.) სახით;
- გამოსახულების წინასწარ გამოყოფილი ფრაგმენტების მახასიათებლებს: ფორმა, ზომა, ტექსტურა და ა.შ.

ამგვარად ასეთი მიდგომის დროს გამოსახულების სეგმენტაცია შეიძლება წარმოვიდგინოთ როგორც ნიშან-თვისებათა სივრცეში გამოყოფილი კლასტერების უკუასახვა გამოსახულების სივრცეში, სადაც ბმული კომპონენტების მარკირების ალგორითმის გამოყენების შედეგად მიიღება ბმული არეები – სეგმენტები.

კლასტერული ანალიზი, რომელიც წარმოადგენს სეგმენტაციის ერთ-ერთ მეთოდს, თვით შეიძლება დაიყოს სამ მიმართულებად. პირველი ეფუძნება იმ დაშვებას, რომ კლასტერები შეესაბამება საწყის მონაცემთა სიმრავლის განაწილების სიმკვრივის ფუნქციის მაქსიმუმებს (მოდებს). ასეთი მიდგომა კითავის მხრივ მოითხოვს სიმკვრივის ფუნქციის შეფასებას და მის საფუძველზე მოდების გამოთვლას. ამ მიდგომაზე დაფუძნებული კლასტერიზაციის მეთოდები თავისი არსით მიეკუთვნება მონაცემთა ანალიზის სტატისტიკურ მიმართულებას და მათ არაპარამეტრული კლასტერიზაციის მეთოდები ეწოდება. ასეთი მიდგომის დროს წინასწარ უცნობია როგორც მონაცემთა განაწილება, ასევე კლასტერთა რაოდენობა.

იმ შემთხვევაში, როცა მონაცემთა განაწილება და კლასტერთა რაოდენობა წინასწარაა ცნობილი, გამოიყენება კლასტერ-ანალიზის მეორე მიმართულება – პარამეტრული კლასტერიზაციის მეთოდები. ასეთ შემთხვევაში მონაცემთა სიმრავლის კლასტერებად დაყოფის კრიტერიუმად შესაძლებელია გამო-

ყენებულ იქნეს კრიტერიუმები, რომლებსაც ევრისტიკული ხასიათი აქვს, ან რომლებიც მიღებულია კლასტერიზაციის ხარისხის მაჩვენებლის მინიმიზაციით ან მაქსიმიზაციით. არსებობს ხარისხის მაჩვენებლების საკმაოდ ფართო სპექტრი. კერძოდ: კლასტერის ცენტრსა და მოცემულ ელემენტს შორის მანძილის კვადრატის მინიმუმი; კლასტერის ელემენტებს შორის მანძილების კვადრატების საშუალო; სხვადასხვა კლასტერების ელემენტებს შორის მანძილების კვადრატების საშუალო და ა.შ. ამ მიმართულებების ალგორითმებს შორის ყველაზე უფრო პოპულარულია ISODATA და მისი სხვადასხვა მოდიფიკაციები.

კლასტერ-ანალიზის მესამე მიმართულებას მიეკუთვნება **იერარქიული ტიპის ალგორითმები**. ასეთ მეთოდებში დასაწყისში მონაცემთა სიმრავლის ყველა ელემენტი განიხილება როგორც ცალკეული კლასტერი. ამის შემდეგ ხდება ყველაზე უფრო მსგავსი ორი უახლოესი კლასტერის გაერთიანება და ეს პროცესი გრძელდება მანამ, სანამ არ მიიღება ერთადერთი კლასტერი. ცხადია, ასეთი სქემა მოითხოვს კლასტერებს შორის მსგავსების შეფასებისათვის შესაბამისი მეტრიკის შემოტანას. ასეთი პროცედურების ჩატარების შედეგად აიგება ხის მსგავსი სტრუქტურა, ე.წ. “დენდროგრამა”, რომელიც ასახავს კლასტერთა თანამიმდევრული გაერთიანების პროცესს კლასტერიზაციის ზომის გათვალისწინებით. ამასთან აღსანიშნავია, რომ ხის გარკვეულ დონეზე ერთ კლასტერად გაერთიანებული ორი კლასტერი ასეთად რჩება შემდეგ დონეებზედაც.

მიუხედავად ზემოაღნიშნულ მიმართულებებს შორის განსხვავებისა, მათი გამოყენებით მონაცემთა სიმრავლის კლასტერებად დანაწევრება უნდა იყოს სასურველი, ანუ მიღებული კლასტერები უნდა აკმაყოფილებდეს ერთგვაროვნების პირობას.

გამოთვლითი თვალსაზრისით კლასტერიზაციის სხვადასხვა პროცედურები ითხოვს მესხიერების განსხვავებულ მოცულობას და დროის დანახარჯს. იქ, სადაც ყოველ პიქსელში ნიშან-თვისების მნიშვნელობა დარდება სხვა დანარჩენ პიქსელებში ნიშან-თვისებების მნიშვნელობებს, დროის დანახარჯი დიდია. ასევე დროის დიდ დანახარჯს მოითხოვს კლასტერიზაციის იტერაციული პროცედურა – ISODATA, იმის გამო, რომ მასში მრავალჯერ ხდება გამოსახულების ელემენტებზე გაგლა. ჰისტოგრამის მოდების ძიებაზე დამყარებულ კლასტერიზაციის ალგორითმებში იმის გამო, რომ მათში მონაცემებზე გაგლა ერთხელ ხდება, დროის დანახარჯი ნაკლებია.

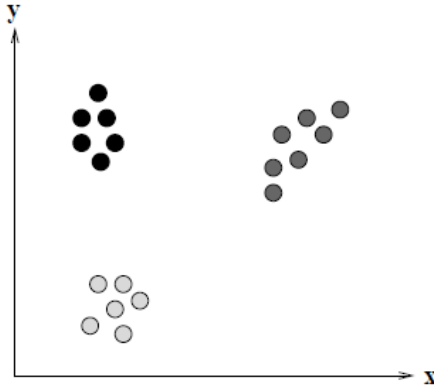
გეომეტრიული თვალსაზრისით ყოველი კლასტერი ნიშან-თვისებათა სივრცეში შეიძლება წარმოვიდგინოთ როგორც ყველაზე უფრო მსგავსი თვისებების მქონე კომპაქტურად განლაგებულ წერტილთა გროვა (ერთობლიობა), ან ღრუბელი. თუ ნიშან-თვისებათა ვექტორები ორგანზომილებიანია და წარმოდგენილია ნამდვილი რიცხვების წყვილების სახით, მაშინ ვექტორთა ასეთი სიმრავლის კლასტერიზაცია შეიძლება განხორციელდეს ევკლიდურ სიბრტყეზე ერთმანეთთან ახლოვანებაზე უფრო წერტილების ქვესიმრავლეების მოძებნით. სწორედ ასეთი შემთხვევაა წარმოდგენილი 13.2 სურათზე.

სურათზე წარმოდგენილი წერტილები წარმოადგენს ევკლიდურ სიბრტყეზე განლაგებულ გაზომვის შედეგებს. წერტილთა ეს სიმრავლე დანაწევრებულია სამ კლასტერად, რომელთაგან თითოეული წარმოადგენს ერთმანეთთან ახლოვანებაზე უფრო წერტილების ქვესიმრავლეს.

ზოგადად, კლასტერიზაციის ამოცანა განიხილება ასეთნაირად: მოცემულია K კლასტერი C_1, C_2, \dots, C_K და მათი შესაბამისი მათემატიკური ლოდინები m_1, m_2, \dots, m_K . სიდიდეს

$$D = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - m_k\|^2,$$

რომელიც განსაზღვრავს საწყისი მონაცემების სიახლოვეს მათ შესაბამის კლასტერებთან, ეწოდება კვადრატული ცდომილება.



სურ. 13.2

თუ კლასტერიზაციისთვის ვიყენებთ უმცირეს კვადრატთა მეთოდს, მაშინ მონაცემთა სიმრავლის K კლასტერად დაყოფის ყველა შესაძლო შემთხვევის გადარჩევის შედეგად შეიძლება ბოდა იმ ვარიანტის ნახვა, რომლის დროსაც D იღებს მინიმალურ მნიშვნელობას. მაგრამ ეს მოითხოვს გამოთვლების დიდ მოცულობას და დროს. ამიტომ იყენებენ მიახლოებით მეთოდებს. ძალზე მნიშვნელოვანია იმის ცოდნა, წინასწარაა ცნობილი კლასტერების K რაოდენობა, თუ არა. ზოგიერთ ალგორითმში K პარამეტრის მნიშვნელობა მომხმარებლის მიერ მოიცემა. სხვებში კი, საუკეთესო K -ს პოვნა ხდება რომელიმე კრიტერიუმის საფუძველზე. მაგალითად, ასეთი: ყოველი კლასტერის დისპერსია არ აღემატებოდეს წინასწარ მოცემული ზღურბლის მნიშვნელობას.

იტერაციული კლასტერიზაცია მათემატიკური ლოდინის გამოყენებით (Iterative K-means clustering). კლასტერიზაციის ეს ალგორითმი წარმოადგენს ექსტრემუმის ძიების მარტივ იტერაციულ ალგორითმს. საწყის მონაცემთა სიმრავლეს წარმოადგენს n -განზომილებიან ვექტორთა სიმრავლე. ამ სიმრავლეში K რაოდენობის კლასტერთა გამოსაყოფად საჭიროა შემდეგი მოქმედებების განხორციელება:

1. იტერაციის n მრიცხველს მივანიჭოთ მნიშვნელობა 1.
2. ავირჩიოთ ნებისმიერად K მათემატიკური $m_1(1), m_2(1), \dots, m_K(1)$ ლოდინის მნიშვნელობები (ანუ ნებისმიერად დავანაწევროთ ვექტორების სიმრავლე K რაოდენობის კლასტერებად და გამოვთვალოთ მათი მათემატიკური ლოდინის მნიშვნელობები).
3. თითოეული x_i ვექტორისთვის გამოვთვალოთ მანძილი $D(x_i, m_k(n))$ ყოველი $k = \overline{1, K}$ -თვის და მოვათავსოთ x_i -ური უახლოესი მათემატიკური ლოდინის ვექტორის მქონე C_j კლასტერში.
4. გავზარდოთ n -ის მნიშვნელობა ერთით და მოვახდინოთ მათემატიკური ლოდინის მნიშვნელობების კორექტირება ახალი $m_1(n), m_2(n), \dots, m_K(n)$ სიმრავლის მისაღებად.
5. გავიმეოროთ ნაბიჯები 3 და 4 მანამ, სანამ ყველა k -თვის არ შესრულდება პირობა $C_k(n) = C_k(n + 1), k = \overline{1, K}$.

წარმოდგენილი ალგორითმი კარგი კრებადობით ხასიათდება, თუმცა მიღებული შედეგი უმცირესი კვადრატული ცდომილების კრიტერიუმის აზრით არ იქნება აუცილებლად გლობალურად ოპტიმალური. ამასთან აღსანიშნავია, რომ ნაბიჯი 2

ვექტორთა სიმრავლის K რაოდენობის კლასტერებად ნებისმიერად დანაწევრებისთვის და შემდგომ მათი მათემატიკური ლოდინის გამოთვლისთვის შეიძლება შეცვლილი იქნეს, ხოლო ნაბიჯზე 5 კი შესაძლებელია ალგორითმის დასრულება, თუ ბოლო იტერაციაზე კლასტერები ვექტორთა მცირე რაოდენობას შეეცვალათ.

კლასტერიზაციის ალგორითმი ISODATA (Iterative self-organizing data analysis techniques). კლასტერიზაციის ერთ-ერთ ყველაზე გავრცელებულ ალგორითმს წარმოადგენს ISODATA, რომელშიც ერთდროულად გამოიყენება როგორც ევრისტიკული მიდგომა, ასევე ხარისხის მაჩვენებელი. დასაწყისში აქ ნებისმიერად ხდება კლასტერიზაციის საწყისი ცენტრების არჩევა, რისთვისაც იყენებენ ამონაკრებს მონაცემთა მოცემული სიმრავლიდან. ამის შემდეგ მიმდინარე მონაცემი მიეკუთვნება იმ კლასტერს, რომლის ცენტრიდან მოცემულ მონაცემამდე მანძილის კვადრატი მინიმალურია (ხარისხის მაჩვენებელი). მას შემდეგ, რაც მოხდება მონაცემების მიკუთვნება შესაბამისი კლასტერების ცენტრებისადმი, ხდება კლასტერების ცენტრების კორექტირება. ამისათვის ყოველი კლასტერის ცენტრი უტოლდება ამ კლასტერის ამოკრებით საშუალოს. ცენტრების კორექტირების პროცედურა იტერაციულია და გრძელდება მანამ, სანამ არ იქნება მიღებული წინა ნაბიჯის იდენტური შედეგი. ამგვარად, ალგორითმში **მანძილის კვადრატის მინიმალურობის** კრიტერიუმის საფუძველზე ხდება კლასტერთა ცენტრების მდებარეობის თანამიმდევრული გაუმჯობესება.

აქაც, ისევე როგორც K -საშუალოების ალგორითმში, ვთქვათ, მოცემული გვაქვს K რაოდენობის C_1, C_2, \dots, C_K კლასტერი შესაბამისი მათემატიკური ლოდინებით m_1, m_2, \dots, m_K

და k -ური კლასტერის კოვარიაციის Σ_k მატრიცა. თუ საწყისი სიმრავლის x_i ელემენტები წარმოადგენს ვექტორებს, ანუ

$$x_i = [v_1, v_2, \dots, v_n],$$

მაშინ მათემატიკური ლოდინის თითოეულ m_k ვექტორს ექნება სახე

$$m_k = [m_{1k}, m_{2k}, \dots, m_{nk}],$$

ხოლო k კლასტერის კოვარიაციის მატრიცა კი იქნება ასეთი

$$\Sigma_k = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2n} \\ \cdot & \cdot & \dots & \cdot \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_{nn} \end{bmatrix},$$

სადაც, $\sigma_{ii} = \sigma_i^2$ არის ვექტორების i -ური v_i კომპონენტის დისპერსია, ხოლო $\sigma_{ij} = \rho_{ij} \sigma_i \sigma_j$ კი ვექტორების i -ურ და j -ურ კომპონენტებს შორის კოვარიაცია (ρ_{ij} -კორელაციის კოეფიციენტია i -ურ და j -ურ კომპონენტებს შორის, σ_i - i -ური კომპონენტის საშუალო კვადრატული გადახრაა, σ_j - j -ური კომპონენტის საშუალო კვადრატული გადახრაა).

