

СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ  
ВОЛОДИМИРА ДАЛЯ

Факультет інформаційних технологій та електроніки

Кафедра інформаційних технологій та програмування

**Пояснювальна записка**  
до магістерської дипломної роботи

магістр

(освітньо-кваліфікаційний рівень)

на тему: Система обробки візуальної інформації для мінімізації  
відходів прокатного виробництва

Виконав: студент 2 курсу, групи ІСТ-23дм  
126 «Інформаційні системи та технології

(шифр і назва спеціальності)

Черепанов І. А.

(прізвище та ініціали)

Керівник Захожай О.І.

(прізвище та ініціали)

Рецензент Меняйленко О.С.

(прізвище та ініціали)

Київ – 2024 року

СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ  
ВОЛОДИМИРА ДАЛЯ

Факультет інформаційних технологій та електроніки  
Кафедра інформаційних технологій та програмування  
Освітньо-кваліфікаційний рівень магістр  
Спеціальність 126 «Інформаційні системи та технології»  
(шифр і назва спеціальності)

ЗАТВЕРДЖУЮ  
Завідувач кафедри ІТП  
\_\_\_\_\_ д.т.н., доц. Захожай О.І.  
(підпис)  
« \_\_\_\_ » \_\_\_\_\_ 2024 р.

ЗАВДАННЯ

на магістерську дипломну роботу студенту

Черепанов Ігор Анатолійович

(прізвище, ім'я, по батькові)

1. Тема роботи: Система обробки візуальної інформації для мінімізації відходів прокатного виробництва,  
керівник роботи доцент, д.т.н. Захожай Олег Ігорович,  
(вчене звання, науковий ступінь, прізвище, ім'я, по батькові)  
затверджені наказом університету від « 06 » 12 2024 року №361/15.15-С
  2. Строк подання студентом роботи: 14 грудня 2024 р.
  3. Вихідні дані до роботи: Матеріали науково-дослідної практики, науково-методична література, дані інтернет-мережі
  4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
    - 4.1 Вступ
    - 4.2 Аналіз питання та постановка завдання
    - 4.3 Дослідження методів і засобів ідентифікації зображень засобами технічного зору.
    - 4.4 Методи і алгоритми інформаційної технології ідентифікації об'єктів засобами технічного зору.
    - 4.5 Технічна реалізація автоматизованої системи ідентифікації контурів об'єктів за допомогою засобів технічного зору
  - 4.4 Висновки
  - 4.5 Перелік використаних джерел
5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
-

6. Консультанти розділів проєкту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 08.11.2024р.

**КАЛЕНДАРНИЙ ПЛАН**

№ з/п	Назва етапів дипломної роботи	Строк виконання етапів роботи	Примітка
1.	Одержання завдання на виконання роботи	08.11.2024	виконано
2.	Укладання і погодження з керівником плану і етапів виконання роботи	11.11.2024	виконано
3.	Узагальнення даних літературних джерел	14.11.2024	виконано
4.	Аналіз шляхів виконання завдання. Вибір і погодження з керівником оптимального шляху виконання завдання	17.11.2024	виконано
5.	Аналіз технічних засобів та існуючих систем	20.11.2024	виконано
6.	Реалізація практичної частини завдання	08.12.2024	виконано
7.	Укладання, оформлення та погодження пояснювальної записки з керівником	13.12.2024	виконано
8.	Надання пояснювальної записки на кафедрі	14.12.2024	виконано
9.	Підготовка доповіді та презентації	16.12.2024	виконано

Студент \_\_\_\_\_ Черепанов І.А.  
(підпис) (прізвище та ініціали)

Керівник роботи \_\_\_\_\_ Захожай О.І.  
(підпис) (прізвище та ініціали)

## ЗМІСТ

ВСТУП .....	6
РОЗДІЛ 1 АНАЛІЗ ПИТАННЯ ТА ПОСТАНОВКА ЗАДАЧІ.....	10
1.1 Дослідження об'єкту автоматизації .....	10
1.2 Аналіз технології процесу різання .....	10
1.3 Аналіз геометричних моделей систем технічного зору.....	14
1.3.1 Перспективна проекція.....	14
1.3.2. Афінна проекція.....	17
РОЗДІЛ 2 ДОСЛІДЖЕННЯ МЕТОДІВ І ЗАСОБІВ ІДЕНТИФІКАЦІЇ ЗОБРАЖЕНЬ ЗАСОБАМИ ТЕХНІЧНОГО ЗОРУ .....	20
2.1 Дескриптори границі .....	20
2.1.1 Ланцюгові коди .....	20
2.1.2 Сигнатури .....	22
2.1.3 Апроксимація багатокутниками.....	22
2.2 Дескриптори області.....	24
2.3 Визначення країв.....	25
2.4 Шум .....	27
2.4.1 Адитивний стаціонарний гаусів шум .....	27
2.4.2 Згладжування.....	28
2.5 Висновки .....	29
РОЗДІЛ 3 МЕТОДИ І АЛГОРИТМИ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ ЗАСОБАМИ ТЕХНІЧНОГО ЗОРУ .....	31
3.1 Сегментація.....	32
3.1.1 Проведення контурів і визначення границі .....	32
3.1.2 Локальний аналіз .....	33
3.1.3 Глобальний аналіз за допомогою перетворення Хоуга .....	35
3.1.4 Глобальний аналіз за допомогою методів теорії графів.....	37
3.2 Визначення граничного рівня.....	38
3.2.1 Глобальні та локальні пороги.....	38
3.2.2 Вибір оптимального порога.....	39

3.2.3	Визначення граничного рівня на основі характеристик границі.....	41
3.2.4	Визначення граничного рівня, засноване на декількох змінних .....	42
3.3	Областно - орієнтована сегментація .....	44
3.3.1	Розширення області за рахунок об'єднання пікселів .....	44
3.3.2	Розбивка і об'єднання області.....	46
3.4	Застосування руху.....	47
3.4.1	Основний підхід.....	47
3.4.2	Акумулятивна різниця.....	49
3.4.3	Визначення еталонного образу.....	50
3.5	Реалізація методу ідентифікації контурів об'єктів з використанням руху .....	50
<b>РОЗДІЛ 4 ТЕХНІЧНА РЕАЛІЗАЦІЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ІДЕНТИФІКАЦІЇ КОНТУРІВ ОБ'ЄКТІВ ЗА ДОПОМОГОЮ ЗАСОБІВ ТЕХНІЧНОГО ЗОРУ .....</b>		
<b>54</b>		
4.1	Схема розміщення прийомного відеопристрою автоматизованої системи ідентифікації контурів об'єктів .....	54
4.2	Структура автоматизованої системи ідентифікації контурів об'єктів ...	56
4.3	Структура програмного забезпечення автоматизованої системи ідентифікації контурів об'єктів .....	57
4.4	Висновки .....	57
<b>ВИСНОВКИ.....</b>		<b>59</b>
<b>ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ .....</b>		<b>63</b>

## ВСТУП

У цей час сильно виріс інтерес до застосування промислових систем машинного зору для рішення різного роду завдань, що виникають у процесі виробництва. Сьогодні системи машинного зору служать ефективним інструментом для контролю виробничих процесів і підвищення рентабельності підприємств. Високопродуктивні обчислювальні машини останнього покоління дозволяють вирішувати багато завдань обробки потоків цифрової відеоінформації й ухвалення рішення в режимі реального часу.

Термін «машинний зір» має багато синонімів: комп'ютерний зір, розпізнавання зорових образів, аналіз зображень і т.д.

У цілому в завдання систем машинного зору входить одержання цифрового зображення, обробка зображення з метою виділення значимої інформації на зображенні й математичний аналіз отриманих даних для рішення поставлених завдань.

Технологія комп'ютерного зору як наукомісткої області знань бере свій початок з 50-х років ХХ століття. Саме в цей період комп'ютери поступово почали ставати загальнодоступним засобом обробки й аналізу інформації. Перші системи оцифровки візуальної інформації були досить примітивними, а зображення - малоформатними й низькоінформативними.

Першим успіхом в області машинного зору можна назвати розробку Френка Розенблатта - перцептрон (від англійського perception - сприйняття).

Перцептрон був уперше змодельований в 1958 році, причому його навчання вимагало біля напівгодини машинного часу на ЕОМ IBM-704. Апаратний варіант - Mark I Perceptron - був побудований в 1960 р. і призначався для розпізнавання зорових образів. Його рецепторне поле складалося з матриці фотоприймачів розміром 20x20, і він цілком успішно вирішував ряд нескладних завдань.

Основними елементами сучасних систем машинного зору можна

назвати камеру, за допомогою якої одержують зображення, плату уведення, оцифровуючу зображення, і плату керування рухом. Розглянувши технологію машинного зору, можна виділити кілька етапів роботи системи. Першим етапом є одержання зображення контрольованого об'єкта. Далі отримане зображення необхідно ввести в промисловий контролер, де відбувається комп'ютерна обробка, аналіз отриманих даних і прийняття рішень у відповідність із закладеною програмою керування. Заключним етапом є вивід керуючих впливів на виконавчі пристрої.

Прогрес цифрової техніки привів до доступності систем машинного зору, здатних якісно вирішувати актуальні завдання в різних галузях промисловості, системах безпеки й ін. Так у даний момент системи машинного зору широко застосовуються в металургії для безконтактного контролю металопрокату, контролю якості труб; у машинобудівному виробництві для збору інформації про процес виготовлення й зборки деталей; в автомобілебудуванні, де необхідний візуальний контроль на кожному етапі зборки машини; а також у медичній і фармацевтичній промисловості, у робототехніці, у наукових дослідженнях і ін.

Теорію комп'ютерного зору повноправно можна вважати самостійним розділом кібернетики. По даній тематиці видаються сотні книг і монографій, проводяться безліч конференцій і симпозіумів, випускається різне програмне й апаратно-програмне забезпечення для рішення будь-яких завдань.

Існує безліч пристроїв формування зображення: від очей тварин до відеокамер і радіотелескопів. Не всі ці пристрої можуть містити лінзи. Наприклад, у першій моделі камери-обскури (що буквально означає "темна кімната"), винайденої в XVI столітті, лінз не було, замість них використався отвір, що дозволяло променям світла фокусуватися на стінці або напівпрозорій пластинці й у такий спосіб демонструвати закони перспективи, які за сто років до цього відкрив Брунеллеччі (Brunelleschi). Уже в 1550 році отвору стали замінити усе більше й більш складними лінзами, а сучасна цифрова або фотокамера - це, по суті, камера-обскура, здатна реєструвати

кількість світла, що попадає на кожен маленьку ділянку її задньої панелі.

Поверхня камери, на якій формується зображення, звичайно являє собою прямокутник, а форма людської сітківки набагато ближче до сферичної поверхні, тому панорамні фотоапарати можуть обладнатися циліндричними чутливими поверхнями. Крім форми, датчики зображення мають і інші характеристики. Вони можуть фіксувати зображення, дискретне в просторі (як, наприклад, у людському оці з паличками й колбочками, 35-міліметрових камерах з емульсійними зернами й цифровими камерами із прямокутними елементами зображень, або пікселями) або безперервне (як, наприклад, у старих телевізійних трубках).

Сигнал, фіксуємий датчиком зображення в певній крапці своєї чутливої поверхні, може бути дискретним або безперервним і складатися з одного числа (чорно-біла камера), декількох значень (наприклад, інтенсивності червоної, зеленої й синьої квіток для кольорової камери або реакції трьох типів колбочок для людського ока), великого набору чисел або навіть безперервної функції довжини хвилі (спектрометри).

**Мета і задачі дослідження.** Підвищення ефективності роботи автоматизованих систем ідентифікації контурів об'єктів шляхом впровадження нових методів аналізу і обробки зображень. Для досягнення цієї мети були поставлені наступні основні задачі:

- дослідження систем технічного зору з метою визначення оптимальних шляхів зниження часової складності процесу ідентифікації контурів об'єктів;
- розробка нового ефективного методу ідентифікації контурів об'єктів;
- розробка ефективного методу видалення шуму з потоку кадрів зображення;
- розробка автоматизованої системи ідентифікації контурів об'єктів за допомогою засобів технічного зору.



**Наукова новизна отриманих результатів** полягає у наступному:

– отримав подальшого розвитку метод локальної сегментації зображень, за рахунок визначення загального значення атрибуту для пікселів одного сегменту, що дозволило знизити часову складність процесу ідентифікації контурів об'єктів;

– запропонований новий метод ідентифікації контурів об'єктів, який завдяки використанню інформації про еталонне зображення і відео потоку, дозволяє визначити момент перекривання одного об'єкту іншим;

– отримав подальшого розвитку метод видалення шуму з потоку відеокадрів, за рахунок введення порівняння сегментованого зображення сусідніх кадрів між собою, що дозволило підвищити якість ідентифікації контурів об'єктів;

– запропонована нова автоматизована система ідентифікації контурів об'єктів за допомогою засобів технічного зору, яка за рахунок використання нових методів сегментації і видалення шуму дозволила підвищити якість процесу ідентифікації.