ალგორითმის რეალიზაცია მოითხოვს შემდეგი ნაბიჯების განხორციელებას:

1. ვექტორი x_i მივაკუთვნოთ იმ l -ურ კლასტერს, რომლისთვისაც გამოსახულება

$$D_{\Sigma} = [x_i - m_i]' \sum_l^{-1} [x_i - m_i]$$

იღებს მინიმალურ მნიშვნელობას.

2. თუ

$$|m_i - m_j| < \tau_v,$$

მაშინ i და j კლასტერები ერთმანეთს ერწყმის (ერთიანდება); τ_v - დისპერსიის ზღურბლური მნიშვნელობაა.

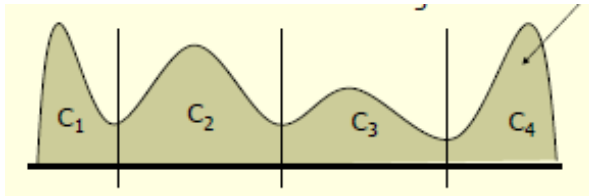
3. თუ k -ური კლასტერის კოვარიაციის Σ_k მატრიცის მაქსიმალური საკუთარი მნიშვნელობა σ_k გადააჭარბებს τ_v ზღურბლს, მაშინ k -ური კლასტერი ნაწევრდება.
4. თუ ყოველი i -ური კლასტერისთვის სრულდება პირობა

$$|m_i(t) - m_i(t + 1)| < \varepsilon,$$

ან თუ შესრულდა იტერაციათა მაქსიმალურად დასაშვები რაოდენობა, მაშინ ალგორითმის მუშაობა სრულდება.

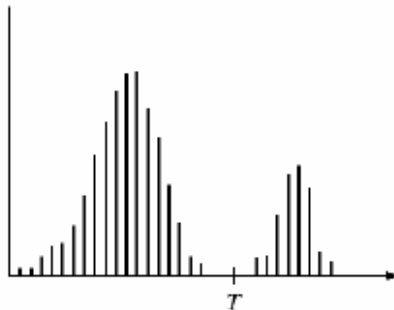
ჰისტოგრამული მეთოდები. ზემოაღწერილი ალგორითმები იტერაციული ხასიათისაა და, აქედან გამომდინარე, მოითხოვს გამოსახულებაზე მრავალჯერად გავლას. ეს კი თავის მხრივ საჭიროებს გამოთვლების დიდ მოცულობას და დროს. ამის თავიდან ასაცილებლად გამოსახულებათა სეგმენტაციისათვის შემუშავებულ იქნა **ჰისტოგრამის მოძებნის ძიებაზე** (histogram mode seeking) დამყარებული მეთოდი, რომელშიც გამოსახულების ჰისტოგრამის გამოთვლის საფუძველზე ხდება მოძებნის (ლოკალური მაქსიმუმების წერტილების) განსაზღვრა. კლასტე-

რებად (C_1, C_2, C_3, C_4) კი ცხადდება ის მნიშვნელობები ჰისტოგრამაზე, რომლებიც მოთავსებულია ჰისტოგრამის ლოკალური მაქსიმუმის ორივე მხარეს განლაგებულ ლოკალურ მინიმუმებს შორის. აღნიშნული შემთხვევა, ანუ როცა ჰისტოგრამაზე გვაქვს რამდენიმე ლოკალური მაქსიმუმი, შეესაბამება მრავალ-მოდულიან ჰისტოგრამას და იგი წარმოდგენილია 13.3 სურათზე.



სურ. 13.3

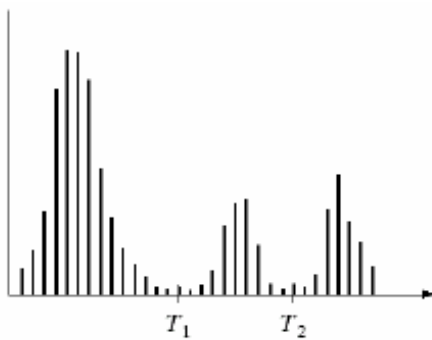
ჰისტოგრამულ მეთოდებში კლასტერიზაცია სორციელდება მონაცემთა ნიშან-თვისებების სივრცეში. მარტივი გამოსახულების შემთხვევაში, როცა გვაქვს რაიმე ერთგვაროვან ფონზე მოთავსებული ერთი ობიექტი, ჰისტოგრამას აქვს სახე (სურ. 13.4).



სურ. 13.4

ასეთ შემთხვევაში ობიექტის ფონიდან გამოსაყოფად საკმარისია T ზღურბლის ერთი მნიშვნელობის აღება. შესაბამისად, ყველა პიქსელი, რომელშიც სიკაშკაშის მნიშვნელობა მეტი იქნება T ზღურბლზე, მიეკუთვნება ობიექტს, დანარჩენი კი ფონს.

იმ შემთხვევაში, როცა ფონზე მოთავსებულია რამდენიმე ობიექტი, მაშინ უკვე თითოეული ობიექტის გამოყოფისთვის ზღურბლის რამდენიმე მნიშვნელობაა საჭირო. ასეთი შემთხვევა წარმოდგენილია 13.5 სურათზე, სადაც გვაქვს ზღურბლის T_1 და T_2 მნიშვნელობები. პიქსელები, რომლებშიც სიკაშკაშის მნიშვნელობები ნაკლებია T_1 -ზე, მიეკუთვნება ფონს, ხოლო ობიექტებს მიეკუთვნება ის პიქსელები, რომლებშიც სიკაშკაშის მნიშვნელობები მოთავსებულია T_1 და T_2 მნიშვნელობებს შორის და ასევე ისინი, რომელთა სიკაშკაშის მნიშვნელობები მეტია T_2 -ზე.



სურ. 13.5

მაგრამ რეალურად მოდების ძიება დაკავშირებულია მნიშვნელოვან სიძნელეებთან, რადგანაც ჰისტოგრამის სირთულის გამო ხშირად ძალიან ძნელი ხდება მოდების ერთმანეთისგან

განცალკევება. მიჩნეულია, რომ ნიშან-თვისებათა სივრცეში მიღებული კლასტერები შეესაბამება გამოსახულებაზე არსებულ ერთგვაროვან არეებს (ობიექტებს). შესაბამისად გამოსახულების სივრცეში სეგმენტების გამოსაყოფად ხორციელდება მიღებული კლასტერების უკუასახვა ნიშან-თვისებათა სივრციდან გამოსახულების სივრცეში. გამოსახულების იმ სეგმენტის პიქსელს, რომლის ინტენსივობის მნიშვნელობა ხვდება ჰისტოგრამაზე მიღებულ *i*-ურ კლასტერში, ენიჭება *i*-ური მნიშვნელობა. ამასთან, სეგმენტების როლში აიღება მაქსიმალური ბმული კომპონენტები. ამ მეთოდის შესაბამის ალგორითმებში საწყის მონაცემებზე გავლა ხდება მხოლოდ ერთხელ და ამის გამო ისინი კლასტერიზაციის იტერაციულ ალგორითმებთან შედარებით უფრო სწრაფია.

რეკურსიული ჰისტოგრამული მეთოდი (ოლანდერის მეთოდი). იგი წარმოადგენს ზემოაღწერილი კლასტერიზაციის ჰისტოგრამული მეთოდის გაუმჯობესებულ ვარიანტს, რომელშიც მოდების ძიებისთვის გამოყენებულია რეკურსიული პროცედურა. ამ მიდგომის ძირითადი არსი მდგომარეობს იმაში, რომ ჯერ მოდა მოიძებნოს მთლიანი გამოსახულების ჰისტოგრამაზე, ხოლო შემდეგ კი კლასტერიზაციის შედეგად მიღებული ყოველი კლასტერის შესაბამისი არეების ჰისტოგრამებზე პროცესი მეორდება მანამ, სანამ არ დამთავრდება სეგმენტირებული არეების შემდგომი დანაწევრება.

სეგმენტაციის პროცესი ჯერ იწყება ისეთი ტრაფარეტის ფორმირებით, რომელიც მოიცავს გამოსახულების ყველა პიქსელს. შემდეგ იგება ამ ტრაფარეტით მოცული გამოსახულების ჰისტოგრამა, რომლის საშუალებითაც კლასტერიზაცია ხორციელდება ნიშან-თვისებათა სივრცეში. შედეგად მიიღება კლასტერთა რაღაც სიმრავლე, რომლებსაც გამოსახულებაზე ერთ-

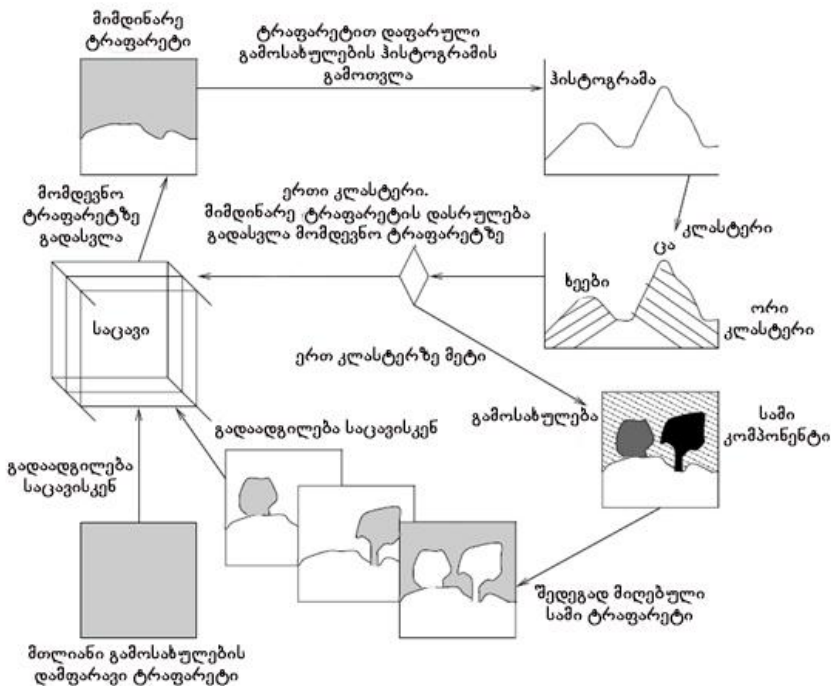
გვაროვანი არეები შეესაბამება; დგინდება გამოსახულების პიქსელების ამ კლასტერებისადმი კუთვნილება. თუ კლასტერთა რიცხვი ნიშან-თვისებათა სივრცეში ერთის ტოლია, მაშინ მოცემული ტრაფარეტით პროცესი სრულდება. თუ კლასტერების რიცხვი ერთზე მეტია, მაშინ თითოეული მათგანისთვის გამოიყენება მარკირების (ჭდეების მინიჭების) პროცედურა და თითოეული კლასტერის ჭდისთვის იგება ბმული არეების სიმრავლე. ამის შემდეგ თითოეული ბმული კომპონენტი გამოიყენება ახალი ტრაფარეტის გენერირებისთვის, რომელიც მოთავსებულია ტრაფარეტების საცავში.

ტრაფარეტები, რომლებიც საცავშია მოთავსებული, შეესაბამება იმ არეებს გამოსახულებაზე, რომლებიც სეგმენტაციის კანდიდატებს წარმოადგენს. საცავიდან აღებული ყოველი მომდევნო ტრაფარეტი განსაზღვრავს ჰისტოგრამის ასაგებად საჭირო პიქსელებს. თვით კლასტერიზაციის პროცესი კი მეორდება საცავის დაცარიელებამდე. კლასტერიზაციის აღნიშნულ პროცესს **რეკურსიულ სივრცით კლასტერიზაციას** უწოდებენ.

ასეთი კლასტერიზაციის მაგალითი წარმოდგენილია 13.6 სურათზე. საწყის გამოსახულებაზე მოცემულია ოთხი ობიექტი: ბალახი, ცა, და ორი ხე. კლასტერიზაციის პროცესის მიმდინარე ტრაფარეტი (ზედა მარცხენა კუთხეში) შეესაბამება ცისა და ხის შემცველ არეს. ამ არის ჰისტოგრამის კლასტერიზაციის შედეგად ფერთა სიცრცეში ვლელულობთ ორ კლასტერს, ერთს ცისთვის, და მეორს ხისთვის. ცის კლასტერი შეიცავს ერთ ბმულ კომპონენტს, ხოლო ხის – ორს. თითოეული ამ კომპონენტისთვის იგება ტრაფარეტი და თავსდება საცავში სეგმენტაციის პროცესში შემდეგ იტერაციებზე მათი გამოყენების მიზნით.

ფერადი გამოსახულებების შემთხვევაში რეკომენდებულია, რომ ჰისტოგრამების გამოთვლა ხდებოდეს სამი RGB ფე-

რადი კომპონენტიდან არა თითოეულისთვის ცალ-ცალკე, არამედ კარუნენ-ლოევის გარდაქმნის (მთავარი კომპონენტების) მსგავსი ცვლადების ასეთი კრებულებისთვის: $(R + G + B)/3$, $(R - B)/2$ და $(2G - R - B)/4$.



სურ. 13.6

თუ გამოსახულება წარმოადგენს მუქ ფონზე მოთავსებულ თეთრ ობიექტს (მარტივი გამოსახულება) და ნიშან-თვისებათა სივრცე ერთგანზომილებიანია, მაშინ კლასტერიზაცია ნიშან-თვისებათა სივრცეში დაიყვანება ზღურბლის ისეთი სიდიდის განსაზღვრაზე, რომელიც მასზე ნაკლებ ან ტოლ ყვე-

ლა წერტილს მიაკუთვნებს ერთ კლასტერს, ხოლო დანარჩენებს კი სხვას. ამგვარად, აქ დგება ზღურბლის გამოთვლის პროცედურის განსაზღვრის ამოცანა. მარტივ შემთხვევაში ეს პროცედურა დაიყვანება ნახევარტონების დონეების ჰისტოგრამის ანალიზზე და შემდგომ ზღურბლისთვის იმ მნიშვნელობის მინიჭებაზე, რომელიც ჰისტოგრამის ორ მაქსიმუმის წერტილს (მოდას) შორის მოთავსებულ არეში მდებარეობს, იქ სადაც სიკაშკაშის მნიშვნელობების სიხშირე დაბალია. რთული გამოსახულებების შემთხვევაში კი მიმართავენ ჰისტოგრამის ფორმის გაუმჯობესების ხერხებს, რომლებშიც ერთმანეთთან შერწყმულია გამოსახულების სივრცითი ინფორმაცია და ინფორმაცია ნახევარტონების სიკაშკაშის მნიშვნელობების შესახებ. მათი არსი მდგომარეობს იმაში, რომ ჰისტოგრამის ფორმის გასაუმჯობესებლად იყენებენ გამოსახულების ლოკალურ თვისებებს, მაგალითად მოცემული პიქსელის გრადიენტის მნიშვნელობებს ან ლაპლასიანს. ასეთი მიდგომის მაგალითს წარმოადგენს ალგორითმი, რომელშიც ჰისტოგრამა გამოითვლება პიქსელის სივრცით-ლოკალური არისთვის. ასეთ არეს შეიძლება წარმოადგენდეს პიქსელის (33 × 33) ან (65 × 65) ზომის მეზობლობა. გამოსახულების საერთო ზღურბლის სიდიდის განსაზღვრავად ხდება გამოთვლილი ჰისტოგრამების ზღურბლების სივრცითი ინტერპოლაცია. სხვა ალგორითმში ზღურბლის გამოსათვლელად ჰისტოგრამა გამოითვლება მხოლოდ იმ პიქსელებისთვის, რომლებშიც ლაპლასიანის მნიშვნელობები დიდია.

ამავე დროს უნდა აღინიშნოს ის ნაკლოვანებები, რომლებიც გააჩნია კლასტერიზაციის მეთოდს. კერძოდ, ისინი მუდამდგება გამოყოფილი საზღვრების არამკაფიობაში, მათ წყვეტებში. ამასთან ერთად, იმის გამო, რომ კლასტერების რიცხვი წინასწარ ყოველთვის ცნობილი არაა, ხოლო თვით კლასტერი-

ზაციის პროცესი ბევრადღა დამოკიდებული ევრისტიკებზე, შესაბამისად სეგმენტების რაოდენობის წინასწარ განჭვრეტაც შეუძლებელია. მაგრამ ამავე დროს ეს მეთოდი ხასიათდება დიდი მდგრადობით ხმაურის მიმართ. ამის გამო კლასტერიზაციის მეთოდმა თავისი გაერცელება ჰპოვა მარტივი გამოსახულებების სეგმენტაციისთვის, სადაც მოსაძიებელი არეების რიცხვი მცირეა, ხოლო ფართი კი დიდი.

13.3. არეების დაგროვების მეთოდი

არეების დაგროვების (region growing) მეთოდი, განსხვავებით გამოსახულების ერთგვაროვან არეებად დანაწევრების მეთოდებისგან, ერთგვაროვანი არეების ფორმირებას იწყებს გამოსახულების რომელიმე წერტილიდან (ჩვეულებრივად პროცესი იწყება ზედა მარცხენა კუთხიდან) და შემდეგ ახდენს მეზობელი პიქსელების თანამიმდევრულ მიერთებას რაიმე კრიტერიუმის საფუძველზე უფრო დიდი არეების მისაღებად, რომლებიც აკმაყოფილებს სეგმენტაციის პირობებს. IV თავში განვსაზღვრეთ ბმულობის არსი ბინარული გამოსახულებისთვის. მაგრამ საზოგადოდ, ბმულობა შეიძლება ასევე განსაზღვრულ იქნეს ნახევარტონური გამოსახულებისთვისაც.

განსაზღვრება 13.1. *ნახევარტონური გამოსახულების R ელემენტთა სიმრავლეს ეწოდება ელემენტარული ბმული არე, თუ:*

1. *R სიმრავლის ყველა ელემენტს აქვს ნახევარტონების ერთნაირი დონე;*

2. R სიმრავლის ნებისმიერი ორი ელემენტი შეერთებულია მომიჯნავე ელემენტების ჯაჭვით, რომელთაგან თითოეული მიეკუთვნება R -ს;
3. ელემენტების ნებისმიერი სიმრავლე, რომელიც მოიცავს R -ს, მაგრამ არ ემთხვევა მას, არ აკმაყოფილებს წინა ორ პირობას.

მეორე პირობა შეიძლება აღებულ იქნეს როგორც **ბმული არის** განმარტება, ხოლო მესამე პირობა აღნიშნავს, რომ ასეთი არე ზომით მაქსიმალური უნდა იყოს.

სეგმენტაციისთვის არეების დაგროვების მიდგომის გამოყენებისას არის ფორმირების ერთგვაროვნების კრიტერიუმი შეიძლება ემყარებოდეს არის შემდეგ მახასიათებლებს:

- საშუალო სიკაშკაშე;
- დისპერსია;
- ფერი;
- ტექსტურა;
- მოძრაობა;
- ფორმა;
- ზომა.

არეების დაგროვების მეთოდის რეალიზებისთვის არსებობს სამი განსხვავებული მიდგომა: **არეების შერწყმა** (region merge), **არეების დანაწევრება** (region split), **არეების დანაწევრება და შერწყმა** (region split and merge).

მეთოდის იდეა მდგომარეობს შემდეგში: გამოსახულების სივრცეში აირჩევა რამდენიმე სასტარტო წერტილი, რომელთა რაოდენობა ითვლება რომ წინასწარაა ცნობილია. პრაქტიკუ-

ლად კი სასტარტო წერტილების არჩევა დამოკიდებულია ამოცანის ბუნებაზე. ამის შემდეგ ეს წერტილები მარკირდება; შემდეგ წინასწარ არჩეული მეტრიკით ხდება მათი შედარება მეზობელ წერტილებთან ერთგვაროვნების პირობის შესრულებაზე. პირობის შესრულების შემთხვევაში მეზობელ წერტილს ენიჭება სასტარტო წერტილის ჭდე. ამის შემდეგ უკვე ეს მეზობლები განიხილება როგორც სასტარტო და იგივე პროცედურა თავიდან მეორდება მათი მეზობლების მიმართაც. საბოლოოდ, გამოსახულების ყოველი წერტილი მიიღებს ამა თუ იმ ჭდეს. მაგრამ რეალობაში, ჩვეულებრივ, აპრიორული ინფორმაციის არქონის გამო, სასტარტო წერტილები აირჩევა ნებისმიერად. ამის გამო არეების დაგროვების მეთოდებს აქვს ემპირიული ხასიათი.

დაგროვების მეთოდები ემყარება შემდეგ ძირითად მომენტებს:

- საწყისი პიქსელების არჩევა;
- იმ ლოგიკური პრედიკატის სახე, რომელიც მართავს არეების დაგროვების პროცესს;
- გამოსახულების გადახედვის მიმართულების განსაზღვრა.

არეების დაგროვების პროცესი გრძელდება მანამ, სანამ შესაძარბებელ პიქსელებს და არეების პიქსელებს შორის განსხვავება საგრძნობი არ გახდება. პიქსელთა შედარებისათვის იყენებენ სტატისტიკურ კრიტერიუმებს. არეების დაგროვებისათვის რ. ჰარალიკისა და ლ. შაპიროს მიერ შემოთავაზებულ იქნა არეების დაგროვების ჰარალიკის მეთოდი, რომლის მიხედვითაც არე წარმოადგენს ბმული პიქსელების სიმრავლეს საერთო ამოკრებითი საშუალოთი და გაფანტულობით.

ვთქვათ, გვაქვს N პიქსელისგან შედგენილი R არე და მისი მეზობელი პიქსელი ინტენსივობის მნიშვნელობით y . R არის ამოკრებითი საშუალო \bar{X} და გაფანტულობა S^2 განისაზღვრება შემდეგი გამოსახულებებით:

$$\bar{X} = \frac{1}{N} \sum_{[i,j] \in R} I[i,j]$$

და

$$S^2 = \sum_{[i,j] \in R} (I[i,j] - \bar{X})^2.$$

იმ დაშვების გათვალისწინებით, რომ R არის ყველა პიქსელი და შესადარებელი y პიქსელი დამოუკიდებელია და მათი ინტენსივობები აღიწერება ერთნაირი ნორმალური განაწილებით, შემთხვევითი სიდიდე

$$T = \left[\frac{(N-1)N}{(N+1)} (y - \bar{X})^2 / S^2 \right]^{\frac{1}{2}}$$

აღიწერება $(N-1)$ თავისუფლების ხარისხის მქონე სტიუდენტის t -განაწილებით. თუ T -ს მნიშვნელობა საკმარის პატარაა, მაშინ y პიქსელი დაემატება R არეს და ამოკრებითი საშუალოსა და გაფანტულობის მნიშვნელობები თავიდან გამოითვლება y პიქსელის მნიშვნელობის გათვალისწინებით შემდეგი ფორმულების გამოყენებით

$$\bar{X}_{new} \leftarrow (N\bar{X}_{old} + y) / (N + 1)$$

და

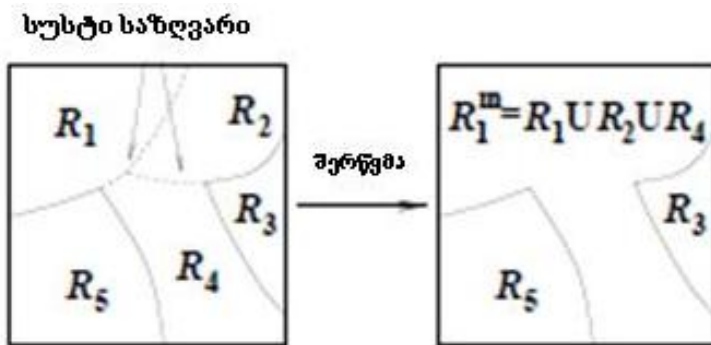
$$S_{new}^2 \leftarrow S_{old}^2 + (y - \bar{X}_{new})^2 + N(\bar{X}_{new} - \bar{X}_{old})^2.$$

იმ შემთხვევაში, თუ T -ს მნიშვნელობა ძალიან დიდია, მაშინ ეს ნიშნავს, რომ y პიქსელის R არისადმი კუთვნილების ალბათობა მცირეა. ხოლო თუ y პიქსელში ინტენსივობის მნიშვნელობა განსხვავდება ინტენსივობის მნიშვნელობებისგან ყველა მეზობელ არეში, მაშინ უკვე ეს პიქსელი აიღება როგორც ახალი არის საწყისი პიქსელი. არეების დაგროვების უფრო მკაცრი კრიტერიუმი კი შეიძლება მოითხოვდეს არა მარტო y პიქსელის მნიშვნელობის სიახლოვეს მეზობელი არის ამოკრებით საშუალოსთან, არამედ იმასაც, რომ ამ არიდან მეზობელი პიქსელის მნიშვნელობაც საკმაოდ ახლოს იყოს y პიქსელის მნიშვნელობასთან.

განვიხილოთ არეების დაგროვების ერთ-ერთი ცნობილი ალგორითმი, რომელშიც შესასვლელი გამოსახულება წარმოიდგინება (2×2) , (4×4) ან (8×8) განზომილების მქონე ტრაფარეტის დადების შედეგად მიღებული უჯრედების სახით. თითოეული ამ უჯრედისთვის ხდება ნახევარტონური დონეების ალბათობათა განაწილების შეფასება. ამის შემდეგ, დაწყებული რომელიმე უჯრედიდან, მაგალითად ზედა მარცხენა მხარეს განლაგებულიდან, ხდება მისი სტატისტიკების შედარება მეზობელ უჯრედებთან მათი მსგავსების დასადგენად. მსგავსების საზომად გამოიყენება კოლმოგოროვ-სმირნოვის ტესტი. თუ ორი უჯრედი აკმაყოფილებს მსგავსების პირობას, მაშინ ისინი ერთიანდება და წარმოქმნის ფრაგმენტს, თუ არა, მაშინ შერწყმას ადგილი არ აქვს. მიღებული ფრაგმენტისთვის ისევ განისაზღვრება სტატისტიკები, დარდება მისი მეზობლების სტატისტიკებს და მეორდება ზემოაღწერილი პროცედურა. მაგრამ რომელიდაც ნაბიჯზე ფრაგმენტის დაგროვება წყდება. სწორედ ასეთი ფრაგმენტი ქმნის ერთგვაროვან არეს. ამის შემდეგ ხდება შემდეგი ფრაგმენტის ფორმირება, რისთვისაც აირჩევა ის

უჯრედი, რომელიც არ მოხვდა ერთგვაროვან არეში და პროცესი ისევ მეორდება თავიდან.

არეების დაგროვების ალგორითმების განვითარება მოხდა ევრისტიკების შემოტანით, რაც გამოიხატა გლობალური ინფორმაციის გამოყენებით ლოკალური გადაწყვეტილებების მისაღებად. კერძოდ, მეზობლების გაერთიანების პროცედურამ უნდა გაითვალისწინოს გლობალური შედეგები. ასეთ ალგორითმში თავიდან საწყისი გამოსახულება ნაწევრდება ნახევარტონების მუდმივი დონეების “ატომარულ” არეებად, რომელთა გაერთიანება ხდება იმ ევრისტიკებით, რომლებიც ეფუძნება მომიჯნავე არეების საზღვრებზე არსებულ ინფორმაციას. საქმე ისაა, რომ ორ “ატომარულ” არეს შორის საზღვარი შეიცავს ნაწილებს, რომელთა გასწვრივ ნახევარტონების დონეებს შორის სხვაობები მცირეა. აქ ევრისტიკის გამოყენების არსი იმაში მდგომარეობს, რომ ორი არე ერთმანეთს ერწყმის, თუ მათ შორის საზღვარი სუსტია (ანუ ადგილი აქვს დაბალ კონტრასტს) და შერწყმის შედეგად მიღებული ახალი არის საზღვარი უფრო მოკლეა, ვიდრე წინა ორის. ასეთი შემთხვევა ნახვენებია 13.7 სურათზე.



სურ. 13.7

არეების დაგროვებისთვის ასევე იყენებენ ფუნქციონალურ აპროქსიმაციას, როგორც მათემატიკურ საფუძველს გამოსახულებაზე ერთგვაროვანი არეების გამოყოფისთვის. ამისათვის აპროქსიმაციის ფუნქციის (მუდმივის ან წრფივის) შერჩევის შემდეგ გამოსახულებაზე იძებნება მათი შესაბამისი ლოკალური არეები, რომელთა აპროქსიმაციაც შესაძლებელია მარტივი აპროქსიმაციის ფუნქციებით. შერწყმა კი ხდება იმ არეების, რომელთა აპროქსიმაციის კოეფიციენტები საკმაოდ ახლოსაა ერთმანეთთან.