**Структура і обсяг роботи.** Робота складається з 4 розділів, 62 сторінок, 13 рисунків.

# РОЗДІЛ 1

## АНАЛІЗ ПИТАННЯ ТА ПОСТАНОВКА ЗАДАЧІ

### 1.1 Дослідження об'єкту автоматизації

У ході виконання роботи був досліджений об'єкт автоматизації – ножиці гільйотин поперечного різання прокату в металургійному виробництві. Технічні характеристики об'єкту автоматизації:

Тип – безмаховічний;

Максимальне посилення різання – 1.22 МПа (125000 кг)

Кут нахилу верхнього ножа –  $4^{\circ}18'$ ;

Хід верхнього ножа – 320 мм;

Число резів в хвилину – 11-12;

Виліт станиці – 900 мм.

Матеріал ножів – сталь 5ХНТ, 5НМТ, 6ЧИ2С, загартована в маслі і опущена на оптимальну твердість 48-50 по Роквелу.

Привід від двигуна типу ДП-82А (140 кВт), 600 об/мин.

Ножиці призначені для різання листів з вуглецевого, конструкційного, низьколегованої і легованої сталі з тимчасовим опором розриву в холодному стані до 780 Н/мм, розмірів (мм):

Товщина – від 4 до 25;

Ширина – від 1 250 до 2 100;

Довжина – від 2 500 до 10 000.

### 1.2 Аналіз технології процесу різання

Процес різання полягає в розподілі листу прокату на мірні довжини відповідно до замовлення, одержаного від фірми-замовника на дану партію прокату. Лист, підготовлений до торцювання подають до рольгангу на ножиці, обрізають передню кромку, після чого здійснюється нанесення першою мітки мірної довжини (за допомогою лінійки та крейди), вимірювання довжини здійснюється щодо ножа. Лист подають по рольгангу

під ніж, коли мітка поєднується з лінією різання, рольганг зупиняють, затискають лист в упорах і опускають ніж. Лист клеймлять, наносять наступну мітку і процес повторюється.

Для поперечного різання гарячого прокату квадратного прямо вугільного і круглого перетину після плющення його на блюмінгах, слябінгах, і сортових станах заготовок, застосовують ножиці з паралельними ножами. Ці ножиці використовують також і для поперечного різання холодного прокату – в цьому випадку профіль ножа відповідає формі поперечного перетину розрізаючого прокату.

У процесі різання на цих ножицях площина, в якій рухається ніж (площина різання), є незмінною. Температура металу, що поступає на ножиці, складає звичайно 800-1000 градусів по Цельсію. Для вуглецевих сталей з тимчасовим опором в холодному стані  $\sigma^e = 400-500$  МПа при вказаних температурах  $\sigma^e = 100-50$  МПа; для легованих сталей з тимчасовим опором в холодному стані  $\sigma^e = 1000$  МПа при вказаних температурах  $\sigma^e = 130-80$  МПа.

Звідси витікає, що при виборі типу і характеристики ножиць слід враховувати їх призначення, щоб правильно визначити максимальне зусилля різання.

Основними параметрами ножиців є: максимальне зусилля різання  $P$ , хід ножів  $R$ , довжина ножа  $L$ , і число ходів (різань) в хвилину  $\Pi$  (продуктивність ножиць).

Припустимі розміри поперечного перетину прокату, що розрізає на ножицях даного типу, визначається максимальним зусиллям різання, на яке розраховані ножиці.

Хід ножів обирають з урахуванням можливості безперешкодного проходження прокату максимального перетину під лапою механічного (гідравлічного) притиску і перекриття ножів в кінці різання ( $A = 10-20$  мм).

Довжину ножів приймають рівною:

а) для невеликих ножиців із зусиллям різу 0,5-2,0 МН – 3-4 кратній максимальній ширині заготовки, виходячи з умови можливості різання на них заготовок дрібніших перетинів, укладених в декілька рядів (пачками);

б) для середніх і крупних ножиців із зусиллям різу до 16 МН, вживаних на блюмінгах, 2 – 2,5-кратній ширині блюма максимального перетину;

в) для ножиців слябінгів із зусиллям реза до 20 МН – 150-200 мм більше максимальної ширини слябів.

Поперечний перетин ножів звичайно приймають із співвідношення  $S/b = 2,5 \dots 3$ , де  $S$  – висота, а  $b$  – товщина ножа.

Форму ножів приймають у вигляді симетричного прямокутника, щоб можна було використовувати при різанні всі чотири кути. Кут загострення роблять рівним  $90^\circ$ . Ножі виготовляють із сталі марок 6ХНМ, 6ХВ2С і ін. з твердістю після термообробки до НВ 400.

По конструкції ножиці поперечного різання з паралельними ножами можна розділити на дві основні групи:

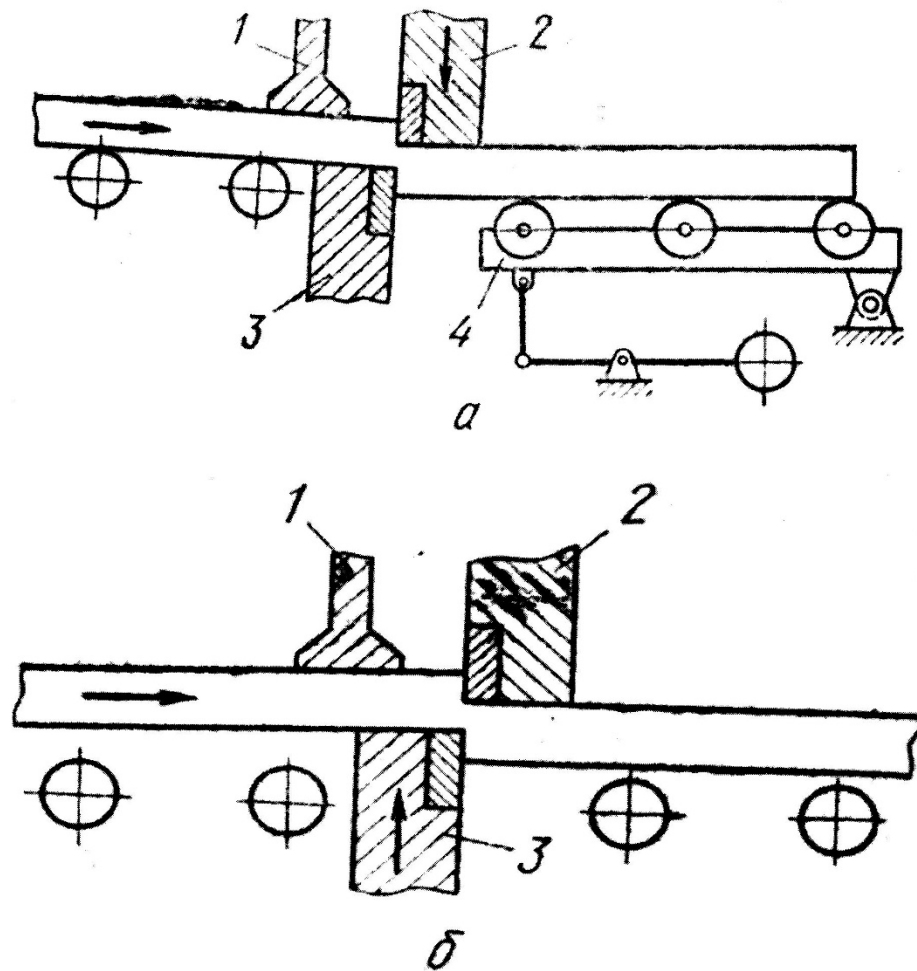
а) ножиці з верхнім рухомим ножем (верхнім різом);

б) ножиці з нижнім рухомим ножем (нижнім різом).

Ножиці з верхнім різом прості по конструкції. Принцип роботи цих ножиців полягає в наступному (рисунок 1.1, а): нижній ніж встановлений нерухомо в станині ножиців; верхній ніж укріплений в супорті (повзуні) і за допомогою кривошипного або гідравлічного приводу рухається вниз і розрізає метал. Слід зазначити, що цим ножицям властиві два істотні недоліки:

1) після різання металу на нижній грані сляба (блюма, заготовки) утворюється завусениць, який заважає при подальшому просуванні металу по рольгангу;

2) з схеми видно, що різання металу можна здійснити лише за наявності столу, що гойдається за ножицями, що ускладнює конструкцію всієї установки.



1 – притиск; 2 – верхній супорт; 3 – нижній супорт; 4 - задній рольганг, що підйомно гойдається, з вантажним урівноваженням

Рисунок 1.1 – Схема різання гуркоту на ножицях

Ножиці з нижнім різом не мають цих недоліків і тому одержали ширше вживаній. Принцип роботи ножиців полягає в наступному (рисунок 1.1, б): нижній ніж змонтований на супорті (повзуні), який може підійматися в гору за допомогою кривошипного або гідравлічного приводу; верхній ніж, встановлений у верхньому супорті (повзуні) і також може переміщатися по вертикалі.

Перед початком різання ножі розкриті, і метал проходить між ними по рольгангу; нижній ніж при цьому знаходиться нижче за рівень поверхні роликів рольганга і не заважає руху метала. Потім метал зупиняється в необхідному положенні ( за допомогою пересувного упору, як на рисунку

1.2), і супорт верхнього ножа опускається до зіткнення з металом; подальше посування верхнього супорта припиняється і починає рухатися супорт нижнього ножа; при цьому відбувається різання металу.

Ножиці з паралельними ножами конструктивно виконують двох типів: з електромеханічним і гідравлічним приводами.

### 1.3 Аналіз геометричних моделей систем технічного зору

#### 1.3.1 Перспективна проекція

Перспективна проекція ґрунтується на проходженні світлового потоку через крапковий отвір. При цьому відбувається дзеркальне формування зображення. На рисунку 1.2 наведений приклад перспективної проекції джерела світла - свічі.

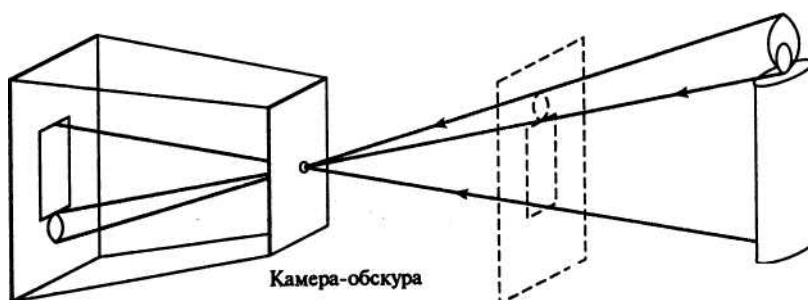


Рисунок 1.2 - Модель перспективної проекції

Це зображення утворять промені світла, що йдуть від сцени, що перебуває перед коробкою. Якщо отвір зменшити до крапки (що, звичайно ж, фізично неможливо), то через площину пластинки (або площина зображення), отвір і деяку крапку сцени буде проходити рівно один промінь.

У дійсності, отвір має кінцевий (хоча й невеликий) розмір, і в кожній крапці площини зображення збирається світло від цілого конуса променів утворюючий певний тілесний кут, так що ця ідеалізована й надзвичайно спрощена геометрична модель формування зображення не зовсім строго описує дійсність. Крім того, дійсні камери звичайно оснащені лінзами, що

також досить ускладнює ситуацію. Проте, модель крапкової перспективи (її ще називають центральною перспективою), уперше запропонована Брунеллеччі на початку XV століття, зручна з математичної точки зору. Незважаючи на свою простоту, вона часто є Прийнятним наближенням описом процесу формування зображення. У результаті перспективної проєкції виникає перевернене зображення, тому іноді замість нього зручно розглядати мнимо зображення, розташоване в площині, що лежить перед отвором на такій же відстані, як і реальна площина зображення (рисунок 1.2). Це мнимо зображення не перевернене, але у всім іншому воно точно відповідає сьогоденню. Залежно від ситуації, зручніше може бути одне або інше. На рисунку 1.3, показаний очевидний ефект перспективної проєкції: відносний розмір об'єктів залежить від відстані до них.

Наприклад, зображення  $B'$  і  $C'$  стовпчиків  $U$  і  $C$  здаються однаковими по висоті, а в дійсності стовпчики  $A$  і  $C$  у два рази менше, ніж  $B$ .