ხშირად არეების დაგროვების ალგორითმის მოქმედების შედეგი დამოკიდებულია ზღურბლის არჩევაზე. ყველაზე უფრო მარტივ შემთხვევას წარმოადგენს მუდმივი ზღურბლის გამოყენება. ასეთ მიდგომაზე დამყარებულ ალგორითმებში დასაგროვებელი არე და მისი მეზობელი წერტილი ერთმანეთს ერწყმის, თუ ამ წერტილისა და არის საშუალო სიკაშკაშის მნიშვნელობებს შორის სხვაობა ნაკლებია ზღურბლის დასაშვებ მნიშვნელობაზე. თვით ზღურბლის მნიშვნელობის შეფასება კი ხდება გამოსახულების შესახებ არსებული აპრიორული ინფორმაციის საფუძველზე.

13.8 სურათზე წარმოდგენილია მაგალითი, რომელშიც ზღურბლი გამოყენებულია არეების ჯერ შერწყმის და შემდეგ დაგროვების ოპერაციების შესრულების შემდეგ. ზღურბლის მნიშვნელობად გამოყენებულია არეში სიკაშკაშეების მნიშვნელობების აბსოლუტური სხვაობის (T) ორი მნიშვნელობა. 13.8,ა სურათზე მოცემულია საწყისი გამოსახულება, რომელზედაც გვაქვს ორი საწყისი წერტილი – 1 და 7, ხოლო 13.8,ბ,გ სურათებზე კი ასეთი დამუშავების შედეგი შესაბამისად $T = 4$ და $T = 8$ მნიშვნელობებისთვის.

0	0	5	6	7
1	1	5	8	7
0	1	6	7	7
2	0	7	6	6
0	1	5	6	5

a	a	b	b	b
a	a	b	b	b
a	a	b	b	b
a	a	b	b	b
a	a	b	b	b

a	a	a	a	a
a	a	a	a	a
a	a	a	a	a
a	a	a	a	a
a	a	a	a	a

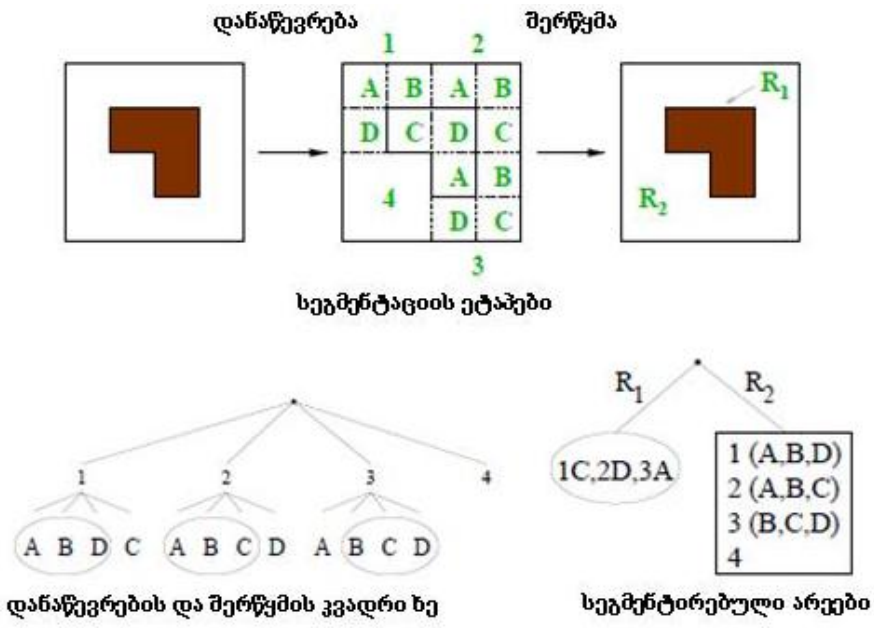
ა
ბ
გ

სურ. 13.8

ამავე დროს, არეების დაგროვების პროცესი შეიძლება განხორციელდეს შერწყმის საწინააღმდეგო მიდგომითაც. ამისათვის საწყის გამოსახულებას განიხილავენ როგორც საწყის სეგმენტს და ახდენენ მის თანამიმდევრულ დანაწევრებას კვადრანტებად (მეოთხედებად). თითოეულ კვადრანტს ამოწმებენ ერთგვაროვნების პირობაზე. იმ შემთხვევაში, თუ პირობა არ სრულდება, მაშინ კვადრანტი ისევ იშლება კვადრანტებად და ეს პროცესი გრძელდება მანამ, სანამ არ მოხდება მთელი გამოსახულების დანაწევრება ერთგვაროვან კვადრანტებად. ამის შემდეგ იწყება მიღებული მომიჯნავე კვადრანტების შერწყმა მათი მსგავსების საფუძველზე უფრო დიდი ზომის ერთგვაროვანი არეების მისაღებად. ერთგვაროვნების შეფასება ხდება სხვადასხვა ხერხებით, მაგალითად სეგმენტში სიკაშკაშის მნიშვნელობების გაფანტულობით, ნახევარტონების დონეების მაქსიმალურ და მინიმალურ მნიშვნელობებს შორის სხვაობის მცირე მნიშვნელობით და ა.შ. თუმცა შერწყმის კრიტერიუმად შეიძლება გამოყენებულ იქნეს სხვა მაჩვენებელიც, მაგალითად, დასამუშავებელი გამოსახულებების კლასის შესახებ სემანტიკური ინფორმაცია. ეს იმას ნიშნავს, რომ არეების დაგროვების პროცესზე შეიძლება გავლენა იქონიოს გამოსახულების გამომსახველობითი ხასიათის ინფორმაციამაც, როგორცაა, მაგალი-

თად, ნახევარტონების დონის სიდიდე, არის ფორმა, საზღვრის სიმკვეთრე.

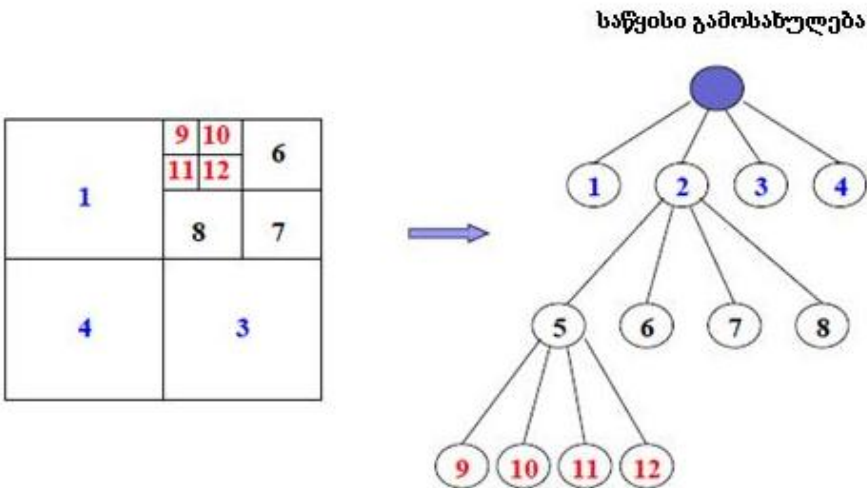
განსხვავებულია ერთმანეთისგან დანაწევრების ხერხებიც. მაგალითად, არსებობს ალგორითმები, სადაც არე ნაწევრდება მეოთხედებად მანამ, სანამ ამ უკანასკნელთა შორის არ მიიღწევა საკმარის მცირე სიდიდის განსხვავება. 13.9 სურათზე წარმოდგენილია სეგმენტაციის ეტაპები გამოსახულების დანაწევრებისა და შერწყმის მიდგომის გამოყენების შემთხვევაში. არის დანაწევრების პროცესი ხორციელდება კვადრი ხის გამოყენებით.



სურ. 13.9

კვადრი ხის გამოყენების მეორე მაგალითი მოცემულია 13.10 სურათზე. როგორც სურათიდან ჩანს, მეორე კვადრანტმა

ვერ დააკმაყოფილა ერთგვაროვნების კრიტერიუმი, ამიტომ მოხდა მისი ისევ დაყოფა კვადრანტებად. შედეგად მიღებული მე-5, მე-6, მე-7, მე-8 კვადრანტიდან მხოლოდ მეხუთემ ვერ დააკმაყოფილა ერთგვაროვნების კრიტერიუმი, ამიტომ ის ისევ დაიყო კვადრანტებად. მიღებულმა მე-9, მე-10, მე-11, მე-12 კვადრანტმა დააკმაყოფილა ერთგვაროვნების კრიტერიუმი და ამით დანაწევრებისა და შერწყმის პროცესი დასრულდა.



სურ. 13.10

არეების დაგროვების მეთოდის ალგორითმების ხასიათი თავისი მოქმედების ბუნებით მიმდევრობითია. ასეთი ალგორითმები უფრო გამოსადეგია ლოკალურ-კონტრასტული წარმონაქმნების გამოსაყოფად. გამოსახულებები აქ მუშავდება იტერაციულ რეჟიმში, რაც მოითხოვს დროის დიდ დანახარჯს და მეხსიერებას. ამავე დროს არ არსებობს ზღურბლის განსაზღვრის თეორიული დასაბუთება. ალგორითმების შედეგად მიღებული ერთგვაროვანი არეები დამოკიდებულია შერწყმის კა-

ნონზე. ამის გამო დაგროვების ალგორითმები კლასტერიზაციის ალგორითმებთან შედარებით უფრო მეტად ემპირიულია.

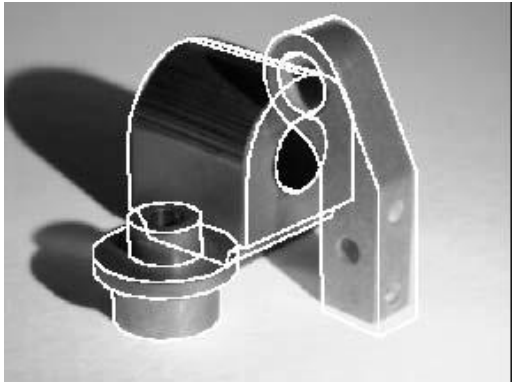
13.4. არეების წარმოდგენის ხერხები

გამოსახულებაზე არეების გამომყოფ ალგორითმებში აუცილებლად გათვალისწინებული უნდა იყოს თითოეული ამ არისტვის დამახასიათებელი თვისებების კომპაქტურად წარმოდგენისა და შენახვის შესაძლებლობა, გამოსახულების ანალიზისთვის მათი შემდგომი გამოყენების მიზნით. გამოყოფილი არეების წარმოდგენისთვის არსებობს სხვადასხვა ხერხები, მათ შორის: საწყის გამოსახულებაზე ზედდების, მარკირებული გამოსახულებები, საზღვრების კოდირება, მონაცემთა გარკვეული სტრუქტურების გამოყენება კვადრანტული (კვადრი) ხეებისა და მახასიათებელ თვისებათა ცხრილის სახით. ჩამოთვლილი ხერხებიდან არეების წარმოდგენისთვის ყველაზე უფრო ხშირად მაინც მარკირებული გამოსახულებები გამოიყენება. განვიხილოთ თითოეული ეს ხერხი ცალ-ცალკე.

13.4.1. საწყის გამოსახულებაზე ზედდების ხერხი

ზედდების (overlay) ხერხის ქვეშ იგულისხმება გამოსახულებაზე გამოყოფილი არეების ისეთი წარმოდგენა, როდესაც საწყის გამოსახულებაზე ერთი ან რამდენიმე ფერის ზედდებით ხდება ჩვენთვის საინტერესო ინფორმაციის ასახვა. ასეთი მიდგომა გამოყენებულია ბევრ გამოსახულების დამუშავების პროგრამულ პაკეტში. ჩვეულებრივ საწყისი გამოსახულება ნახევარტონურია, ხოლო ზედდების ფერად კი გამოიყენება ისეთი ფერი, რომელიც გამოსახულებაზე კარგად გამოჩნდება. ამასთან ეს ორი გამოსახულება მუშავდება ცალ-ცალკე და

ერთმანეთთან კავშირი არ აქვს. მაგალითად, სეგმენტაციის შედეგების სადემონსტრაციოდ გამოყოფილი არეების სასახლო-რო პიქსელებს (კონტურებს) შეიძლება სიკაშკაშის მნიშვნელობებად თეთრი ფერი მივანიჭოთ. ხშირად არეების საზღვრების თვალსაჩინოებისთვის კონტურების სისქე ერთ პიქსელზე მეტია. 13.11 სურათზე წარმოდგენილია რამდენიმე კარკასული ტიპის დეტალისგან შედგენილი სცენის გამოსახულება, რომელზედაც დეტალების საინსპექციო კონტროლისათვის ამოსაცნობი ობიექტების მოდელების პროექციები ზემოდანაა დადებული საწყის ნახევარტონურ გამოსახულებაზე.



სურ. 13.11

13.4.2. მარკირებული გამოსახულებები

გამოსახულების მარკირება გამოსახულების ანალიზის პროცესის ერთ-ერთ მნიშვნელოვან შუალედურ ნაბიჯს წარმოადგენს. მისი არსი მდგომარეობს თითოეული გამოყოფილი არისთვის უნიკალური იდენტიფიკატორის (ჭდის) (ჩვეულებრივად წარმოადგენს მთელ რიცხვს, მაგრამ შეიძლება სიმბოლოც იყოს) მინიჭებაში. ამის შესახებ უკვე აღვნიშნეთ ბმული კომპონენ-

ტების მარკირებისას (IV თავი). იმ შემთხვევაში, თუ ჭდეების მნიშვნელობები მცირე სიდიდეებს წარმოადგენს, მაშინ ნახევარტონური გამოსახულების სახით ასახვისას მარკირებული გამოსახულება საკმაოდ ბნელი იქნება. ამის თავიდან ასაცილებლად, ანუ ვიზუალურად უკეთესი გამოსახულების მისაღებად შეიძლება მარკირებული გამოსახულების სიკაშკაშის მნიშვნელობების დინამიკური დიაპაზონის გაჭიმვა ან ჰისტოგრამის გაფართოების მეთოდის გამოყენება.