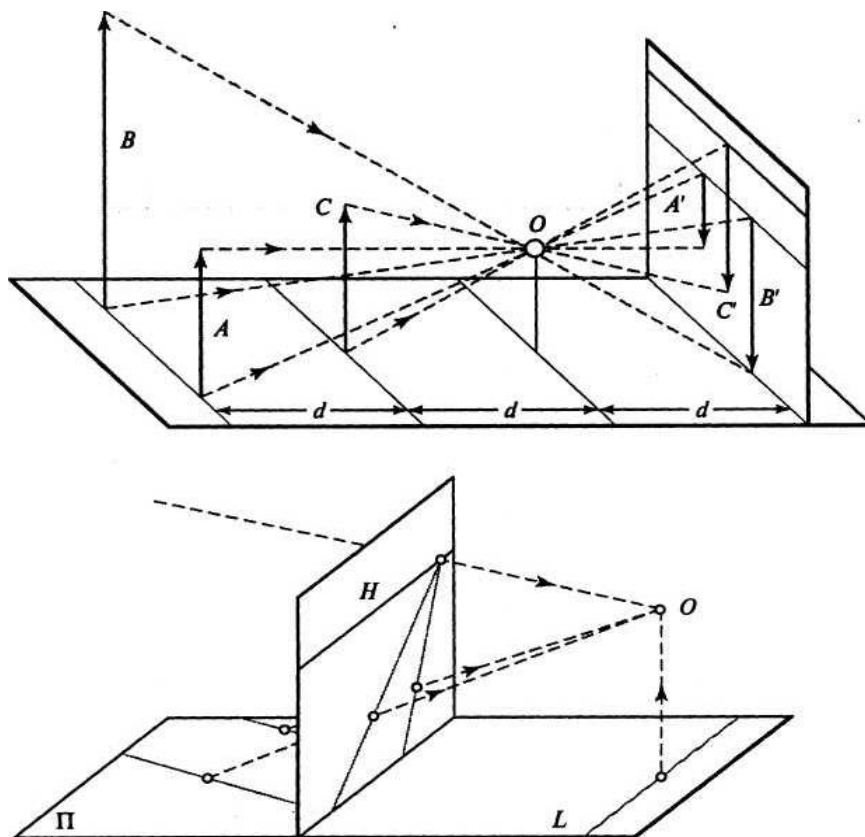


Рисунок 1.3 - Ефекти перспективи

На рисунку 1.3 (нижня ілюстрація), наведений ще один загальновідомий ефект: проєкції двох паралельних прямих, що лежать у деякій площині  $\Pi$ , сходяться на горизонтальній лінії  $H$ , утвореної перетинанням площини зображення із площиною, паралельною  $\Pi$  і минаючої через отвір. Помітимо, що пряма  $L$  на площині  $\Pi$ , паралельна площини зображення, взагалі не дає ніякого зображення.

Названі властивості легко довести чисто геометрично. У той же час, часто буває зручно (крім того, елегантно) оперувати такими поняттями, як система відліку, координати й рівняння. Розглянемо, наприклад, систему координат  $(O, i, j, k)$ , пов'язану з камерою-обскурою, початок координат  $Про$  яку збігається з отвором, а вектори  $i$  і  $j$  утворюють базис векторної площини, паралельної до площини зображення  $\Pi'$ , що перебуває на відстані  $f'$  від отвору в позитивному напрямку вектора  $k$  (рисунок 1.4). Пряма, що проходить через отвір і перпендикулярна до  $\Pi'$ , називається оптичною віссю, а крапка  $C'$ , у якій вона перетинається з  $\Pi'$ , називається центром зображення. Цю крапку можна використати в якості початку системи координат, пов'язаної із площиною зображення, і вона відіграє важливу роль у процедурах калібрування камер.

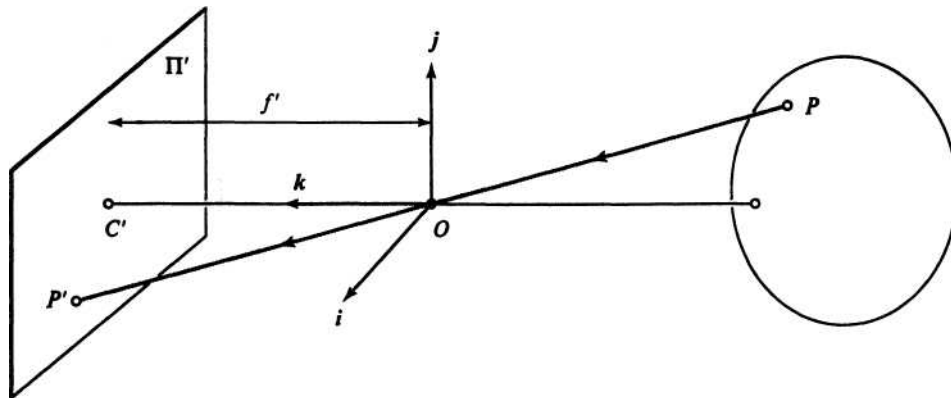


Рисунок 1.4 - Система координат пов'язана з камерою-обскурою

Позначимо через  $P$  крапку сцени з координатами  $(x, y, z)$ , а через  $P'$  зображення з координатами  $(x', y', z')$ . Тому що  $P'$  лежить у площині зображення, то  $z' = f'$ . Оскільки три крапки  $P, O$  і  $P'$  лежать на одній прямій,



те  $OP' = \lambda OP$  для деякого числа  $\lambda$ . Звідси треба, що

$$\begin{cases} x' = \lambda x \\ y' = \lambda y \\ f' = \lambda z \end{cases} \iff \lambda = \frac{x'}{x} = \frac{y'}{y} = \frac{f'}{z},$$

$$\begin{cases} x' = f' \frac{x}{z}, \\ y' = f' \frac{y}{z}. \end{cases}$$

(1.1)

### 1.3.2. Афінна проекція

Розглянемо дві афінні моделі: слабоперспективну й ортогональну.

Фронтально-паралельна площина  $\Pi^o$  визначається рівнянням  $z = z^o$  (рисунок 1.5). Для будь-якої точки  $P$  на  $\Pi^o$  рівняння перспективної проекції можна переписати як

$$\begin{cases} x' = -mx \\ y' = -my \end{cases} \quad m = -\frac{f'}{z_0}.$$

(1.2)

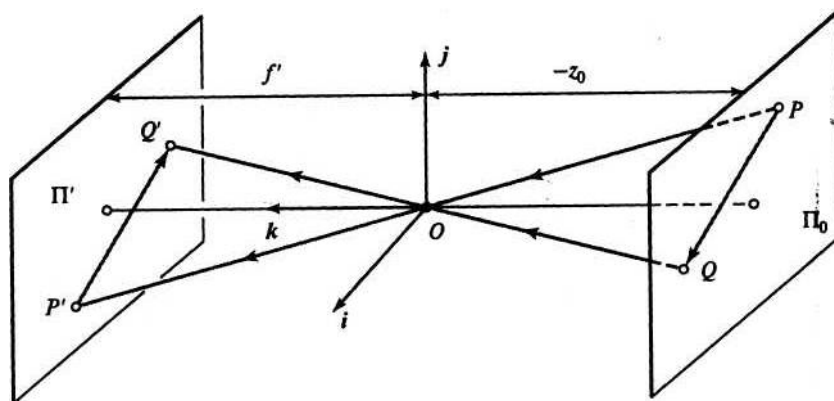


Рисунок 1.5 - Слабоперспективна проекція

У слабopersпективній проєкції всі елементи на площині  $\Pi^o$  проєктується з однаковим збільшенням. На відміну від інших геометричних моделей процесу формування зображення, при ортогональному проєктуванні елементи зображення не перевертаються. Отже, збільшення береться негативним, що небагато неприродно, але зате спрощує проєкційні рівняння.

По фізичних міркуваннях  $z^o$  повинне бути негативним (площина повинна перебувати перед отвором), так що коефіцієнт збільшення, пов'язаний із площиною  $\Pi^o$ , більше нуля. Назва параметра пояснюється наступними міркуваннями: розглянемо дві крапки  $P$  і  $Q$  на площині  $\Pi^o$  і їхнього зображення  $P'$  і  $Q'$  (мал. 1.4); очевидно, що вектори  $PQ$  і  $P'Q'$  паралельні, тому  $|P'Q'| = m|PQ|$ . Це вираження й представляє залежність розміру зображення від відстані до об'єкта, про яку згадувалося раніше. Коли глибина сцени мала в порівнянні із середньою відстанню до камери, збільшення можна вважати постійним. Ця проєкційна модель називається слабкою проєкцією, або масштабної ортографії. Якщо заздалегідь відомо, що камера постійно залишається на приблизно однаковій відстані від сцени, можна нормувати координати зображення таким чином, щоб коефіцієнт збільшення = -1. Це буде вже ортогональна проєкція, що визначається як

$$\begin{cases} x' = x, \\ y' = y. \end{cases} \quad (1.3)$$

Причому всі промені світла паралельні осі  $k$  і перпендикулярні площини зображення  $\Pi'$  (рисунок 1.6). Хоча слабopersпективна модель прийнятна при описі багатьох процесів формування зображення, застосування чистої ортогональної проєкції звичайно дає нереалістичні результати.

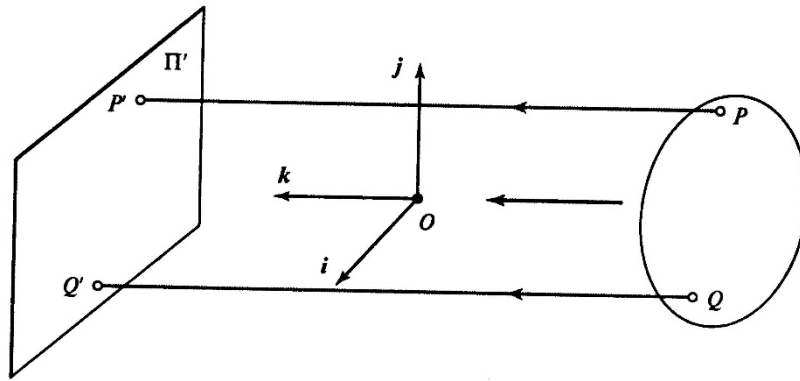


Рисунок 1.6 - Ортогональна проекція

#### 1.4 Висновки

1. На основі аналізу об'єкта автоматизації – гільйотинних ножиців для металургійного виробництва, можна зробити висновок, що необхідним є впровадження нових методів аналізу положення кромки листа при обрізанні.

2. Для якісного визначення моменту зупинки прокатного стану для обрізання листа необхідне використання методів, що дозволяють контролювати усю довжину кромки листа.

3. Використання засобів технічного зору дозволяє забезпечити необхідну якість контролю положення кромки листа при обрізанні.

4. Аналіз методів і способів формування зображення об'єктів показав, що кожен з них відрізняється максимальною точністю оцифрування і часовою складністю.

5. Для забезпечення максимальної точності і зниження часової складності аналізу зображення для ідентифікації контурів об'єктів доцільне використання ортогональної або слабо перспективної проекції.

## **РОЗДІЛ 2**

### **ДОСЛІДЖЕННЯ МЕТОДІВ І ЗАСОБІВ ІДЕНТИФІКАЦІЇ ЗОБРАЖЕНЬ ЗАСОБАМИ ТЕХНІЧНОГО ЗОРУ**

У системах технічного зору проблемою опису називається виділення властивостей (деталей) об'єкта з метою розпізнавання. В ідеальному випадку дескриптори не повинні залежати від розмірів, розташування й орієнтації об'єкта, але повинні містити достатню кількість інформації для надійної ідентифікації об'єктів. Опис є основним результатом при конструюванні систем технічного зору в тому розумінні, що дескриптори повинні впливати не тільки на складність алгоритмів розпізнавання, але також і на їхню роботу. Розглянемо три основні категорії дескрипторів: дескриптори границі, дескриптори області й дескриптори для опису тривимірних структур.

#### **2.1 Дескриптори границі**

##### **2.1.1 Ланцюгові коди**

Ланцюгові коди застосовуються для подання границі у вигляді послідовності відрізків прямих ліній певної довжини й напрямку. Звичайно в основі цього подання лежить 4- або 8-зв'язні прямокутні ґрати. Довжина кожного відрізка визначається дозволом ґрат, а напрямки задаються обраним кодом. Відзначимо що для подання всіх напрямків в 4-спрямованому ланцюговому коді досить 2 біт, а для 8-спрямованого ланцюгового коду потрібно 3 біт. Для породження ланцюгового коду заданої границі спочатку вибираються ґрати. Тоді, якщо площа осередку, розташованої усередині границі, більше певного числа (звичайно 50%), їй привласнюється значення 1; у протилежному випадку цьому осередку привласнюється значення 0. Остаточо ми кодуємо границю між двома областями, використовуючи напрямку. Результат кодування в напрямку по годинній стрілці з початком у місці, позначеному крапкою. Альтернативна процедура складається в розбивці границі на ділянки рівної довжини (кожна ділянка має те саме число

пікселів) і з'єднанні граничних крапок кожної ділянки прямою лінією, а потім присвоювання кожної лінії напрямку, найближчого до одному із припустимих напрямків ланцюгового коду. Важливо відзначити, що ланцюговий код даної границі залежить від початкової крапки. Однак можна нормувати код за допомогою простої процедури. Для створення ланцюгового коду початкова крапка на ґратах вибирається довільним образом. Розглядаючи ланцюговий код як замкнуту послідовність індексів напрямків, ми знову вибираємо початкову крапку таким чином, щоб результуюча послідовність індексів була цілим числом, що має мінімальну величину. Також можна нормувати повороти, якщо замість ланцюгового коду розглядати його першу різницю. Перша різниця обчислюється в результаті відлічування (у напрямку проти вартовий стрілки) числа напрямків, що розділяють два сусідніх елементи коду. Наприклад, перша різниця для ланцюгового коду з 4 напрямками 10103322 є 3133030. Якщо розглядати код як замкнуту послідовність, тоді перший елемент різниці можна обчислити, використовуючи перехід між останнім і першим компонентами ланцюга. У даному прикладі результатом є 33133030. Нормування можна здійснити шляхом розбивки всіх границь об'єкта на однакове число рівних сегментів і наступним припасуванням довжин сегментів коду з метою їхньої відповідності цій розбивці.

Викладені методи нормування є точними тільки в тому випадку, коли самі границі інваріантні до повороту й зміни масштабу. Цей випадок рідко зустрічається на практиці. Наприклад, той самий об'єкт, розбитий на елементи у двох різних напрямках, як правило, має різну форму границі, причому ступінь розходження пропорційний розв'язної здатності зображення. Цей ефект можна зменшити, якщо вибирати довжини елементів ланцюга більшими, ніж відстані між пікселями дискретного образу, або ж вибирати орієнтацію ґрат уздовж головних осей об'єкта, який підлягає кодуванню.

### **2.1.2 Сигнатури**

Сигнатурою називається одномірне функціональне подання границі. Відомо кілька способів створення сигнатур. Одним з найбільш простих є побудова відрізка із центра до границі як функції кута. Очевидно, що такі сигнатури залежать від периметра області й початкової крапки. Нормування периметра можна здійснити, пронормувавши криву максимальним значенням. Проблему вибору початкової крапки можна вирішити, визначивши спочатку ланцюговий код границі, а потім застосувавши метод, викладений у попередньому розділі. Звичайно, відстань, що залежить від кута, не є єдиним способом визначення сигнатури. Наприклад, можна провести через границю пряму лінію й визначити кут між дотичною до границі й цією лінією як функцію положення уздовж границі. Отримана сигнатура, хоча й відрізняється від кривій  $r()$ , несе інформацію про основні характеристики форми границі. Наприклад, горизонтальні ділянки кривої відповідали б прямим лініям уздовж границі, оскільки кут дотичної тут постійний. Один з варіантів цього методу як сигнатура використовує так називану функцію щільності нахилу. Ця функція являє собою гістограму значень кута дотичній. Оскільки гістограма є мірою концентрації величин, функція щільності нахилу строго відповідає ділянкам границі з постійними кутами дотичній (прямі або майже прямі ділянки й має глибокі провали для ділянок, що відповідають швидкій зміні кутів (виступи або інші види вигинів)).

### **2.1.3 Апроксимація багатокутниками**

Дискретну границю з довільною точністю можна апроксимувати багатокутниками. Для замкнутої кривої апроксимація є точною, коли число сегментів у багатокутнику дорівнює числу крапок границі, так що кожна пара сусідніх крапок визначає сегмент багатокутника. На практиці метою апроксимації багатокутниками є якісне визначення форми границі за допомогою мінімального числа багатокутних сегментів. Хоча звичайно ця проблема нетривіальна й досить швидко зводиться до ітеративного пошуку,

що вимагає більших тимчасових витрат, є ряд методів апроксимації багатокутниками, відносна простота яких і вимоги до обробки даних роблять їх придатними для додатків в області технічного зору.

У завданні апроксимації багатокутниками застосовуються методи об'єднання, засновані на помилці або інших критеріях. Один з підходів складається в з'єднанні крапок границі лінією по методу найменших квадратів. Лінія проводиться доти, поки помилка апроксимації не перевищить раніше заданий поріг. Коли поріг перевищується, параметри лінії заносяться на згадку, помилка покладається рівної нулю й процедура повторюється; нові крапки границі з'єднуються доти, поки помилка знову не перевищить поріг. Наприкінці процедури утворюються вершини багатокутника в результаті перетинання сусідніх ліній. Одні з основних труднощів, пов'язана із цим підходом, полягає в тому, що ці вершини звичайно не відповідають вигинам границі (таким, як кути), оскільки нова лінія починається тільки тоді, коли помилка перевищить поріг. Якщо, наприклад, довга пряма лінія перетинає кут, то числом (щозалежить від порога) крапок, побудованих після перетинання, можна зневажити раніше, ніж буде перевищене значення граничного рівня. Однак для усунення цих труднощів поряд з методами об'єднання можна використати методи розбивки.

Один з методів розбивки сегментів границі складається в послідовному розподілі сегмента на дві частини доти, поки задовольняється заданий критерій. Наприклад, можна зажадати, щоб максимальна довжина перпендикуляра, проведеного від сегмента границі до лінії, що з'єднує дві крайні крапки цього сегмента, не перевищувала раніше встановленого значення граничного рівня. Якщо це має місце, найбільш далека крапка стає вершиною, розділяючи, таким чином, вихідний сегмент на два підсегменту. Цей метод володіє тим перевагою, що він адаптований до найбільш підходящих крапок вигину. Для замкнутої границі найкращою початковою парою крапок звичайно є крапки, найбільш вилучені від границі.

## 2.2 Дескриптори області

Область, що представляє інтерес, можна описати формою її границі або ж шляхом завдання її характеристик. Важливо відзначити, що методи, розглянуті вище, застосовуються для опису областей.

Існуючі системи технічного зору ґрунтуються на досить простих дескрипторах області, що робить їх більше привабливими з обчислювальної точки зору. Як варто очікувати, застосування цих дескрипторів обмежено ситуаціями, у яких об'єкти, що представляють інтерес, розрізняються настільки, що для їхньої ідентифікації досить кілька основних дескрипторів.

Площа області визначається як число пікселів, що втримуються в межах її границі. Цей дескриптор корисний при зборі інформації про взаємне розташування й форму об'єктів, від яких камера розташовується приблизно на тому самому відстані. Типовим прикладом може служити розпізнавання системою технічного зору об'єктів, що рухаються по конвеєрі.

Більша й мала осі області корисні для визначення орієнтації об'єкта. Відношення довжин цих осей, називане ексцентриситетом області, також є важливим дескриптором для опису форми області.

Периметром області називається довжина її границі. Хоча іноді периметр застосовується як дескриптор, частіше він використовується для визначення міри компактності області, рівної квадрату периметра, діленому на площу. Відзначимо, що компактність є безрозмірною величиною (і тому інваріантно до зміни масштабу) і мінімальної для поверхні, що має форму диска.

Зв'язковий називається область, у якій будь-яка пара крапок може бути з'єднана кривій, що повністю лежить у цій області. Для безлічі зв'язних областей (деякі з них мають отвору) як дескриптор корисно використати число Ейлера, що визначається як різниця між числом зв'язних областей і числом отворів. Наприклад, числа Ейлера для букв *A* і *B* відповідно дорівнює 0 і -1.



## 2.3 Визначення країв

Різка зміна яскравості зображення цікаво з кількох причин. По-перше, такі різкі зміни найчастіше виникають на границях об'єктів - це може бути зображення світлого предмета на темному тлі або темному предметі на світлому тлі. По-друге, різкі зміни яскравості часто бувають наслідком зміни відбивної здатності на досить характерних структурах - як смуги на зебрі або плями на леопарді. Нарешті, до різких змін яскравості зображення також часто приводять різкі зміни орієнтації поверхні.

Крапки зображення, у яких яскравість змінюється особливо сильно, часто називають краями, або крайовими крапками. Потрібно зв'язати крайові крапки із границями предметів або інших значущих елементів. При цьому складно точно визначити зміни, які варто виділяти, - чи можна визначити границю предмета по блакитному небу, що просвічується крізь листи? Як правило, складно виділити край, що має певне семантичне значення, оскільки для цього необхідна інформація більше високого рівня. Проте, досвід створення зорових систем підказує, що вивчення країв необхідно й корисно, тому варто знати, де ці краї проходять.

Визначення країв - тема суперечок, що не припиняються, багато хто з яких виявляються неінформативними.

Границі об'єкта не слід ототожнювати з різкими змінами значень на зображенні. По-перше, у найгіршій ситуації об'єкти можуть не дуже відрізнятися по кольорах від тла. По-друге, на об'єкти часто накладається текстура або оцінки, які мають власні краї, і їх буває настільки багато, що часто складно відшукати серед них значимі ділянки границь об'єкта. Нарешті, краї, що не мають ніякого відношення до границь об'єктів, можуть виникати завдяки тіням і т.п. Назвемо кілька способів уникнути цих труднощів.

У деяких ситуаціях є можливість управляти висвітленням; якщо можна вибирати висвітлення, то за допомогою ретельного підбора можна досягти величезної різниці в кольорі й усунути тіні. Потім, установивши більші

параметри згладжування й високе граничне значення контрастності, можна переконатися, що краю, обумовлені текстурою, згладжені й більше не виділяються. Цей метод сумнівний, оскільки важко підібрати потрібні параметри згладжування й граничні значення, а текстура не тільки заважає, але й несе певну інформацію.

Існують інші способи боротьби з неприємними розходженнями між краями й границями об'єктів. По-перше, можна працювати над удосконаленням детекторів краю.

По-друге, можна повністю заперечувати користь від детекторів країв. Такий підхід розвився зі спостереження, що на деяких етапах визначення країв, особливо при немаксимальному придушенні, відкидається інформація, що надзвичайно складно відновити. Це відбувається внаслідок ухвалення важливого рішення - перевіряти результат за допомогою граничного значення. Замість пропонується зберігати цю інформацію в "неявному" (термінологія з області теорії ймовірностей) видах. Вивчивши ці доводи, автори зійшлися на думці не приймати дану точку зору, оскільки на сьогоднішній день немає практичних механізмів обробки неявної інформації.

Нарешті, визначення країв можна розглядати як питання, зв'язаний з усіма елементами структури системи - фаталістична позиція, що складається в тім, що майже в кожного процесу візуалізації є свої труднощі, а правильний підхід до цієї проблеми полягає в тому, щоб зрозуміти, як інтегрується візуальна інформація й побудувати систему, толерантну до такої інтеграції

Всі детектори країв погано поводяться в кутах; різні тільки деталі поводження.

Краї, на які реагують детектори країв, іноді називають східчастими краями, оскільки для них характерно стрибкоподібна зміна значень, що іноді моделюють у вигляді сходів. Вивчається безліч інших видів країв. Найпоширеніший приклад — це крышеподібний край, що складається зі зростаючого сегмента, з'єданого зі спадаючої, що досить схоже на наслідок взаємного відбиття. Іншим наслідком взаємного відбиття є комбінація

східчастого й дахоподібних країв.

На практиці це робиться не часто, чому є два пояснення.

По-перше, у теорії (і на практиці) немає зручного базису для прийнятих моделей. Які саме складені краї варто шукати? Проста відповідь - ті, для яких легко знайти оптимальний фільтр - самий незадовільний.

По-друге, семантика дахоподібних і більше складних складових країв ще невизначених, чим для східчастих країв. Крім того, невідомо, що робити із дахоподібним краєм після того, як він буде знайдений.

Краю складно визначити й важко знайти, але завдання з виходом схеми визначення країв вирішуються. Дахоподібні краї точно так само важко визначити й складно знайти; до того ж жодна завдання з виходом схеми детекторів дахоподібних країв не була вирішена. Реальна проблема полягає в тому, що не існує надійного механізму пророкування того, що варто шукати.

## **2.4 Шум**

Основна проблема завдання визначення країв - це шуми зображення. Детектори країв побудовані таким чином, що дають великий вихід при різких змінах, а одна із причин виникнення різких змін - це додаток до пікселів сторонніх значень (оскільки шуми в різних пікселях, як правило, не корелюють, тобто можуть дуже сильно відрізнятися друг від друга).

Терміном шум звичайно позначають такі дані, з яких неможливо (або невідомо як, або непотрібно) витягти інформацію; все інше називається сигналом. Неправильно вважати, що шум не містить ніякої інформації - з його допомогою, наприклад, можна оцінити температуру камери, обробивши зображення, отримані в темній кімнаті із закритою лінзою. Більше того, оскільки без моделі шуму сказати про шум щось істотне не можна, буде невірним уважати, що шуми не моделюється. Взагалі, шум - це все те, чим не бажають користуватися, і в цьому полягає його суть.