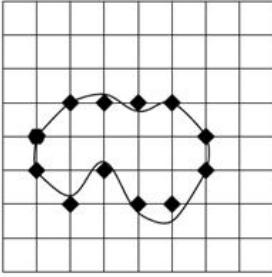
13.4.3. საზღვრების კოდირება

გამოსახულებაზე გამოყოფილი არეების წარმოდგენისთვის შეიძლება გამოყენებული იქნეს მათი საზღვრები. ამასთან საზღვრების შესახებ ინფორმაცია დევს არა უშუალოდ გამოსახულებაზე, როგორც ეს ზედღების ხერხის გამოყენებისას ხდება, არამედ მონაცემთა გარკვეულ სტრუქტურაში. ასეთი სტრუქტურის ერთ-ერთ უმარტივეს მაგალითად შეიძლება მოვიყვანოთ არის სასაზღვრო პიქსელებისგან შედგენილი წრფივი სია. განვიხილოთ გამოყოფილი არის კონტურის კომპაქტური წარმოდგენისთვის ასეთი სიის ნაირსახეობა, რომელსაც **ფრიმანის ჯაჭვური კოდი** ეწოდება (Freeman chain code). სასაზღვრო წერტილების სიის ჯაჭვურ კოდში გარდაქმნისას ხდება ამ სიიდან აღებული ინფორმაციის კიდევ ერთხელ დამატებითი დისკრეტიზაცია. განვიხილოთ ახლა ჯაჭვური კოდირების მთავარი პრინციპი. კოდირებისთვის განკუთვნილი სასაზღვრო მრუდი იდება კვადრატულ ბადეზე. ამ ბადის უჯრედების ზომები განსაზღვრავს კოდირების სივრცით გარჩევითობას. ამის შემდეგ სასაზღვრო მრუდზე აირჩევა საწყისი წერტილი და, დაწყებული ამ წერტილიდან, კვადრატულ ბადეზე ხდება საკვანძო წერტილების განსაზღვრა. ასეთ წერტილებად ბადეზე აირჩევა

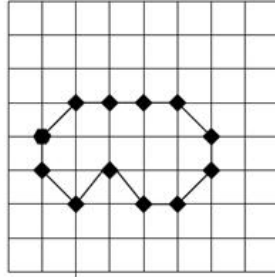
ბადის ის წერტილები, რომლებიც ყველაზე ახლოსაა განლაგებული სასახლვრო მრუდთან. აღნიშნული საკვანძო წერტილები განიხილება როგორც იმ ტეხილის წვეროები, რომლითაც სასახლვრო მრუდის აპროქსიმაცია ხდება. ამ ტეხილის თითოეული გვერდი წარმოადგენს მონაკვეთს, რომელიც ერთმანეთთან აერთებს ბადის საკვანძო წერტილს მის ერთ-ერთ მეზობელთან. ამ მონაკვეთების მიმართულების კოდირება 8-ბმულობის შემთხვევაში ხდება მცირე სიდიდის მთელი რიცხვებით, დაწყებული 0-დან და დამთავრებული კოდირებისთვის გამოყენებული მომიჯნავე მეზობლების მიმართულებების რაოდენობის რიცხვამდე, ანუ 7-მდე. ტეხილი წირის თითოეული ასეთი მონაკვეთის კოდების ერთობლიობა ქმნის ჯაჭვურ კოდს.

ჯაჭვური კოდი ბინარული გამოსახულების საზღვრისთვის წარმოადგენს მთელი $c = \{c_0, c_1, \dots, c_{n-1}\}$ რიცხვების მიმდევრობას, სადაც $c_i, i = \overline{0, n-1}$ მნიშვნელობებს იღებს $\{0, 1, \dots, 7\}$ სიმრავლიდან. ელემენტების რაოდენობას c მიმდევრობაში ეწოდება ჯაჭვური კოდის სიგრძე. c_0 და c_{n-1} ელემენტები წარმოადგენს კოდის საწყის და ბოლო წერტილებს.

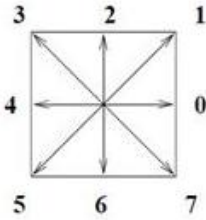
13.12 სურათზე წარმოდგენილია ჯაჭვური კოდირების მაგალითი, რომელშიც წრფივი მონაკვეთების დახრის კუთხეების (მიმართულებების) კოდირებისთვის გამოყენებულია 8-ბმული ჯაჭვური კოდი, ანუ კოდირებისთვის განიხილება არა მარტო ჰორიზონტალური და ვერტიკალური, არამედ დიაგონალური მიმართულებებიც. კოდის რიცხვითი მნიშვნელობა 0 შეესაბამება 0^0 -ით მიმართულ მონაკვეთს, მნიშვნელობა 1 – 45^0 -ით მიმართულ მონაკვეთს, და ასე შემდეგ, მნიშვნელობა 7 – 315^0 -ით მიმართულ მონაკვეთს. სურათზე სასახლვრო მრუდის საწყისი წერტილი აღნიშნულია ექვსკუთხედით, ხოლო ბადისა და მრუდის გადაკვეთის წერტილები კი – რომბებით.



სასაზღვრო მრუდი



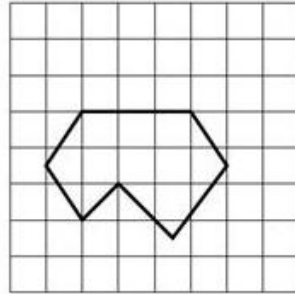
ჯაჭვური კოდი



კოდირების სქემა

100076543532

მრუდის ჯაჭვური კოდი



პოლიგონალური აპროქსიმაცია

სურ. 13.12

სასაზღვრო მრუდის საწყისი წერტილის კოორდინატები და ჯაჭვური კოდი სრულიად საკმარისია ამ მრუდის აღსადგენად ბადის სივრცითი გარჩევის მიხედვით. ჯაჭვური კოდი, გარდა იმისა, რომ მკვეთრად ამცირებს საზღვრის მონაცემების შენახვისათვის საჭირო მეხსიერების მოცულობას, ასევე შეიძლება გამოყენებული იყოს მრუდების დამუშავების ოპერაციებში, მაგალითად, ობიექტების ფორმის მიხედვით ამოცნობისას. თუ გამოყოფილ არეს აქვს არა მარტო გარე საზღვარი, არამედ შეიცავს ერთ ან რამდენიმე ნახვრეტს, მაშინ თითოეული მათ-

განი ცალ-ცალკე შეიძლება წარმოვადგინოთ ჯაჭვური კოდის სახით.

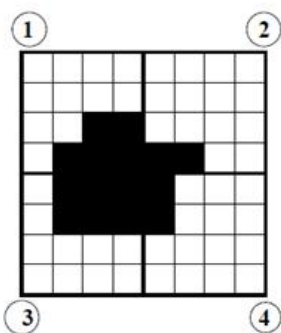
იმ შემთხვევაში, თუ გამოსახულებაზე გამოყოფილი არის მესხიერებაში შესანახად საჭირო არ არის მისი ყველა სასახლდრო პიქსელის კოორდინატების შენახვა, მაშინ მრუდი შეიძლება ასევე მიახლოებით წარმოვადგინოთ მონაკვეთების სახით, მაგრამ განსხვავებით ჯაჭვური კოდირების მეთოდისგან, ამ შემთხვევაში მონაკვეთების წვეროების კოორდინატები ნამდვილი რიცხვებია და არ არის აუცილებელი, რომ ისინი ბადის კვანძებს ემთხვეოდეს. გამოსახულებაზე გამოყოფილი არის სასახლდრო მრუდის წარმოდგენის ამ მეთოდს **პოლიგონალური აპროქსიმაცია** ეწოდება (სურ. 13.12). აღნიშნული მიდგომა ამარტივებს სახდურების დამუშავებისთვის საჭირო გამოთვლებს და იძლევა მესხიერების ეკონომიის საშუალებას.

13.4.4. კვადრი ხეები

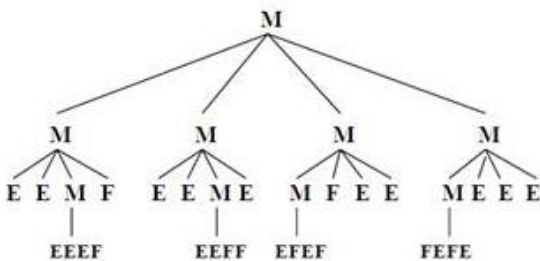
ციფრულ გამოსახულებაზე გამოყოფილი არეების კომპაქტური წარმოდგენის კიდევ ერთ ხერხს წარმოადგენს **კვადრანტული ანუ კვადრი ხეები** (quad tree). ამ შემთხვევაში არე აღიწერება არა მარტო მისი სახდურის შესახებ მონაცემების საშუალებით, არამედ მთლიანობაში. კვადრი ხის ყოველი კვანძი შეესაბამება გამოსახულებაზე არსებულ რაღაც კვადრატულ არეს. ამასთან კვანძის აღნიშვნისთვის გამოიყენება სამიდან ერთ-ერთი ჭდე. კერძოდ, ეს ჭდეებია: *შევსებულია* (full, F), *ცარიელია* (empty, E), *ნაწილობრივ შევსებულია* (mixed, M). თუ კვანძი მონიშნულია შევსებულის აღმნიშვნელი ჭდით, მაშინ კვადრატული არის თითოეული პიქსელი მიეკუთვნება ინტერესის არეს. თუ კვანძი მონიშნულია ცარიელის აღმნიშვნელი ჭდით, მაშინ შესაბამისი კვადრატული არე და ინტერესის არე ერთმანეთს

არ კვეთს, ხოლო თუ არე მონიშნულია როგორც ნაწილობრივ შევსებული, მაშინ კვადრატული არის პიქსელების მხოლოდ ნაწილი წარმოადგენს ინტერესის არის პიქსელებს, დანარჩენები კი არა. კვადრანტულ ხეზე შეიღობილი კვანძები მხოლოდ ნაწილობრივ შევსებულ კვანძებს გააჩნია, რაც შეეხება შევსებულ და ცარიელ კვანძებს, ისინი ფოთლოვანებია.

გამოსახულებაზე გამოყოფილი (8×8) პიქსელი განზომილების მქონე არის კვადრი ხის საშუალებით წარმოდგენის შემთხვევა ნაჩვენებია 13.13 სურათზე.



არე გამოსახულებაზე



კვადრი ხე

სურ. 13.13

როგორც ვხედავთ, არის საზღვრები არ წარმოადგენს სწორხაზოვან მონაკვეთებს. ამის გამო შესაბამისი კვადრი ხე შედგება 4 დონისაგან. სურათზე წარმოდგენილი 4 შეიღობილი კვანძი შეესაბამება ზედა მარცხენა, ზედა მარჯვენა, ქვედა მარცხენა და ქვედა მარჯვენა კვადრანტებს. ხის პირველი დონისთვის კვადრანტების გადათვლის რიგითობა აღნიშნულია წრესხაზებში ჩაწერილი რიცხვების საშუალებით.

13.4.5. თვისებათა ცხრილები

გამოსახულებათა ანალიზის ზოგიერთ ამოცანაში გამოყოფილი არის წარმოდგენა საჭიროა არა მისი შემადგენელი პიქსელების, არამედ გამოყოფილი ნიშან-თვისებების საშუალებით. სწორედ არის ასეთ წარმოდგენას უწოდებენ **თვისებათა ცხრილს** (property table). ასეთი ცხრილი ჰგავს რელაციურ მონაცემთა ბაზის ცხრილს. კერძოდ, გამოსახულების თითოეულ არეს ცხრილში შეესაბამება ცხრილის სტრიქონი, ხოლო თითოეულ თვისებას – სვეტი. არის თვისებებად შეიძლება გამოვიყენოთ მისი გეომეტრიული თვისებები – ზომა, ფორმა, ასევე ინტენსივობა, ფერი, ტექსტურა და ჩვენ მიერ წინა თავებში უკვე განხილული არის ყველა მახასიათებელი ნიშან-თვისება.

13.5. საზღვრების ანალიზის მეთოდი

გამოსახულებათა ანალიზის ზოგიერთი ამოცანა მოითხოვს არა ერთგვაროვანი არეების, არამედ მათი საზღვრების ან ზოგიერთი სხვა სტრუქტურის, მაგალითად, სწორხაზოვანი მონაკვეთების ან წრეწირების რკალების გამოყენებას. აქ განვიხილავთ სწორედ ასეთი სტრუქტურების მოძიებასთან დაკავშირებულ საკითხებს. საზღვრების ანალიზის მეთოდი ემყარება გამოსახულების იმ განსაკუთრებულ თვისებას, რომ გამოსახულების შესახებ ყველაზე მეტი ინფორმაცია დევს სხვადასხვა არაერთგვაროვანი არეების საზღვრებზე, ანუ იქ, სადაც ადგილი აქვს სიკაშკაშეების მკვეთრ ვარდნას, სხვანაირად რომ ვთქვათ, ერთგვაროვნობის წყვეტას. საზღვრების ანალიზის ამოცანა მოიცავს ორ ეტაპს: კიდის ელემენტების მოძიება და შეპირაპირება.

13.5.1. არეების საზღვრების მიღვენება

გამოსახულებაზე სეგმენტაციის ან ბმული კომპონენტების მარკირების შედეგად მიღებული არეების გამოყოფის შემდეგ შესაძლებელია თითოეული ამ არის საზღვრის მოძებნა. მეთოდს, რომელიც ამ ამოცანის გადასაწყვეტად გამოიყენება, **საზღვრების გამოყოფა (boundary extraction)** ეწოდება. გამოსახულების სეგმენტაცია საზღვრების გამოყოფის ხერხის გამოყენებით შეიძლება განხორციელდეს ორი გზით: მიმდევრობითი და პარალელური. პირველ შემთხვევაში მთელ გამოსახულებაზე იმ წერტილების პოვნის ხარჯზე, რომელთა მიდამოშიც ადგილი აქვს სიკაშკაშის მკვეთრ ვარდნებს, ხდება საზღვრების ფრაგმენტების გამოყოფა. მეორე შემთხვევაში, დაწყებული რომელიღაც წერტილიდან, თანამიმდევრულად ხდება ამავე საზღვარზე არსებული მომიჯნავე წერტილების მოძიება. ხშირად საზღვრების მოძიებული მონაკვეთები, იმის გამო, რომ მათი ნაწილი შეიძლება მცდარად იყოს გამოყოფილი, ხოლო ნაწილი კი გამოტოვებული, არ ქმნის შეკრულ კონტურს. აქედან გამომდინარე, საჭირო ხდება საზღვრების წვეტების ადგილების შევსება. ამისთვის შემუშავებულია სპეციალური მეთოდები, რომელთაც **შეპირაპირების (ან გადაბმის) მეთოდები** ეწოდება. კიდის ელემენტების შეპირაპირების არსი მდგომარეობს საზღვრის ბმული ელემენტების შეერთებაში და მათ შევსებაში უწყვეტი საზღვრის მიღებამდე. ამასთან პროცესის მიმდინარეობის დროს პარალელურად უნდა მოხდეს მცდარი საზღვრების გაცხრილვა. ამ ამოცანის შესასრულებლად გამოიყენება ევრი-სტიკული ძიების და დინამიკური პროგრამირების მეთოდები, წირებისა და მრუდეების შეპირაპირებისა და რელაქსაციის (აღდგენის) მეთოდები.