### **2.4.1 Адитивний стаціонарний гаусів шум**

У моделі адитивного стаціонарного гаусового шуму до кожному

пікселю додається значення, що вибирається на основі гаусового розподілу ймовірностей. Майже завжди середнє значення цього розподілу дорівнює нулю. Як параметр цієї моделі використовується середньоквадратичне відхилення.

При буквальному розгляді модель адитивного стаціонарного гаусового шуму виявляється поганою моделлю шуму зображення. По-перше, ця модель припускає існування позитивних (і, що більш тривожно, негативних!) значень пікселів довільної величини. Втім, при правильному виборі середньоквадратичного відхилення для звичайних камер, що знімають поза приміщенням або при денному висвітленні, це не представляє великої проблеми, оскільки такі значення пікселів навряд чи з'являться на практиці. При відновленні зображення із шумами проблемним пікселям привласнюється, відповідно, нульове або максимальне значення.

По-друге, шумові значення зовсім незалежні, тобто в цій моделі не враховується можливість існування груп пікселів, реакція яких зкорельована, можливо, через особливості електроніки камери або через теплові плями в інтегральній схемі камери. Цю проблему вирішити небагато складніше, оскільки моделі шуму, що описують дане явище, досить складні з аналітичної точки зору. Нарешті, ця модель не описує "мертві" пікселі (пікселі, які при будь-яких обставинах не дають ніякої реакції на падаюче випромінювання або постійно перенасичені). "Мертві" пікселі зустрічаються тоді, коли середньоквадратичне відхилення досить велике й існує деяке граничне значення відображуваних пікселів, при цьому середньоквадратичне відхилення може бути настільки більшим, що іншу частину зображення описати неможливо. Вирішальна перевага моделі адитивного гаусового шуму - це простота оцінки реакції фільтрів на змодельовані в такий спосіб випадкові процеси. Це, у свою чергу, допомагає оцінити, наскільки ефективно фільтр реагує на сигнал і не сприймає шум.

#### **2.4.2 Згладжування**

У загальному випадку будь-яка зміна значення відбивається на цілому

блоці пікселів. Наприклад, контури предмета можуть перетворитися в довгий ланцюжок крапок, у яких похідна зображення буде великою. У багатьох моделях шуму більші похідні зображення, обумовлені шумом, є, по суті, локальними подіями. Це означає, що згладжування продиференційованого зображення - це деяке об'єднання шуканих змін і придушення ефекту шумів. Альтернативна інтерпретація: шукані зміни не будуть придушуватися при згладжуванні, при якому придушуються шумові ефекти.

Існує ще одне пояснення того, чим корисне згладжування. Припустимо, що зображення із шумами було згладжено, а потім продиференційовано. По-перше, ядро згладжування знижує дисперсію шуму, оскільки звичайно намагаються застосовувати позитивні ядра згладжування.

По-друге, прагнення пікселів бути схожими на своїх сусідів зростає: якщо взяти стаціонарний адитивний гаусів шуми й згладити його, то значення пікселів в отриманому сигналі більше не будуть незалежними. По суті в цьому й складається згладжування — згадаєте, що згладжування було уведено як спосіб пророкування значення пікселя за значеннями сусідніх з ним пік-сіней. У той же час, якщо пікселі прагнуть бути схожими на сусідні, те похідні повинні бути менше (оскільки вони показують те, наскільки пікселі відрізняються від сусідніх з ними).

## **2.5 Висновки**

1. Дескриптори не повинні залежати від розмірів, розташування й орієнтації об'єкта, але повинні містити достатню кількість інформації для надійної ідентифікації об'єктів.

2. Дослідження методів визначення дескрипторів границі показало, що нормування є точним тільки в тому випадку, коли самі границі інваріантні до повороту й зміни масштабу.

3. Точність нормування може бути збільшена, якщо вибирати довжини елементів ланцюга більшими, ніж відстані між пікселями дискретного образу, або ж вибирати орієнтацію ґрат уздовж головних осей кодуємого

об'єкта.

4. Дискретну границю з довільною точністю можна апроксимувати багатокутниками. Для замкнутої кривої апроксимація є точною, коли число сегментів у багатокутнику дорівнює числу крапок границі, так що кожна пара сусідніх крапок визначає сегмент багатокутника.

5. Існуючі системи технічного зору ґрунтуються на досить простих дескрипторах області, що робить їх більше привабливими з обчислювальної точки зору. Як варто очікувати, застосування цих дескрипторів обмежено ситуаціями, у яких об'єкти, що представляють інтерес, розрізняються настільки, що для їхньої ідентифікації досить кілька основних дескрипторів.

6. Як правило, складно виділити край, що має певне семантичне значення, оскільки для цього необхідна інформація більше високого рівня. Проте, досвід створення зорових систем підказує, що вивчення країв необхідно й корисно, тому варто знати, де ці краї проходять.

7. Основна проблема завдання визначення країв - це шуми зображення. Детектори країв побудовані таким чином, що дають великий вихід при різких змінах, а одна із причин виникнення різких змін - це додаток до пікселів сторонніх значень.

8. Ядро згладжування знижує дисперсію шуму, оскільки звичайно намагаються застосовувати позитивні ядра згладжування.

9. На деяких етапах визначення країв, особливо при немаксимальному придушенні, відкидається інформація, що надзвичайно складно відновити. Це відбувається внаслідок ухвалення важливого рішення - перевірити результат за допомогою граничного значення.

## РОЗДІЛ 3

### МЕТОДИ І АЛГОРИТМИ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ ЗАСОБАМИ ТЕХНІЧНОГО ЗОРУ

З метою класифікації методів і підходів, використовуваних у системах технічного зору, зір розбитий на три основних підкласи: зір низького, середнього й високого рівнів. Системи технічного зору низького рівня призначені для обробки інформації з датчиків параметрів і стану об'єкту інформатизації.

Ці системи можна віднести до класу «інтелектуальних» машин, якщо вони мають наступні ознаки (ознаками інтелектуального поведіння):

- 1) можливістю виділення істотної інформації з безлічі незалежних ознак;
- 2) здатністю до навчання на прикладах і узагальненню цих знань із метою їхнього застосування в нових ситуаціях;
- 3) можливістю відновлення подій за неповною інформацією;
- 4) здатністю визначати мети й формулювати плани для досягнення цих цілей.

Створення систем технічного зору з такими властивостями для обмежених видів робочого простору в принципі можливо, але характеристики таких систем далекі від можливостей людського зору. В основі технічного зору лежить аналітична формалізація, спрямована на рішення конкретних завдань. Машини із сенсорними характеристиками, близькими до можливостей людини, очевидно, з'являться ще не швидко. Однак слід зазначити, що копіювання природи не є єдиним рішенням цієї проблеми.

Системи технічного зору середнього рівня пов'язані із завданнями сегментації, описи й розпізнавання окремих об'єктів. Ці завдання охоплюють безліч підходів, заснованих на аналітичних поданнях. Системи технічного зору високого рівня вирішують проблеми, розглянуті вище. Для більше

ясного розуміння проблем технічного зору високого рівня і його зв'язку з технічним зором низького й середнього рівнів уведемо ряд обмежень і спростимо розв'язуване завдання.

### **3.1 Сегментація**

Сегментацією називається процес підрозділу сцени на складові частини або об'єкти. Сегментація є одним з основних елементів роботи автоматизованої системи технічного зору, тому що саме на цій стадії обробки об'єкти виділяються зі сцени для подальшого розпізнавання й аналізу. Алгоритми сегментації, як правило, ґрунтуються на двох фундаментальних принципах: розривності й подібності. У першому випадку основний підхід ґрунтується на визначенні контурів, а в другому - на визначенні граничного рівня й розширенні області. Ці поняття застосовні як до статичних, так і до динамічного (залежним від часу) сценам. В останньому випадку рух може служити потужним засобом для поліпшення роботи алгоритмів сегментації.

На рисунку 3.1 проілюстрований процес сегментації зображення контуру в деякій площині  $xu$ . В процесі сегментації область  $xu$  розбивається на малі ділянки після реалізуються алгоритми визначення контуру і порогового рівня. В цьому випадку здійснюється аналіз подібності пікселів у межах дискретної ділянки і у випадку відповідності пороговому рівню здійснюється об'єднання точок у контур

#### **3.1.1 Проведення контурів і визначення границі**

Методи обчислення градієнта й граничний поділ - визначають розриви в інтенсивності подання образу об'єкта. В ідеальному випадку ці методи визначають пікселі, що лежать на границі між об'єктом і фоном. На практиці даний ряд пікселів рідко повністю характеризує границю через шум, розривів на границі внаслідок нерівномірної освітленості й інших ефектів, що приводять до розмитію зображення. Таким чином, алгоритми виявлення контурів супроводжуються процедурами побудови границь об'єктів з відповідних послідовностей пікселів. Нижче розглянуто кілька методик,



використовуваних для досягнення цієї мети.

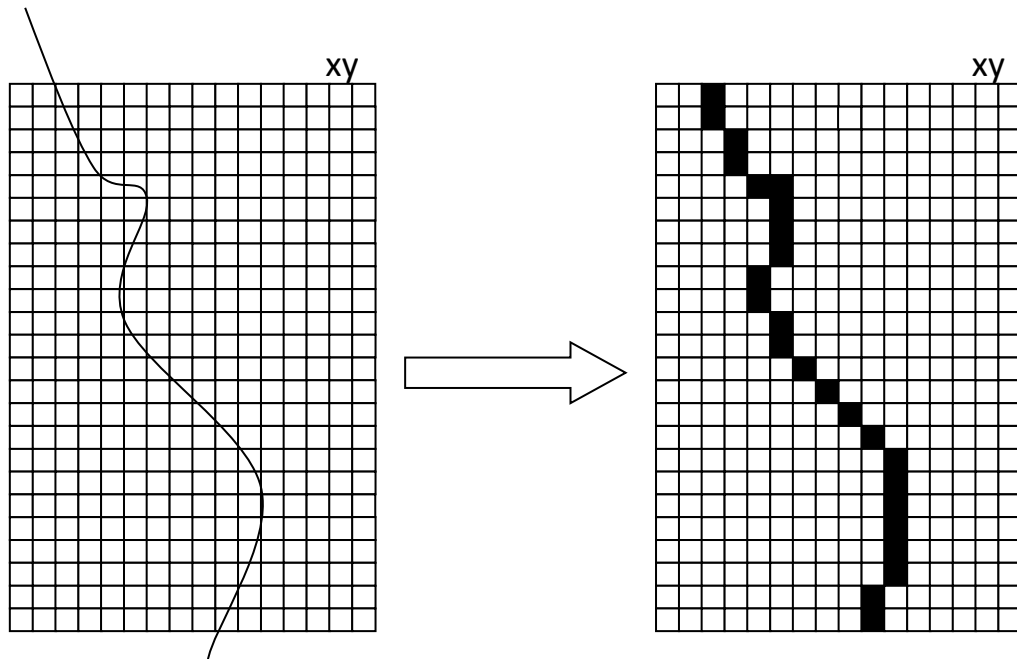


Рисунок 3.1 – Сегментація і визначення контуру об'єкту

### 3.1.2 Локальний аналіз

Одним з найбільш простих підходів з'єднання точок контуру є аналіз характеристик пікселів у невеликій області (наприклад, в області розміром 3 X 3 або 5 X 5) кожної крапки  $(x, y)$  образа, що вже піддався процедурі виявлення контуру. Всі крапки, що є подібними (визначення критерію подоби дано нижче), з'єднуються, образуєючи границю з пікселів, що володіють деякими загальними властивостями.

При такому аналізі для встановлення подоби пікселів контуру необхідно визначити:

- 1) величину градієнта, необхідного для побудови контурного пікселя,
- 2) напрямок градієнта.

Перша характеристика позначається величиною  $G[f(x, y)]$ .

$$G[f(x, y)] = [G_x^2 + G_y^2]^{1/2} = \left[ \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 \right]^{1/2} \quad (3.1)$$

$$G[f(x, y)] \cong |G_x| + |G_y|$$

Таким чином, піксель контуру з координатами  $(x', y')$  подібний по величині в певній раніше околиці  $(x, y)$  пікселю з координатами  $(x, y)$ , якщо справедливо нерівність

$$|G[f(x, y)] - G[f(x', y')]| \leq T \quad (3.2)$$

де  $T$  — граничне значення.

Напрямок градієнта встановлюється по куту вектору градієнта в рівнянні

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} \quad (3.3)$$

$$\theta = \text{arctg} \begin{bmatrix} G_x \\ G_y \end{bmatrix} \quad (3.4)$$

де  $\theta$ -кут (щодо осі  $x$ ), уздовж якого швидкість зміни має найбільше значення.

Тоді можна сказати, що кут пікселя контуру з координатами  $(x', y')$  у деякій околиці  $(x, y)$  подібний до кута пікселя з координатами  $(x, y)$  при виконанні наступної нерівності:

$$|\theta - \theta'| < A \quad (3.5)$$

де  $A$  — граничне значення кута.

Необхідно відзначити, що напрямок контуру в точці  $(x, y)$ , у дійсності, перпендикулярний напрямку вектору градієнта в цій точці. Однак для порівняння напрямків нерівність дає еквівалентні результати.

Ґрунтуючись на цих припущеннях, можна з'єднати точку в деякій околиці  $(x, y)$  з пікселем, що має координати  $(x, y)$ , якщо задовольняються критерії по величині й напрямку. Рухаючись від пікселя до пікселю й

представляючи кожну приєднувати точку, що, як центр околиці, процес повторюється для кожної крапки образу. Для встановлення відповідності між рівнями інтенсивності висвітлення й послідовностями пікселів контуру застосовується стандартна бібліотечна процедура.

Ціль складається у визначенні розмірів прямокутників, за допомогою яких можна побудувати якісне зображення. Побудова таких прямокутників здійснюється в результаті визначення строго горизонтальних і вертикальних контурів. Подальший процес складається в з'єднанні сегментів контуру, розділених невеликими проміжками, і в об'єднанні окремих коротких сегментів.

### 3.1.3 Глобальний аналіз за допомогою перетворення Хоуга

Цей метод ґрунтується на з'єднанні граничних крапок шляхом визначення їхнього розташування на кривій спеціального виду. Спочатку припускаючи, що на площині  $xu$  образу дано  $n$  крапок, потрібно знайти підпоследовності крапок, що лежать на прямих лініях. Одне з можливих рішень складається в побудові всіх ліній, що проходять через кожну пару крапок, а потім у знаходженні всіх підпоследовностей крапок, близьких до певних ліній. Завдання, пов'язана із цією процедурою, полягає в знаходженні  $n(n-1)/2 \sim n^2$  ліній і потім у здійсненні  $n[n(n-1)]/2 \sim n^3$  порівнянь кожної крапки з усіма лініями. Цей процес трудомісткий з обчислювальної точки зору за винятком найпростіших додатків.

Дане завдання можна вирішити по-іншому, застосовуючи підхід, запропонований Хоугом і називаний перетворенням Хоуга. Розглянемо крапку  $(x_i, y_i)$  і загальне рівняння прямої лінії  $v := ax_i + b_i$ . Є нескінченне число ліній, що проходять через крапку  $(x_i, y_i)$ , але всі вони задовольняють рівнянню  $v := ax_i + b_i$  при різних значеннях  $a$  й  $b$ . Однак записавши це рівняння у вигляді  $b = -x_i a + y_i$  і розглянувши площину  $ab$  (простір параметрів), одержимо рівняння однієї лінії для фіксованої пари чисел  $(x_i, y_i)$ . Більше того, друга крапка  $(x_j, y_j)$  також має в просторі параметрів пов'язану з нею лінію, що перетинає іншу лінію, пов'язану із крапкою  $(x_i, y_i)$  у крапці  $(a', b')$ , де

значення  $a'$  і  $b'$ -параметри лінії, на якій розташовані крапки  $(x_i, y_i)$  і  $(x_j, y_j)$  у площині  $xu$ . Фактично всі крапки, розташовані на цій лінії, у просторі параметрів будуть мати лінії перетинання в крапці  $(a', b')$ .

Обчислювальна ефективність перетворення Хоуга полягає в поділі простору параметрів на так називані елементи, що збирають. Елемент, що збирає,  $A(i, j)$  відповідає площі, пов'язаної з координатами простору параметрів  $(a_i, b_j)$ . Спочатку ці елементи вважаються рівними нулю. Тоді для кожної крапки  $(x_k, y_k)$  у площині образа представляємо параметр  $a$  рівним кожному із припустимих значень на осі  $a$  й обчислюємо відповідне  $b$ , використовуючи рівняння  $b = -x_k + y_k$ . Отримане значення  $b$  потім округляється до найближчого припустимого значення на осі  $b$ . Якщо вибір  $a_p$  приводить до обчислення  $b_q$ , думаємо  $A(p, q) = A(p, q) + 1$ . Після завершення цієї процедури значення  $M$  у елементі  $A(i, j)$  відповідає  $M$  крапкам у площині  $xu$ , що лежить на лінії  $y = ax + b$ . Точність розташування цих крапок на одній прямій залежить від числа розбивок площини  $ab$ . Відзначимо, що, якщо розбивається вісь  $a$  на  $K$  частин, тоді для кожної крапки  $(x_k, y_k)$  виходить  $K$  значень  $b$ , що відповідають  $K$  можливих значень  $a$ . Оскільки є  $n$  крапок образа, процес складається із  $nK$  обчислювальних операцій. Тому наведена вище процедура лінійна відносно  $n$  і має менше число обчислювальних операцій, чим процедура, описана вище, якщо  $K \leq n$ .

Проблема, пов'язана з поданням прямої лінії рівнянням  $v = ax + b$ , полягає в тому, що обидва параметри  $a$  й  $b$  прагнуть до нескінченності, якщо лінія приймає вертикальне положення. Для усунення цього використається нормальне подання прямої лінії у вигляді

$$x \cos \theta + y \sin \theta = \beta. \quad (3.6)$$

Це подання для побудови таблиці елементів, що збирають, використається так само, як метод, викладений вище, але замість прямих ліній ми маємо синусоїдальні криві в площині  $\rho$ . Як і колись,  $M$  точок, що

лежать на прямій  $x\cos\theta_i + y\sin\theta_i = \rho_i$ , відповідають  $M$  синусоїдальним кривим, які перетинаються в точці  $(\theta_i, \rho_i)$  простору параметрів. Якщо використається метод зростання  $\theta$  і знаходження для нього відповідного  $\rho$ , процедура дає  $M$  крапок в елемент, що збирає,  $A(i, j)$ , пов'язаний із точкою  $(\theta_i, \rho_i)$ .

### 3.1.4 Глобальний аналіз за допомогою методів теорії графів

Викладені вище методи засновані на завданні послідовності крапок контуру, отриманих у результаті градієнтного перетворення. Цей метод не часто застосовується для попередньої обробки даних у ситуаціях, що характеризуються високим рівнем шуму, внаслідок того, що градієнт є похідний і підсилює коливання інтенсивності. Існує глобальний підхід, заснований на поданні сегментів контуру у вигляді графа й пошуку на графі шляху найменшої вартості, що відповідає значимим контурам. Цей підхід представляє наближений метод, ефективний при наявності шуму. Як і варто очікувати, ця процедура значно складніше й вимагає більше часу обробки, чим методи, викладені вище.

Граф  $G = (N, A)$  являє собою кінцева, непуста безліч вершин  $N$  разом з безліччю  $A$  неупорядкованими парами різних елементів з  $N$ . Кожна пара з  $A$  називається дугою.

Граф, у якому дуги є спрямованими, називається спрямованим графом. Якщо дуга виходить із вершини  $n_i$ , до вершини  $n_j$ , тоді  $n_j$  називається спадкоємцем вершини  $n_i$ . У цьому випадку вершина  $n_i$  називається попередником вершини  $n_j$ . Процес ідентифікації спадкоємців кожної вершини називається розширенням цієї вершини. У кожному графі визначаються рівні таким чином, щоб нульовий рівень складався з єдиної вершини, названої початковою, а останній з вершин, названих цільовими. Кожній дузі  $(n_i, n_j)$  приписується вартість  $c(n_i, n_j)$ . Послідовність вершин  $n_1, n_2, \dots, n_k$ , де кожна вершина  $n_i$  є спадкоємцем вершини  $n_{i-1}$ , називається шляхом від  $n_1$  до  $n_k$ , а вартість шляху визначається формулою

$$c = \sum_{i=2}^k c(ni - 1, nj) \quad (3.7)$$

Елемент контуру визначається як границя між двома пікселями  $p$  и  $q$ . У даному контексті під контуром розуміється послідовність елементів контуру.

### 3.2 Визначення граничного рівня

Поняття граничного рівня (порога) – тест виду

$$T = T [x, v, p(x, y), f(x, y)], \quad (3.8)$$

де  $f(x, y)$  -інтенсивність у крапці  $(x, y)$ ,

$p(x, y)$ -деяка локальна властивість, обумовлена в околиці цієї крапки.

Граничне зображення дається наступним вираженням:

$$g(x, y) = \begin{cases} 1, & \text{если } f(x, y) > T \\ 0, & \text{если } f(x, y) \leq T \end{cases} \quad (3.9)$$

так що пікселі в  $g(x, y)$ , що мають значення 1, відповідають об'єктам, а пікселі, що мають значення 0, відповідають тлу. У рівнянні передбачається, що інтенсивність об'єктів більше інтенсивності тла. Протилежна умова виходить шляхом зміни знаків у нерівностях.

#### 3.2.1 Глобальні та локальні пороги

Якщо значення  $T$  у рівнянні залежить тільки від  $f(x, y)$ , те, поріг називається глобальним. Якщо значення  $T$  залежить як від  $f(x, y)$ , так і від  $p(x, y)$ , поріг називається локальним. Якщо, крім того,  $T$  залежить від просторових координат  $x$  а  $v$ , у цьому випадку він називається динамічним порогом.

Глобальні пороги застосовуються в ситуаціях, коли є явне розходження між об'єктами й фоном і де освітленість досить однорідна. Методи зворотної

й структурованої освітленості, звичайно дають зображення, які можуть бути сегментовані шляхом застосування глобальних порогів. Але, як правило, довільне висвітлення робочого простору приводить до зображень, які, якщо виходити з визначення граничного рівня, вимагають локального аналізу для компенсації таких ефектів, як неоднорідність висвітлення, тіні й відбиття.

Нижче ми розглянемо ряд методів для вибору порогів, використовуваних при сегментації. Хоча деякі з них можуть застосовуватися для вибору глобального порога, вони звичайно використовуються в ситуаціях, що вимагають аналізу локального порога.

### 3.2.2 Вибір оптимального порога

Часто розглядають гистограму, що складається із суми значень функції щільності ймовірності. У випадку бімодальної гистограми апроксимуюча її функція дається рівнянням

$$p(z) = P_1 p_1(z) + P_2 p_2(z), \quad (3.10)$$

де інтенсивність  $z$ -випадкова змінна величина,

$p_1(z)$  і  $p_2(z)$ -функції щільності ймовірності,

$P_1$  і  $P_2$  – апіорні ймовірності.

У цьому випадку апіорні ймовірності означають поява двох видів рівнів інтенсивності на образі. Повна гистограма може бути апроксимована сумою двох функцій щільності ймовірності. Якщо відомо, що об'єкт складається зі світлих пікселів і вони займають 20 % площі образу, то  $P_1 = 0,2$ . Необхідно, щоб

$$P_1 + P_2 = 1. \quad (3.11)$$

У цьому випадку це означає, що на іншу частину образу доводиться 80 % пікселів тла. Уведемо дві наступні функції від  $z$ :

$$\begin{aligned}
 d_1(z) &= P_1 p_1(z), \\
 d_2(z) &= P_2 p_2(z).
 \end{aligned}
 \tag{3.12}$$

Відповідно до теорії прийняття рішень відомо, що середня помилка визначення пікселя об'єкта як фоном (і навпаки) мінімізується за допомогою наступного правила: розглядаючи піксель зі значенням інтенсивності  $z$ , підставляється це значення  $z$  у рівняння (3.11) і (3.12). Потім визначається піксель як піксель об'єкта, якщо  $d_1(z) > d_2(z)$ , або як піксель тла, якщо  $d_2(z) > d_1(z)$ . Тоді оптимальний поріг визначається величиною  $z$ , для якої  $d_1(z) = d_2(z)$ . Таким чином, думаючи в рівняннях  $z = T$ , виходить, що оптимальний поріг задовольняє рівнянню

$$P_1 p_1(T) = P_2 p_2(T). \tag{3.13}$$

Отже, якщо відомі функціональні залежності  $p_1(z)$  і  $p_2(z)$ , це рівняння можна використати для знаходження оптимального порога, що відокремлює об'єкти від тла. Якщо цей поріг відомий, рівняння може бути використане для сегментації даного образу.