არეების საზღვრების გამოყოფისთვის იყენებენ **საზღვრების მიდევნების (border tracing)** ალგორითმებს. ამ სახის მარტივი ალგორითმი ასეთია: მოვახდინოთ გამოსახულების ყველა წერტილის გადარჩევა და თითოეული ბმული კომპონენტის (არის) პირველი სასაზღვრო წერტილებისგან შევადგინოთ მათი კოორდინატების სია. ამის შემდეგ თითოეული არისთვის, დაწყებული პირველი სასაზღვრო წერტილიდან, მოვახდინოთ ბმული კომპონენტის საზღვრის მიდევნება საათის ისრის მიმართულებით მანამ, სანამ ისევ არ დავბრუნდებით საწყის წერტილში. ეს ალგორითმი ემყარება იმ დაშვებას, რომ ყოველი საზღვარი შემოსაზღვრავს შეკრულ არეს. ასეთი ალგორითმი ადვილი გამოსაყენებელია მცირე ზომის გამოსახულებისთვის. მაგრამ დიდი გამოსახულების შემთხვევაში, რომელიც მთლიანად შეიძლება არ მოთავსდეს მეხსიერებაში, ასეთი ალგორითმი არ გამოდგება, რადგანაც საზღვრის მიდევნების პროცესში შეიძლება საჭირო გახდეს მრავალჯერადი მიმართვა შეყვანა-გამოყვანის მოწყობილობისადმი, სადაც ინახება საწყისი გამოსახულება.

განვიხილოთ ალგორითმი, რომელიც ახდენს არეების საზღვრების გამოყოფას გამოსახულების მარცხნიდან-მარჯვნივ, ზემოდან-ქვემოთ მიმართულებით დათვალიერების გზით. აქ საწყის მონაცემებად გამოიყენება მარკირებული გამოსახულება, ხოლო გამოსასვლელ მონაცემებს უნდა წარმოადგენდეს თითოეული არის საზღვარი, წარმოდგენილი საათის ისრის მიმართულებით მოწესრიგებულ წერტილთა კოორდინატების სის სახით. მარკირებული გამოსახულების პიქსელების მნიშვნელობები, რომელთაც ალგორითმი იყენებს როგორც შესასვლელ მონაცემებს, შეესაბამება არეების ჭდეებს. ივარაუდება, რომ გამოსახულებაზე გვაქვს მხოლოდ ერთი ფონური არე. მისი ჭდის მნიშვნელობა ენიჭება იმ არეებს, რომელთა საზღვრე-

ბის გამოყოფა არ მოითხოვება. ალგორითმი, იმის ნაცვლად, რომ მოახდინოს თითოეული არის საზღვრის მიდევნება, გამოსახულების პიქსელებს ამუშავებს მარცხნიდან-მარჯვნივ, ზემოდან-ქვემოთ და ახდენს სასაზღვრო პიქსელების ჯაჭვების ფორმირებას, რომლებიც თავის მხრივ ქმნის არეების საზღვრების ბმულ ფრაგმენტებს. ალგორითმის მუშაობის ნებისმიერ მომენტში გამოსახულებაზე გვექნება მიმდინარე არეების (current regions) სიმრავლე, რომელთა საზღვრები ნაწილობრივ, მაგრამ ჯერ კიდევ არასრულად იყო გამოყოფილი, დამუშავებული არეების (past regions) სიმრავლე, რომელთა საზღვრები სრულად გამოიყო და მომავალი არეების (future regions) სიმრავლე, რომლებიც გამოსახულების პიქსელების გადარჩევის პროცესში ჯერ არ გამოვლენილა. მიმდინარე არეების სასაზღვრო პიქსელების ჯაჭვები ინახება მონაცემთა სტრუქტურებში. თუ არე სრულად არის დამუშავებული, ანუ საზღვარი სრულად გამოიყო და მისი საზღვრის შესახებ მონაცემები მოთავსებულია გამოსასვლელ მონაცემებში, მაშინ ამ არის შესახებ არსებული ინფორმაცია ჰემ-ცხრილიდან ამოიღება. როცა გამოსახულების ალგორითმით დამუშავებისას გვხვდება ახალი არე, მაშინ მის შესახებ ინფორმაცია ჩაიწერება ჰემ-ცხრილში. მოცემული არის შესახებ ჰემ-ცხრილში არსებული ჩანაწერი შეიცავს მოცემული არის შესახებ დროის მოცემული მომენტისთვის ფორმირებული ბმული ჯაჭვების სიის შესახებ მახვენებელს. თითოეული ჯაჭვი წარმოადგენს სასაზღვრო პიქსელების კოორდინატების ბმულ სიას, დალაგებულს მათი აღმოჩენის მიხედვით. საზღვრის მიდევნების აღნიშნულ ალგორითმს გამოსახულების დამუშავების თითოეულ ნაბიჯზე საქმე აქვს მარკირებული გამოსახულების სამ სტრიქონთან. ეს არის დამუშავებაში მყოფი მიმდინარე სტრიქონი, და ასევე სტრიქონები მის ზემოთ და ქვემოთ. გამოსახულების ზედა და ქვედა კიდურა სტრიქონების

შემთხვევაში მათ ემატება ფონური პიქსელების თითო ფიქტიური სტრიქონი. თითოეული სტრიქონის დამუშავების შემდეგ იმ მიმდინარე არეების ჯაჭვები, რომელთა საზღვრების სრულად გამოყოფა უკვე დასრულდა, ერთიანდება ერთ მთლიან საზღვრად და თავსდება გამოსასვლელ მონაცემებში, ხოლო ჰეშ-ცხრილიდან და არეების კოორდინატების სიიდან კი ამოიღება ინფორმაცია სრულად დამუშავებული არეების შესახებ. ასეთი ალგორითმის საშუალებით არეების საზღვრების გამოყოფის პროცედურა მარკირებული გამოსახულებისთვის ნაჩვენებია 13.14 სურათზე, სადაც (ა) მარკირებული გამოსახულებაა, ხოლო (ბ) საზღვრების გამოყოფის პროცედურის შედეგები.

	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0
2	0	0	0	0	2	2	0
3	0	1	1	1	2	2	0
4	0	1	1	1	2	2	0
5	0	1	1	1	2	2	0
6	0	0	0	0	2	2	0
7	0	0	0	0	0	0	0

ა

არე	სიგრძე	არის კოორდინატების სია
1	8	(3,2)(3,3)(3,4)(4,4)(5,4)(5,3)(5,2)(4,2)
2	10	(2,5)(2,6)(3,6)(4,6)(5,6)(6,6)(6,5)(5,5) (4,5)(3,5)

ბ

სურ. 13.14

პარალელური სეგმენტაციის ხერხის გამოყენებისას კიდის ელემენტების საპოვნელად გამოიყენება:

- **მაღალსიხშირული სივრცითი ფილტრაცია**, რომელიც ეფუძნება იმ ფაქტს, რომ სპექტრის მაღალსიხშირული მდგენელები (შეესაბამება სიკაშკაშის მკვეთრ ვარდნებს) გამოსახულებაზე გვაძლევს ინფორმაციას ობიექტების საზღვრების შესახებ. ამ მიზნით ხდება გამოსახულების ინტენსივობის ფუნქციის ფურიე-გარდაქმნა, რაც შედეგად გვაძლევს სივრცითი კოორდინატების გარდაქმნას სიხშირეთა სპექტრში, რომლის მაღალსიხშირული მდგენელები შეესაბამება გამოსახულებაზე არსებული ობიექტების მკვეთრ კიდეებს, ხოლო დაბალსიხშირული მდგენელები კი – თითქმის ერთნაირი ნახევარტონების მქონე არეებს. სპექტრის მაღალსიხშირული მდგენელების გამოყოფისა და შემდეგ ფურიე უკუაქსახვის განხორციელების შედეგად, ანუ ისევ სივრცით კოორდინატებში გადასვლით, მიიღება გამოსახულება, რომელზედაც კიდეები მკვეთრად ხაზგასმული იქნება. ასეთი გამოსახულების ზღურბლური დამუშავებით კი მიიღება ბინარული კონტურული გამოსახულება.
- **გრადიენტული ოპერატორი**, რომელიც ეფუძნება გამოსახულების იმ თვისებას, რომ კიდის წერტილებში ინტენსივობის ფუნქციის გრადიენტის მოდულის სიდიდე უფრო მეტია. აქედან გამომდინარე, გამოსახულებაზე გრადიენტული გარდაქმნის გამოყენების შედეგად მიიღება გრადიენტული გამოსახულება, რომელიც ზღურბლური დამუშავების შედეგად გვაძლევს ბინარულ კონტურულ გამოსახულებას, რომლის ერთეულოვანი ელემენტები განსაზღვრავს სიკაშკაშის ფუნქციის საზღვრებს. გრა-

დიენტის მოდულის სიდიდე განისაზღვრება გამოსახულების მეზობელი ელემენტების მნიშვნელობებს შორის დისკრეტული სხვაობების გამოთვლით როგორც სტრიქონის, ასევე სვეტის გასწვრივ. ეს გვაძლევს საშუალებას, რომ გამოსახულებაზე ვიპოვოთ შესაბამისად ვერტიკალური და ჰორიზონტალური ვარდნები. მაგრამ ასეთი დისკრეტული დიფერენცირება ხასიათდება ხმაურის მიმართ მგრძობიარობით. ამ უკანასკნელის შესასუსტებლად მიმართავენ გამოსახულების დაგლუვებას, რისთვისაც იყენებენ ე.წ. **არეების შენიღბვას (masking area)**, რაც გულისხმობს სხვაობების აღებას არა მეზობელ ელემენტებს, არამედ ელემენტების მეზობელ ჯგუფებს შორის. ასევე გამოიყენება ჩვენ მიერ VII თავში განხილული დისკრეტული დიფერენცირების არაწრფივი მეთოდები, კერძოდ, რობერტსის ჯვარედინი ოპერატორი, სობელის ოპერატორი და სხვ.

- **ფუნქციონალური აპროქსიმაცია**, რომელიც არეების საზღვრების გამოსაყოფად იყენებს ოპტიმიზაციის ალგორითმებს. ვიცით, რომ იდეალური ვარდნა განიხილება როგორც საფეხუროვანი სიგნალი (სურ. 7.1,ა). აქედან გამომდინარე, რეალური გამოსახულების ფრაგმენტი შეიძლება აპროქსიმირებული იქნეს ცვლადი პარამეტრების მქონე იდეალური ვარდნით. თუ მოცემული ფრაგმენტისთვის ასეთი აპროქსიმაციის ჩატარება საკმაოდ ზუსტად მოხდება, შეიძლება ჩაითვალოს, რომ ამ ადგილას ვარდნას ადგილი აქვს აპროქსიმაციის მოძიებულ პარამეტრებით.

მიმდევრობითი სეგმენტაციის ხერხს კი მიეკუთვნება საზღვრის მიდევნების მეთოდები. აქ საზღვრის მონაკვეთების მო-

ძიება ხორციელდება მიმდევრობითი ძიების გზით და მოიცემა სამი წესის საშუალებით:

- საწყისი წერტილის არჩევა;
- მეზობელი წერტილის მიერთება;
- მიღვენების პროცესის შეჩერება.

საწყისი წერტილის განსაზღვრა შეიძლება სხვადასხვა ხერხით განხორციელდეს, მაგალითად ობიექტის სიკაშკაშის მნიშვნელობებს შორის შუალედური მნიშვნელობით, ექსპერტის მითითებით. მიღვენების პროცესის შეჩერება ხდება მაშინ, როცა ვერ ხერხდება იმ შემდეგი წერტილის მოძიება, რომელიც აკმაყოფილებს მიერთების პირობას. აღსანიშნავია, რომ მიერთება ხორციელდება თითო ელემენტის დონეზე და ამასთან ისეთის, რომელიც უზრუნველყოფს საზღვრის შეკვრას, მაგრამ არა მის თვითგადაკვეთას. ამასთან, საზღვრის ყოველი შემდეგი წერტილის მიერთებისას ჩნდება შესაძლებლობა, რომ გათვალისწინებულ იქნეს როგორც ლოკალური, ისე გლობალური მონაცემები, და ასევე აღნიშნული საზღვრის აგებისას დაგროვილი გამოცდილება.

საზღვრის გამოყოფის პროცესისთვის ძალიან მნიშვნელოვანია ზღურბლის სიდიდის განსაზღვრა. მისი მნიშვნელობა ისე უნდა შეირჩეს, რომ მან, ერთი მხრივ, ხელი შეუწყოს საზღვრის რაც შეიძლება მეტი ჭეშმარიტი მონაკვეთების და, მეორე მხრივ, რაც შეიძლება ნაკლები მცდარი მონაკვეთების გამოყოფას. ზღურბლის სიდიდის შერჩევა ხშირად ემპირიულად ხდება, თუმცა არსებობს მისი შერჩევის ავტომატური ხერხი, დამყარებული გამოსახულების გლობალურ თვისებებზე და ასევე პისტოგრამის ანალიზის შედეგად. იმის გამო, რომ საზღვრების გამოყოფის მეთოდი ემყარება გამოსახულებაზე არსე-

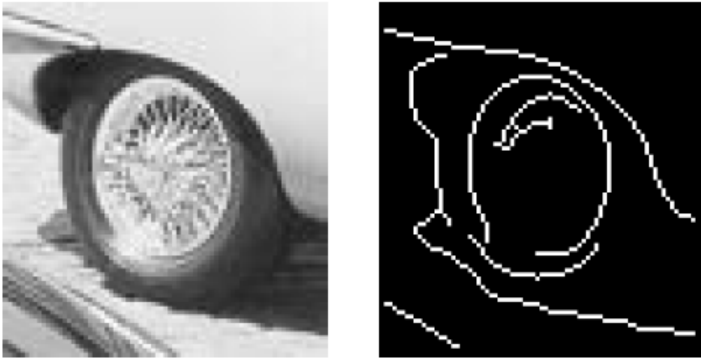
ბული ლოკალური არაერთგვაროვნობების ანალიზს, შესაბამისად, ბუნებრივია მათი მგრძობიარობა ხმაურის მიმართ. შედეგად ბული ელემენტების სიმრავლისაგან ფორმირებული გამოყოფილი საზღვრები გამოსახულებაზე შეიძლება მოიცავდეს არაბმულ არეებსაც. ამიტომ, ხშირად საზღვრების ანალიზის მეთოდით გამოსახულების სეგმენტაციამდე მიმართავენ გამოსახულების წინასწარ ფილტრაციას. 13.15,ა,ბ სურათზე წარმოდგენილია შემთხვევა, როცა საზღვრის გამოყოფისთვის გამოყენებული იქნა ქენის ოპერატორი და აგრეთვე ის პრობლემები, რომლებიც საზოგადოდ წარმოიშობა კიდის გამოყოფის და საზღვრის მიდევნების ალგორითმების გამოყენებისას. როგორც სურათიდან ჩანს, ობიექტების ნაწილების კონტურების სეგმენტები მცდარადაა შერწყმული განათებული ან ჩრდილოვანი არეების კონტურებთან. კერძოდ, გამოსახულების (სურ. 13.15,ა) ზედა ნაწილში არსებული სარკული ანარეკლი ხელს უშლის კიდის დეტექტორს მაშუქის ქრომირებული რგოლის საზღვრის ფორმირებაში.