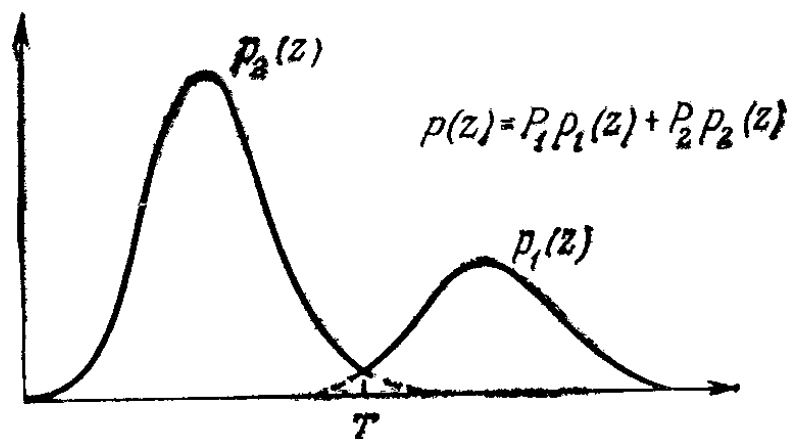


Рисунок 3.2 - Гістограма інтенсивності і її апроксимація у вигляді суми двох функцій щільності ймовірності



### 3.2.3 Визначення граничного рівня на основі характеристик границі

Одним з найбільш важливих аспектів при виборі граничного рівня є можливість надійно ідентифікувати модові піки для даної гістограми. Це важливо при автоматичному виборі граничного рівня в ситуаціях, коли характеристики образи міняються внаслідок великого розкиду інтенсивності. З викладеного вище треба, що можливість вибору «гарного» граничного рівня може бути істотно збільшена у випадку, якщо піки гістограм є високими, вузькими, симетричними й розділені глибокими провалами.

Одним з підходів для поліпшення виду гістограм є розгляд тільки тих пікселів, які лежать на границі (або біля її) між об'єктами й фоном. Одне з очевидних поліпшень полягає в тому, що цей підхід дозволяє одержувати гістограми менш залежними від відношення між об'єктом і фоном. Наприклад, гістограма інтенсивності образу, складеного з маленького об'єкта на великій площі постійного тла, визначалася б більшим піком внаслідок концентрації пікселів тла. З іншого боку, що результуючі гістограми мали б піки з більше збалансованими висотами, якби розглядалися пікселі, що лежать тільки на (або біля) границі між об'єктом і фоном. Крім того, імовірність розташування пікселя на границі об'єкта практично дорівнює ймовірності того, що він лежить на границі тла, що поліпшує симетрію гістограмних піків. Остаточо, як показано нижче, використання пікселів, які задовольняють деяким простим критеріям, заснованим на операторах градієнта й Лапласа, приводить до збільшення провалів між піками гістограм.

Вище неявно малосся на увазі, що границя між об'єктами й фоном відоме. Очевидно, що під час проведення сегментації ця інформація відсутня, оскільки знаходження роздягнула між об'єктами й фоном є остаточною метою наведеної тут процедури. Однак, обчисливши градієнт пікселя, можна визначити, чи лежить він на контурі. Крім того, лапласіан може дати інформацію про те, чи лежить даний піксель на темній (тобто фон) або

світлої (об'єкт) стороні контуру. Із внутрішньої сторони ідеального контуру лапласіан дорівнює нулю, тому на практиці можна чекати, що провали гістограм, утворених пікселями, обраними за критерієм градієнт/лапласіан, будуть розташовуватися досить нечасто й мати бажану висоту.

Градієнт  $G[f(x,y)]$  будь-якої крапки образу й лапласіан  $L[f(x, y)]$  можна використати для формування трирівневого образу:

$$S(x, y) = \begin{cases} 0, & \text{якщо } G[f(x, y)] < T, \\ +, & \text{якщо } G[f(x, y)] \geq T; uL[f(x, y)] \geq 0, \\ -, & \text{якщо } G[f(x, y)] \geq T; uL[f(x, y)] < 0, \end{cases} \quad (3.14)$$

де символи «0», «+», «-» представляють три різних рівні освітленості, а  $T$ —пороговий рівень.

Припустимо, що темний об'єкт розташовується на світлому тлі, тоді застосування рівняння дає образ  $s(x, y)$ , у якому всі пікселі, що не лежать на контурі (для них значення  $G[f(x, y)]$  менше  $T$ , позначені «0», всі пікселі на темній стороні контуру позначені «+» і всі пікселі на світлій стороні контуру позначені «-». Для світлого об'єкта на темному тлі символи «+» і «-» у рівнянні міняються місцями.

Викладена процедура може застосовуватися для створення сегментованого, бінарного образу, у якому 1 відповідає об'єктам, що представляють інтерес, і «0» для фонового зображення. Слід зазначити, що переміщення (уздовж горизонтальних або вертикальних ліній сканування) від світлого тла до темного об'єкта повинне характеризуватися заміною знака «-» тла на -1- об'єкта  $s(x, y)$ . Внутрішня область об'єкта складається з пікселів, позначених або «0» або «+». Остаточне переміщення від об'єкта до тла характеризується заміною знака «+» на «-».

### **3.2.4 Визначення граничного рівня, засноване на декількох змінних**

Викладені вище методи пов'язані з визначенням граничного рівня для єдиного змінного значення інтенсивності. У деяких додатках можна

використати більше однієї змінної для характеристики кожного пікселя образу, збільшуючи в такий спосіб не тільки ступінь розходження між об'єктом і фоном, але й між самими об'єктами. Одним з найбільш значимих прикладів є кольоровий зір, де використовуються червоні, зелені й блакитні компоненти (КЗГ) для формування складеного кольорового образу. У цьому випадку кожний піксель характеризується трьома змінними й це дозволяє будувати тривимірну гістограму. Основна процедура та ж, що й для однієї змінної. Нехай, наприклад, є три 16-уровневих зображення, що відповідають КЗГ компонентам датчика кольори. Сформуємо кубічні ґрати  $16 \times 16 \times 16$  і помістимо в кожний елемент пікселі, КЗГ компоненти яких мають інтенсивності, що відповідають координатам, що визначають положення цього елемента. Число крапок у кожному елементі ґрат може бути потім розділене на загальне число пікселів образу для формування нормованої гістограми.

Тепер вибір порога полягає в знаходженні груп крапок у тривимірному просторі, де кожна «компактна» група аналогічна основній моді гістограми однієї змінної. Наприклад, припустимо, що ми шукаємо дві значимі групи крапок даної гістограми, де одна група відповідає об'єкту, а інша - тлу. Беручи до уваги, що тепер кожний піксель має три компоненти й може бути розглянутий як крапка тривимірного простору, можна сегментувати образ за допомогою наступної процедури. Для кожного пікселя образу обчислюється відстань між цим пікселем і центром кожної групи. Тоді, якщо піксель розташовується поруч із центром групи крапок об'єкта, ми позначаємо його «1»; у протилежному випадку ми позначаємо його «0». Це поняття легко поширити на більшу частину компонентів пікселя й відповідно на більшу частину груп. Основна складність полягає в тому, що визначення значимих груп, як правило, приводить до досить складного завдання, оскільки число змінних зростає.

### 3.3 Областно - орієнтована сегментація

#### 3.3.1 Розширення області за рахунок об'єднання пікселів

Розширення області зводиться до процедури групування пікселів або підобластей у більші об'єднання. Найпростішою з них є агрегування пікселів. Процес починається з вибору безлічі вузлових крапок, з яких відбувається розширення області в результаті приєднання до вузлових крапок сусідніх пікселів зі схожими характеристиками (інтенсивність, текстура або кольори). Нехай цифри усередині осередків указують інтенсивність. Нехай крапки з координатами  $(3, 2)$  і  $(3, 4)$  використовуються як вузлові. Вибір двох початкових крапок приведе до сегментації образу на дві області: області  $R1$ , пов'язаної з вузлом  $(3, 2)$ , і області  $R2$ , пов'язаної з вузлом  $(3, 4)$ . Властивість  $P$ , що буде використатися для того, щоб віднести піксель до тієї або іншої області, полягає в тому, що модуль різниці між інтенсивностями пікселя й вузлової крапки не перевищує граничний рівень  $T$ . Любою піксель, що задовольняє цій властивості одночасно для обох вузлів, довільно попадає в область  $Ri$ . У цьому випадку сегментація проводиться для двох областей, причому крапки в  $R1$  позначаються буквою  $a$ , точки в  $R2$  буквою  $b$ . Необхідно відзначити, що незалежно від того, у який із цих двох областей буде взята початкова крапка, остаточний результат буде той самий. Якщо, з іншої сторони вибрати  $T = 8$ , була б отримана єдина область.

Попередній приклад, незважаючи на його простоту, ілюструє деякі важливі проблеми розширення області. Двома очевидними проблемами є: вибір початкових вузлів для правильного подання областей, що представляють інтерес, і визначення підходящих властивостей, для включення крапок у різні області в процесі розширення. Вибір безлічі, що складає з однієї або декількох початкових крапок, треба з постановки завдання. Наприклад, у військових додатках об'єкти, що представляють інтерес, мають більше високу температуру, чим фон, і тому проявляються

більш яскраво. Вибір найбільш яскравих пікселів є природним початковим кроком в алгоритмі процесу розширення області. При відсутності апріорної інформації можна почати з обчислення для кожного пікселя набору властивостей, що напевно буде використаний при встановленні відповідності пікселя тієї або іншої області в процесі розширення. Якщо результатом обчислень є групи точок (кластери), тоді в якості вузлових беруться ті пікселі, властивості яких близькі до властивостей центрів цих груп. Так, у прикладі, наведеному вище, гістограма інтенсивностей показала б, що точки з інтенсивністю від одного до семи є домінуючими. Вибір критерію подоби залежить не тільки від завдання, але також від виду наявних даних про образ. Наприклад, аналіз інформації, отриманої із супутників, істотно залежить від використання кольорів. Завдання аналізу значно ускладниться при використанні тільки монохроматичних образів. На жаль, у промисловому технічному зорі можливість одержання мультиспектральних і інших додаткових даних про образ є скоріше виключенням, чим правилом. Звичайно аналіз області повинен здійснюватися за допомогою набору дескрипторів, що включають інтенсивність і просторові характеристики (моменти, текстуру) одного джерела зображення. Відзначимо, що застосування тільки одних дескрипторів може приводити до неправильних результатів, якщо не використається інформація про умови зв'язку в процесі розширення області. Це легко продемонструвати при розгляді випадкового розташування пікселів із трьома різними значеннями інтенсивності. Об'єднання пікселів в «область» на основі ознаки однакової інтенсивності без обліку умов зв'язку приведе до безглузлого результату при сегментації.

Іншою важливою проблемою при розширенні області є формулювання умови закінчення процесу. Звичайно процес розширення області закінчується, якщо більше не існує пікселів, що задовольняють критерію приналежності до тієї або іншої області. Вище згадувалися такі критерії, як інтенсивність, текстура й кольори, які є локальними по своїй природі й не враховують «історію» процесу розширення області. Додатковий критерій, що

підвищує потужність алгоритму розширення області, включає поняття розміру, схожості між пікселем-кандидатом і тільки що створеними пікселями (порівняння інтенсивності кандидата й середньої інтенсивності області), а також форми області, що підлягає розширенню. Використання цих типів дескрипторів засновано на припущенні, що є неповна інформація про очікувані результати.

### 3.3.2 Розбивка і об'єднання області

Викладена вище процедура розширення області починає роботу із заданої безлічі вузлових крапок. Однак можна спочатку розбити образ на ряд довільних непересічних областей і потім поєднувати й/або розбивати ці області з метою задоволення умов. Ітеративні алгоритми розбивки й об'єднання, робота яких спрямована на виконання цих обмежень, можуть бути викладені в такий спосіб.

Нехай  $R$  є повною областю образу, на якій визначений предикат  $P$ . Один зі способів сегментації  $R$  складається в успішній розбивці площі образу на всі менші квадратні області, так що для кожної області  $R_i$ ,  $P(R_i) = \text{ІСТИНА}$ . Процедура починає роботу з розгляду всієї області  $R$ . Якщо  $P(R) = \text{НЕПРАВДА}$ , область розбивається на квадранти. Якщо для якого-небудь квадранта  $P$  приймає значення НЕПРАВДА, цей квадрант розбивається на підквадранти й т.д. Цей метод розбивки звичайно представляється у вигляді так названого квадродрева (дерева, у якого кожна вершина має тільки чотири нащадків). Відзначимо, що корінь дерева відповідає всьому образу, а кожна вершина - розбивці. У цьому випадку тільки  $R_4$  підлягає подальшій розбивці. Якщо застосовувати тільки операцію розбивки, очікується, що в результаті остаточної розбивки всієї площі образу на підобласті останні будуть мати однакові властивості. Це можна усунути припустимим об'єднанням так само, як і розбивкою. Для того щоб задовольнити умовам сегментації, уведеним вище, необхідно поєднувати тільки ті сусідні області, пікселі яких задовольняють предикату  $P$ , таким чином, дві сусідні області  $R_i$  і  $R_k$  поєднуються тільки в тому випадку, якщо  $P(R_i \cup R_k) = \text{ІСТИНА}$ .