სურ. 13.15,ა

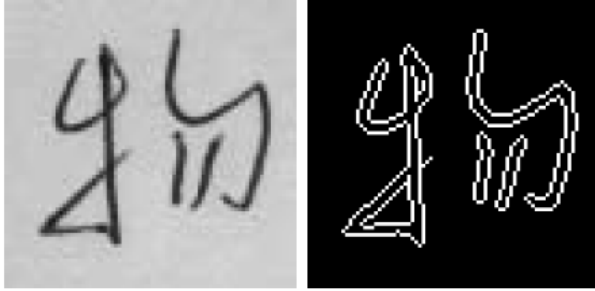
ხოლო ქვედა გამოსახულებაზე (სურ. 13.15,ბ) კი ავტომობილის ჩრდილი ერწყმის საბურავს (რომელიც ასევე შერწყმუ-

ლია ფრთასთან), რის შედეგადაც როგორც საბურავის, ასევე მისი სოლების საზღვრები არასაკმარისად კარგად არის გამოყოფილი.



სურ. 13.15,ბ

წარმოდგენილი შემთხვევიდან ჩანს, რომ გამოყოფილი საზღვრების სეგმენტები ზუსტად არ ემთხვევა გამოსახულების არეებს. ხანდახან ასეთ დამთხვევას ადგილი აქვს და მაშინ არეების აღმოჩენა შესაძლებელია გამოყოფილი საზღვრების შესახებ მონაცემების საფუძველზე. ასეთი დამთხვევის მაგალითი მოცემულია 13.16 სურათზე, სადაც სიმბოლოები ხასიათდება ფონის მიმართ მაღალი კონტრასტით, რამაც გააადვილა გამოსახულებაზე მათი აღმოჩენა. ამის შემდეგ უკვე შესაძლებელია გამოყოფილი საზღვრების სეგმენტების გაანალიზება, თუ მოვახდენთ სასაზღვრო პიქსელების სეგმენტირებას სწორხაზოვან მონაკვეთებად, წრეწირების რკალებად და ა.შ. მაგალითად, სწორკუთხა შენობის საზღვარი შეიძლება დაიყოს ოთხ სწორხაზოვან მონაკვეთად.



სურ. 13.16

13.6. თავსებადი მეზობელი კონტურული ფრაგმენტების დაჯგუფება მრუდებად

ხშირად საზღვრების ანალიზის მეთოდების გამოყენებით ვერ ხერხდება უწყვეტი საზღვრების მთლიანად გამოყოფა, ანუ ზოგიერთ უბანზე შესაძლოა ადგილი ჰქონდეს წვეტებს. ასეთი წვეტების აღსადგენად, როგორც უკვე აღვნიშნეთ, გამოიყენება შეპირაპირების მეთოდები, რომელთათვისაც საწყის გამოსახულებას წარმოადგენს მარკირებული სასაზღვრო ელემენტების მქონე ბინარული გამოსახულება. ამ ამოცანის გადასაწყვეტად, ძირითადად არსებობს ორი გზა: ჯერ აღიწეროს საზღვრის ყველა გამოყოფილი მონაკვეთი და შემდეგ მოხდეს მათი გაანალიზება, ან ვიმუშაოთ საზღვრის გამოყოფილ მონაკვეთებთან როგორც ერთ მთლიან გამოსახულებასთან.

შეპირაპირების პროცესის თავისებურება მდგომარეობს იმაში, რომ საზღვარზე წვეტის ადგილის აღდგენა უნდა მოხდეს სწორად. საჭიროა მხედველობაში მივიღოთ ის, რომ მეზობელი მონაკვეთების შეერთების შედეგად უნდა მივიღოთ საჭირო ფორმის შეკრული საზღვარი. თვით კონტურის ფორმის განსაზღვრისათვის უნდა გამოვიყენოთ გლობალური ინფორმაცია, მაგალითად გამოსახულების ჰისტოგრამა.

იმ შემთხვევაში, თუ საზღვრის ფორმა წინასწარაა ცნობილი, მაშინ მონაკვეთების საპოვნელად შეიძლება გამოვიყენოთ აპროქსიმაცია, შეთანხმებული ფილტრაცია ან ჰაფის გარდაქმნა. თუ წყვეტის ადგილას ელემენტების რიცხვი ორზე მეტია, მაშინ დგება გადარჩევის ამოცანა. ამიტომ წყვეტის შესაბამისი ჩანართის ასარჩევად იყენებენ გამოსახულების შესახებ აპრიორულ ინფორმაციას.

ახლა დაუვშვათ, რომ 13.5.1. პარაგრაფში განხილული კონტურის მიდევნების ალგორითმის შესასვლელს წარმოადგენს არა მარკირებული გამოსახულება, არამედ მარკირებული კონტურული გამოსახულება, რომელზედაც კონტურის პიქსელებს მინიჭებული აქვს მნიშვნელობა 1, ხოლო დანარჩენებს – 0. ასეთ შემთხვევაში არეების საზღვრების ცალკეული სეგმენტების მიდევნების ამოცანა უფრო რთული ხდება. ასეთ დროს უკვე აუცილებელი არაა, რომ შეკრული არეები კონტურული პიქსელებით იყოს შემოსაზღვრული. კონტურების სეგმენტები შედგება ბმული კონტურული პიქსელებისაგან. სეგმენტები შეიძლება იწყებოდეს რომელიმე წვეროდან, კუთხის წერტილიდან ან შეპირაპირების ადგილიდან და მთავრდებოდეს რომელიმე სხვა წვეროზე, კუთხის წერტილზე ან შეპირაპირების ადგილზე. სეგმენტები ამ დროს არ შეიცავენ შუალედურ შეპირაპირების ადგილებს ან კუთხეებს. 13.17 სურათზე წარმოდგენილია მარკირებული კონტურული გამოსახულების მაგალითი. პიქსელი [3,3] წარმოადგენს **შეპირაპირების პიქსელს** (junction pixel), რომელშიც ერთმანეთს უერთდება სამი სხვადასხვა სწორხაზოვანი კონტურული სეგმენტი. პიქსელი [5,3] – კუთხის პიქსელია, რომელიც შეიძლება განხილულ იქნეს როგორც სეგმენტის ბოლო წერტილი იმ შემთხვევაში, თუ მოცემულ ამოცანაში დაშვებულია, რომ სეგმენტები მთავრდებოდეს კუთხის წერტილებით.

	1	2	3	4	5
1	1	0	0	0	1
2	0	1	0	1	0
3	0	0	1	0	0
4	0	0	1	0	0
5	0	0	1	1	1

სურ. 13.17

ასეთი სეგმენტების მიღენების ალგორითმში მოითხოვება შემდეგი ოპერაციების შესრულება:

- ახალი სეგმენტის მიღენება;
- შიდა პიქსელის დამატება სეგმენტისათვის;
- სეგმენტის დასრულება;
- შეპირაპირების ადგილის მოძებნა;
- კუთხის მოძებნა.

ისევე როგორც საზღვრების მიღენების ზემოგანხილული ალგორითმის შემთხვევაში, კონტურების დაჯგუფების პროცედურის ყოველ ნაბიჯზე შესაბამისი ინფორმაციის შესანახად მიზანშეწონილია გამოყენებული იქნეს მონაცემთა ეფექტური სტრუქტურები. ასეთი მონაცემთა სტრუქტურები ზემოაღწერილი საზღვრების მიღენების ალგორითმში გამოყენებული მონაცემთა სტრუქტურების მსგავსია. დამუშავებული, მიმდინარე და მომავალი არეების ნაცვლად აქ გვაქვს დამუშავებული, მიმდინარე და მომავალი კონტურების სეგმენტები. სეგმენტები წარმოადგენს გამოსახულებაზე არსებული სწორხაზოვანი ან მრუდხაზოვანი წირების შესაბამისი კონტურული წერტილების

სიებს. მიმდინარე სეგმენტები ინახება ოპერატიულ მეხსიერებაში. მათი წვდომისთვის გამოიყენება ჰეშ-ცხრილი. დამუშავებული სეგმენტები კი ინახება დისკურ ფაილში და მათი შესაბამისი ადგილი ჰეშ-ცხრილში თავისუფლდება. მთავარი განსხვავება მდგომარეობს შესაპირაპირებელი პიქსელების და მათში შემავალი სეგმენტების აღმოჩენაში. კერძოდ, გამოსახულების დათვალიერება ხორციელდება ზემოდან-მარცხნივ მიმართულებით, ხოლო მათგან გამომავალი სეგმენტების კი – მარჯვნივ-დან-ქვემოთ მიმართულებით. მაგალითისთვის განვიხილოთ ოპერატორი, რომელიც ადგენს, წარმოადგენს თუ არა მოცემული პიქსელი იზოლირებულ წერტილს, ახალი სეგმენტის საწყის წერტილს, ძველი სეგმენტის შიდა წერტილს, ძველი სეგმენტის ბოლო წერტილს, შეპირაპირების თუ კუთხის წერტილს. იმის მიხედვით, თუ რას წარმოადგენს მოცემული პიქსელი, ოპერატორის მოქმედებაც შესაბამისია. ასეთი ოპერატორის მოქმედების შედეგი მარკირებულ კონტურულ გამოსახულებაზე (სურ. 13.17) წარმოადგენილია 13.18 სურათზე.

სეგმენტი	სიგრძე	სეგმენტის კოორდინატების სია
1	3	(1,1)(2,2)(3,3)
2	3	(1,5)(2,4)(3,3)
3	3	(3,3)(4,3)(5,3)
4	3	(5,3)(5,4)(5,5)

სურ. 13.18

მოცემულ სურათზე წერტილი (5,3) განიხილება როგორც კუთხის წერტილი. იმ შემთხვევაში, თუ სეგმენტების დასამთავრებლად კუთხის წერტილები არ იქნებოდა გამოყენებული, მა-

შინ მე-3 სეგმენტის სიგრძე 5-ის ტოლი იქნებოდა და ის წარმოიდგინებოდა ასეთი სიის სახით:

$$((3,3),(4,3),(5,3),(5,4),(5,5)).$$

13.7. მოძრავი ობიექტების კონტურების გამოყოფა

ჩვენ უკვე განვიხილეთ ორგანზომილებიან გამოსახულებებზე მოძრავი ობიექტების აღმოჩენის და მათი ნიშან-თვისებების გამოყოფის საკითხები (XII თავი). ისევე, როგორც სტატიკური სცენების შემთხვევაში, აქაც დგას მოძრავი ობიექტების სეგმენტაციის, ანუ მათი კონტურების გამოყოფის ამოცანა. ამ პარაგრაფში ჩვენ გავეცნობით ორგანზომილებიან გამოსახულებებზე მოძრავი ობიექტების სეგმენტაციის საკითხს.

13.7.1. მოძრავი ობიექტების საზღვრების დადგენა

ორგანზომილებიან გამოსახულებაზე მოძრავი ობიექტების კონტურების აღმოსაჩენად გამოიყენება მონაცემები როგორც სივრცითი, ასევე დროითი კონტრასტის შესახებ. სტატიკური სცენების გამოსახულებების შემთხვევაში მხედველობაში მიიღებოდა ერთ გამოსახულებაზე რომელიმე ნიშან-თვისების (სიკაშკაშე, ტექსტურა და ა.შ.) სიცრცითი კონტრასტის ცვლილება. მაგრამ დინამიკური სცენების შემთხვევაში უკვე სხვა შემთხვევასთან გვაქვს საქმე. თუ გვაქვს, მაგალითად, სცენის ორი $I[x, y, t]$ და $I[x, y, t + \Delta t]$ გამოსახულება, მაშინ მათთვის შეიძლება გამოვითვალოთ და ერთობლივად განვიხილოდ სივრცითი და დროითი გრადიენტები. ასეთ სიდიდედ მოიაზრება **სივრცით-დროითი გრადიენტი (ს-დგ)** (spatio-temporal gradient magnitude, STG), რომელიც განისაზღვრება როგორც სივრცითი და

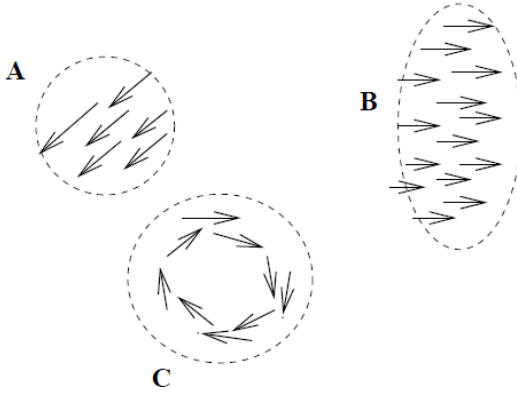
დროითი გრადიენტების მოდულების ნამრავლი და ის შემდეგნაირად გამოსახება

$$\mathbf{STG}[x, y, t] = \|[x, y, t]\|(|\mathbf{I}[x, y, t] - \mathbf{I}[x, y, t + \Delta t]|).$$

ამ სიდიდის მნიშვნელობები გამოსახულებათა წყვილის ყველა პიქსელისათვის შეიძლება წარმოვადგინოთ $\mathbf{STG}[\cdot]$ გამოსახულების სახით და შემდეგ მასზე გამოვიყენოთ ჩვენ მიერ კონტურების გამოყოფისთვის აქამდე განხილული ყველა მეთოდი. მაგრამ მიღებული კონტურები უკვე იქნება არა უძრავი, არამედ მოძრავი ობიექტების კონტურები.

13.7.2. მოძრაობის ტრაექტორიების დაგროვება

დავუშვათ, რომ ვიდეომიმდევრობის ორი კადრის საშუალებით შევძელით გადაადგილების ვექტორების გამოთვლა. ვთქვათ, რომ ისინი განსაზღვრულია გამოსახულების არა ყველა, არამედ მხოლოდ ინტერესის რამდენიმე წერტილისთვის ან არისთვის (XII თავი). გადაადგილების ვექტორები შეიძლება გამოყენებულ იქნეს გამოსახულებაზე რაიმე ნიშნით ერთგვაროვანი არეების გამოსაყოფად. ხოლო რაც შეეხება კლასტერიზაციას, ის უნდა მოვახდინოთ გამოსახულებაზე გადაადგილების ვექტორების ადგილმდებარეობის, მათი სიჩქარის და მოძრაობის მიმართულების მხედველობაში მიღებით. 13.19 სურათზე წარმოდგენილ შემთხვევაში **A** და **B** ობიექტები გადატანით მოძრაობას ასრულებს, ხოლო **C** ობიექტი კი ბრუნვითს. აღსანიშნავია, რომ გამოსახულებაზე გადატანითად მოძრავი ობიექტის აღმოჩენა უფრო ადვილია, ვიდრე მბრუნავის.



სურ. 13.19

ამასთან, იმის გამო, რომ გადატანითი მოძრაობისას ობიექტების ყველა წერტილს ერთნაირი სიჩქარე აქვს, მათი შესაბამისი კლასტერები გამოსახულებაზე ერთმანეთთან მჭიდროდ იქნება განლაგებული. რაც შეეხება მბრუნავ ობიექტებს, მათი აღმოჩენა გამოსახულების უფრო რთულ დამუშავებას მოითხოვს.

ლიტერატურა

1. Azad P., Gockel T., Dillmann R., Computer Vision - Principles and Practice, Elektor International Media BV, 2008.
2. Ballard D. H., Brown Ch. M., Computer Vision, Prentice-Hall, Englewood Cliffs, New Jersey, 1982.
3. Gonzales R.C., Woods R.E., Digital Image Processing, Prentice-Hall, 2002.
4. Haralick R.M., Shapiro L.G., Image Segmentation Techniques, Computer Vision, Graphics and Image Processing, Vol. 29, #1, pp. 100-132, 1985.
5. Jahne B., Digital Image Processing, 6th revised and extended edition, Springer-Verlag Berlin Heidelberg, 2005.
6. Juneja M., Sandhu P.S., Performance Evaluation of Edge Detection Techniques for Images in Spatial Domain, Int. Jour. of Computer Theory and Engineering, Vol. 1, #5, pp. 614-621, 2009.
7. Kastrinaki V., Zervakis M., Kalaitzakis K., A Survey of Video Processing Techniques for Traffic Applications, Image and Vision Computing, Vol. 21, Issue 4, pp. 359-381, 2003.
8. Maini R., Aggarwal H., Study and Comparison of Various Image Edge Detection Techniques, Int. Jour. of Image Processing (IJIP), Vol. 3, Issue 1, pp. 1-11, 2009.
9. Nixon M. S., Aguado A. S., Feature Extraction and Image Processing, Newnes, Oxford, 2002.
10. NeTra: A Toolbox for Navigating Large Image Databases, Multimedia Systems 7, pp. 184-199, Springer-Verlag, 1999.
11. Nomiya H., Uehara K., Content-based Image Classification via Visual Learning, Data Mining and Knowledge Discovery in Real Life Applications, Book Edited by: Ponce J., Karahoca A., pp. 438, February 2009, I-Tech, Vienna, Austria, 2009.

12. Singhai N., Shandilya S., A Survey on: Content Based Image Retrieval Systems, Int. Jour. of Computer Applications, Vol. 4, #2, pp. 22-26, 2010.
13. Szelisky R., Computer Vision: Algorithms and Applications, Springer-Verlag London Limited, 2011.
14. Tuceryan M., Jain A. K., Texture Analysis, The Handbook of Pattern recognition and Computer Vision (2nd Edition), by Chen C.H., Pau L.F. and Wang P. S. P. (eds.), pp. 207-248, World Scientific Publishing Co., 1998.
15. Yilmaz A., Javed O., Shah, Object Tracking: A Survey, ACM Computing Surveys, Vol. 38, #4, Article 13, 2006.
16. Young I.T., Gerbrands J. J., Van Vliet L.J., Fundamentals of Image Processing, Version 2.2, Delft University of Technology, Netherlands, 1998.
17. Zhang D. and Lu.G., Review of Shape Representation and Description techniques, Pattern recognition, Vol. 37, pp. 1-19, 2004.
18. Дуда Р., Харт П., Распознавание образов и анализ сцен, -М.: Мир, 1976.
19. Мэпп Д., Зрение. –М.: “Радио и связь”, 1987.
20. Методы компьютерной обработки изображений, Под ред. Сойфера В.А., 2-е изд., -М.:ФИЗМАТЛИТ, 2003.
21. Павлидис Т., Алгоритмы машинной графики и обработки изображений, -М.: Радио и связь, 1986.
22. Прэтт У.К., Цифровая обработка изображений, -М.: Мир, т.1, 2, 1982.
23. Розенфельд А., Распознавание и обработка изображений, -М.: Мир, 1972.
24. Хорн Б.К.Л., Зрение роботов, -М.: Мир, 1989.
25. Шапиро Л., Стокман Дж., Компьютерное зрение, Москва, изд-во БИНОМ, 2006.

სარჩევი

ავტორისაგან	3
I თავი. შესავალი	4
1.1. კომპიუტერული ხედვის საგნის არსი და მიზანი	4
1.2. ციფრული გამოსახულების სტრუქტურა	14
1.3. ციფრული გამოსახულებების გამოყენების მაგალითები	16
1.4. ოპერაციები გამოსახულებებზე	19
1.4.1. პიქსელების მოდიფიკაცია მცირე მიდამოებში	21
1.4.2. გამოსახულების ხარისხის გაუმჯობესება	22
1.4.3. რამდენიმე გამოსახულების კომბინაცია	23
1.5. გამოსახულებების სიმბოლური (არაგრაფიკული) აღწერის ფორმირება	23
II თავი. გამოსახულების ფორმირება და წარმოდგენა	25
2.1. მხედველობის ფსიქოფიზიკური თვისებები	25
2.2. გამოსახულების ფორმირების მოდელი	29
2.3. კომპიუტერში გამოსახულების შექვანის ტექნიკური საშუალებები	37
2.4. გამოსახულების მათემატიკური მოდელი	39
2.4.1. უწყვეტი გამოსახულების აღწერა	40
2.4.2. დისკრეტული გამოსახულების აღწერა	47
2.5. ციფრული გამოსახულების ფორმატები	52
III თავი. გამოსახულების დისკრეტიზაცია და კვანტირება	57
3.1. სივრცითი გაზომვები გამოსახულებაზე	57

3.2.	გამოსახულების დისკრეტიზაცია	59
------	-----------------------------	----

IV თავი. ბინარული გამოსახულებების

ანალიზი	67
----------------	----

4.1.	პიქსელი და მისი გარემო	67
4.2.	ტრაფარეტების გამოყენება გამოსახულებების დამუშავებისას	68
4.3.	ბმული კომპონენტები და მათი მარკირება	71
4.4.	ბინარული გამოსახულებების მორფოლოგია	74
4.4.1.	ბინარული მორფოლოგიის ოპერაციები	77
4.5.	ბინარულ გამოსახულებაზე არეების მომიჯნაობის აღწერა გრაფების საშუალებით	86
4.6.	ზღურბლის გამოყენება ნახევარტონური გამოსახულებების ბინარიზაციისათვის	90

V თავი. ფორმის წარმოდგენა და აღწერა

5.1.	არის ფორმის წარმოდგენა და აღწერა	97
5.2.	კონტურზე დამყარებული ფორმის წარმოდგენა და აღწერა	100
5.3.	არეზე დამყარებული ფორმის წარმოდგენა და აღწერა	109
5.4.	ფორმის მარტივი აღმწერები	115
5.5.	ობიექტების ამოცნობა ფორმაზე დაყრდნობით	125

VI თავი. გამოსახულებების გაუმჯობესება და რესტავრაცია

6.1.	შესავალი	132
6.2.	ტონების განაწილების ცვლილება გამოსახულებაზე	134
6.2.1.	კონტრასტის შეცვლა	134
6.2.2.	პრეპარირება	141

6.2.3	ჰისტოგრამის გარდაქმნა	145
6.3.	მცირე ზომის არეების მოშორება გამოსახულებაზე	155
6.3.1.	“მარილი და წიწაკა” ტიპის ხმაურისგან გაწმენდა	156
6.3.2.	მცირე კომპონენტების მოშორება	161
6.4.	გამოსახულების დაგლუვება	162
6.5.	მედიანური ფილტრი	166
 VII თავი. კიდის აღმოჩენა გამოსახულებაზე		170
7.1.	კიდის არსი გამოსახულებაზე	170
7.2.	კონტურები	175
7.3.	კიდის აღმოჩენის დიფერენციალური მეთოდები	181
7.4.	გრადიენტზე დამყარებული ოპერატორები	183
7.4.1.	რობერტსის ოპერატორი	184
7.4.2.	სობელის და პრევიტის ოპერატორები	187
7.5.	ლაპლასიანზე დამყარებული ოპერატორები (ლაპლასიანები)	191
7.5.1.	ლაპლასის ოპერატორი	192
7.6.	კიდების აღმოჩენა გაუსის ფილტრაციით და ლოგ - ფილტრაციით	195
7.6.1.	ქენის დეტექტორი	202
 VIII თავი. სცენის გამოსახულებათა ანალიზი		206
8.1.	შესავალი	206
8.2.	სცენის გამოსახულებათა ამოცნობა: ზოგადი მოდელი	208
8.2.1	კლასიფიკაციის სისტემის ცდომილებების შეფასება	215
8.2.2.	მცდარი ამოქმედებები და მცდარი გამოტოვებები	217

8.3.	კავშირი სისტემის სიზუსტესა და ამოკრების სისრულეს შორის	218
8.4.	სცენის გამოსახულებაზე ორგანზომილებიანი ობიექტების ამოცნობის ზოგადი სტრუქტურა	220
8.5.	ობიექტების წარმოდგენა ნიშან-თვისებათა ვექტორის საშუალებით	223
8.6.	კლასიფიკატორის რეალიზაციის ალგორითმები	226
	8.6.1. კლასიფიკაცია უახლოესი საშუალო მნიშვნელობის მიხედვით	227
	8.6.2. კლასიფიკაცია უახლოეს მეზობლამდე მანძილის მიხედვით	232
8.7.	ამოცნობის სტრუქტურული მეთოდები	234
8.8.	უზუსტობათა მატრიცა	239
8.9.	გადაწყვეტილებათა ხე	240
8.10.	გადაწყვეტილების მიღების ბაიესის წესი	245
8.11.	გამოსახულებების ნიშან-თვისებები	250
	8.11.1. სიკაშკაშის ნიშან-თვისებები	252
	8.11.2. ალბათური ნიშან-თვისებები	252
	8.11.3. ჰისტოგრამული ნიშან-თვისებები	253
	8.11.4. გეომეტრიული ნიშან-თვისებები	254
	8.11.5. ტოპოლოგიური ნიშან-თვისებები	257

IX თავი. ფერი და განათებულობა 263

9.1.	სინათლისა და ფერის ფიზიკური თვისებები	263
9.2.	ფერების RGB სისტემა	266
9.3.	ფერადი ჰისტოგრამები	272
9.4.	ფერადი გამოსახულებების სეგმენტაცია	274
9.5.	განათებულობა	277
	9.5.1. სინათლის ერთეულოვანი წყაროს გამოსხივება	277
	9.5.2. დიფუზიური არეკვლა	279

9.5.3.	სარკული არეკვლა	281
9.5.4.	ინტენსივობაზე მოქმედი ფაქტორები	283
9.6.	ფერის აღქმა ადამიანის მიერ	285
9.7.	მრავალსპექტრალური გამოსახულებები	286
X თავი. ტექსტურა		287
10.1.	ტექსტურის არსი	287
10.2.	ტექსტურის აღწერა ტექსტების საშუალებით	290
10.3.	ტექსტურის რაოდენობრივი მახასიათებლები	292
10.3.1.	კიდების სიმკვრივე და მიმართულება	293
10.3.2.	მომიჯნობის მატრიცები	296
XI თავი. გამოსახულებების მოძიება მათი შინაარსის მიხედვით		301
11.1.	გამოსახულებების მონაცემთა ბაზები	301
11.2.	მოთხოვნების აგება ნიმუშის მიხედვით	304
11.3.	გამოსახულებებს შორის მსგავსების განსაზღვრა	306
11.3.1.	ფერთა მსგავსების მახასიათებლები	307
11.3.2.	ტექსტურული მსგავსების მახასიათებლები	311
11.3.3.	ფორმის მსგავსების მახასიათებლები	314
11.3.4.	ფორმის ჰისტოგრამები	318
11.3.5.	ესკიზური შეპირისპირება	319
XII თავი. მოძრაობის აღმოჩენა დინამიკურ სცენებზე		322
12.1.	მოძრაობის არსი გამოსახულებაზე	323
12.2.	გამოსახულებების გამოკლება	326
12.3.	სცენაზე წერტილთა გადაადგილების ვექტორების გამოთვლა	329
12.4.	ოპტიკური ნაკადი	333

12.5.	ოპტიკური ნაკადის განტოლება	338
12.6.	მოძრავი ობიექტების თვალთვალი ვიდეომიმდევრობებზე	345
12.7.	ვიდეომიმდევრობების სეგმენტაცია	352
	XIII თავი. გამოსახულებების სეგმენტაცია	355
13.1.	სეგმენტაციის არსი და მიზნები	355
13.2.	კლასტერიზაციის მეთოდები	362
13.3.	არეების დაგროვების მეთოდი	377
13.4.	არეების წარმოდგენის ხერხები	387
	13.4.1. საწყის გამოსახულებაზე ზედღების ხერხი	387
	13.4.2. მარკირებული გამოსახულებები	388
	13.4.3. საზღვრების კოდირება	389
	13.4.4. კვადრი ხეები	392
	13.4.5. თვისებათა ცხრილები	394
13.5.	საზღვრების ანალიზის მეთოდი	394
	13.5.1. არეების საზღვრების მიდევნება	395
13.6.	თავსებადი მეზობელი კონტურული ფრაგმენტების დაჯგუფება მრუდეებად	404
13.7.	მოძრავი ობიექტების კონტურების გამოყოფა	408
	13.7.1. მოძრავი ობიექტების საზღვრების დადგენა	408
	13.7.2. მოძრაობის ტრაექტორიების დაგროვება	409
	ლიტერატურა	411