Викладене вище можна представити у вигляді процедури, де на кожному кроці виконуються наступні операції:

1. Розбивка області  $R_i$ , для якого  $P(R_i) = \text{НЕПРАВДА}$ , на чотири непересічні квадранти.
2. Об'єднання сусідніх областей  $R_i$  і  $R_k$ , для яких  $P(R_i \cup R_k) = \text{ІСТИНА}$ .
3. Вихід на останок, коли подальше об'єднання або розбивка неможливо.

Можливі варіанти цього алгоритму. Наприклад, можна спочатку розбити образ на квадратні блоки. Подальша розбивка виконується по викладеному вище способу, але спочатку об'єднання обмежується групами із чотирьох блоків, що є у квадродреві нащадками й задовольняють предикатом  $P$ . Коли подальше об'єднання цього типу стає неможливим, процедура завершується остаточним об'єднанням областей відповідно до кроку 2. У цьому випадку поєднувані області можуть мати різний розмір. Основною перевагою цього підходу є використання одного квадродреву для розбивки й об'єднання до кроку, на якому відбувається остаточне об'єднання.

### **3.4 Застосування руху**

Рух являє собою потужний засіб, що використовується людиною й тваринами для виділення їхніх об'єктів, що цікавлять, із тла. У системах технічного зору рух використовується при виконанні різних операцій на конвеєрі, при переміщенні руки, оснащеної датчиком, більш рідко при переміщенні всієї робототехнічної системи.

#### **3.4.1 Основний підхід**

Один з найбільш простих підходів для визначення змін між двома кадрами зображення (образами)  $f(x, v, t_i)$  і  $f(x, v, t_j)$ , узятими відповідно в моменти часу  $t_i$  і  $t_j$ , ґрунтується на порівнянні відповідних пікселів цих двох образів. Для цього застосовується процедура, що полягає у формуванні так названої різниці образів.

Припустимо, що ми маємо еталонний образ, що має тільки стаціонарні компоненти. Якщо зрівняємо цей образ із у такий же спосіб, що має об'єкти, що рухаються, то різниця двох образів виходить у результаті викреслювання стаціонарних компонентів (тобто залишаються тільки ненульові записи, які відповідають нестаціонарним компонентам зображення).

Різниця між двома кадрами зображення, узятими в моменти часу  $t_i$  і  $t_j$ , можна визначити в такий спосіб:

$$d_{ij}(x,y) = \begin{cases} 1, \text{ якщо } |f(x, y, t_i) - f(x, y, t_j)| > \theta \\ 0, \text{ протилежному випадку} \end{cases} \quad (2.15)$$

де  $d_{ij}(x,y)$  – значення граничного рівня.

Відзначимо, що  $d_{ij}(x, y)$  приймає значення 1 для просторових координат  $(x, y)$  тільки в тому випадку, якщо два образи в крапці із цими координатами істотно розрізняються по інтенсивності, що визначається значенням граничного рівня  $\theta$ .

При аналізі образу, що рухається, всі пікселі зображень різниці  $d_{ij}(x, y)$ , що мають значення 1, розглядаються як результат руху об'єкта. Цей підхід примітимо тільки в тому випадку, якщо два образи зареєстровані й освітленість має відносно постійну величину в межах границь, установлюваних граничним рівнем  $\theta$ . На практиці запису в  $d_{ij}(x, y)$ , що мають значення 1, часто з'являються в результаті дії шуму. Звичайно на різниці двох кадрів зображення такі значення виглядають як ізольовані крапки. Для їхнього усунення застосовується простий підхід, що полягає у формуванні 4- або 8-зв'язних областей з одиниць в  $d_{ij}(x, y)$ , і потім зневажають будь-якою областю із числом записів, меншим за задане. При цьому можна не розпізнати малі й/або повільно, що рухаються об'єкти, але це збільшує ймовірність того, що записи, що залишаються, у різниці двох кадрів зображення дійсно відповідають руху.



### 3.4.2 Акумулятивна різниця

Як говорилося вище, різниця кадрів завдяки шуму часто містить ізольовані записи. Незважаючи на те що число таких записів може бути скорочене або повністю ліквідоване в результаті аналізу зв'язності граничних рівнів, цей процес може також привести до втрати зображень малих або повільно, що рухаються об'єктів. Нижче викладається підхід для рішення цієї проблеми шляхом розгляду зміни в розташуванні пікселів на декількох кадрах, тобто в процес уводиться «пам'ять». Основна ідея полягає в зневазі тими змінами, які виникають випадково в послідовності кадрів  $i$ , таким чином, можуть бути віднесені до випадкового шуму.

Розглянемо послідовність кадрів зображення  $f(x, y, t_1)$ ,  $f(x, y, t_2)$ , ...,  $f(x, y, t_n)$  і допустимо, що  $f(x, y, t_1)$  є еталонним образом. Зображення акумулятивної різниці формується в результаті порівняння еталонного образа з кожним образом у даній послідовності. У процедурі побудови зображення акумулятивної різниці є лічильник, призначений для обліку розташування пікселів. Його значення збільшується щораз, коли виникає розходження в розташуванні відповідних пікселів еталонного образа й образа з розглянутої послідовності. Таким чином, коли  $k$ -й кадр рівняється з еталонним, запис у даному пікселі акумулятивної різниці означає, у скільки разів інтенсивність пікселя  $k$ -го кадру відрізняється від інтенсивності пікселя еталонного образа. Розходження встановлюються, наприклад, за допомогою рівняння (2.15).

Нерідко корисно розглядати три типи зображень акумулятивної різниці: абсолютне, позитивне й негативне. Останні два виходять із рівняння (2.15), у якому немає модуля, а замість  $f(x, y, t_i)$  підставляється значення еталонного кадру. Припускаючи, що числові значення інтенсивності об'єкта перевищують значення тла у випадку, коли різниця позитивна, вона рівняється з позитивним значенням граничного рівня; якщо негативно, порівняння виконується з негативним значенням граничного рівня. Це визначення заміняється на протилежне якщо інтенсивність об'єкта менше

фона.

### **3.4.3 Визначення еталонного образу**

Успіх застосування методів залежить від еталонного образу, щодо якого проводяться подальші порівняння. Як уже говорилося вище, розходження між двома образами в завданні розпізнавання об'єктів, що рухаються, визначається шляхом виключення стаціонарних компонентів при збереженні елементів, що відповідають шуму й об'єктам, що рухаються. Проблема виділення образу із шуму вирішується методом фільтрації або за допомогою формування зображення акумулятивної різниці.

На практиці не завжди можна одержати еталонний образ, що має тільки стаціонарні елементи, і це приводить до необхідності побудови еталона з набору образів, що містять один або більше, що рухаються об'єктів. Це особливо характерно для ситуацій, що описують сцени з багатьма об'єктами, що швидко змінюються, або у випадках, коли виникають часті зміни сцен. Розглянемо наступну процедуру генерації еталонного образу. Припустимо, що ми розглядаємо перший образ послідовності в якості еталонного. Коли нестаціонарний компонент повністю вийшов зі свого положення в еталонному кадрі, що відповідає фон в даному кадрі може бути перенесений у положення, спочатку займане об'єктом в еталонному кадрі. Коли всі об'єкти, що рухаються, повністю покинули свої первісні положення, у результаті цієї операції відтвориться еталонний образ, що містить тільки стаціонарні компоненти. Переміщення об'єкта можна визначити за допомогою операції розширення позитивного зображення акумулятивної різниці.

### **3.5 Реалізація методу ідентифікації контурів об'єктів з використанням руху**

У випадку фіксованої області контролю зображення з використанням руху пропонується визначення порогового рівня через різницю інтенсивності точок контролюємого об'єкта і еталонної моделі зображення. Таким чином,

відповідність поточного стану зображення еталонному визначатиметься за ступенем подібності.

На рисунку 3.3 проілюстрований запропонований принцип для випадку перекриття усієї контрольованої області. При цьому усі сегменти області повинні мати однакові атрибути пік селів. При цьому усі сегменти, які знаходяться зі сторони перекриття, за атрибутами дорівнюються до сегментів контуру. Таким чином, при використанні порогового рівня відповідності сегментованого зображення еталонному існує можливість з визначеним припущенням установити момент перекриття контрольованої області в процесі руху.

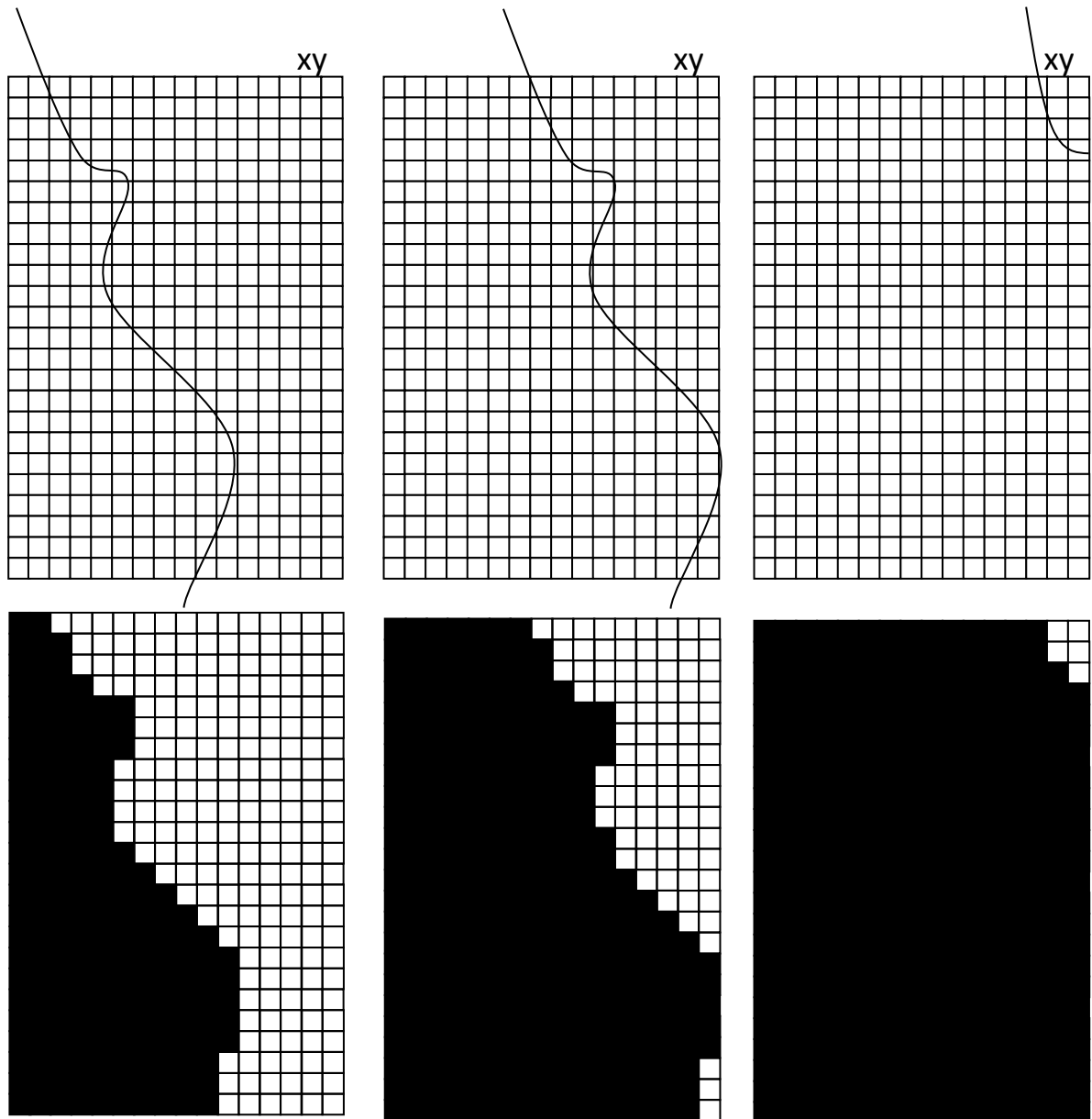


Рисунок 3.3 – Визначення моменту перекриття контрольованої області в

## процесі руху

З метою виділення шуму в процесі ідентифікації пропонується сумісна обробка послідовності кадрів. З урахуванням того що шум створює динамічні зміни атрибутів пік селів у контрольованому зображенні, порівняння послідовності кадрів дозволить визначити області, які змінюються несистематично і скорегувати ці одиничні помилки. На рисунку 3.4 проілюстрований принцип видалення помилок, викликаних появою шуму в потоці відеоінформації.

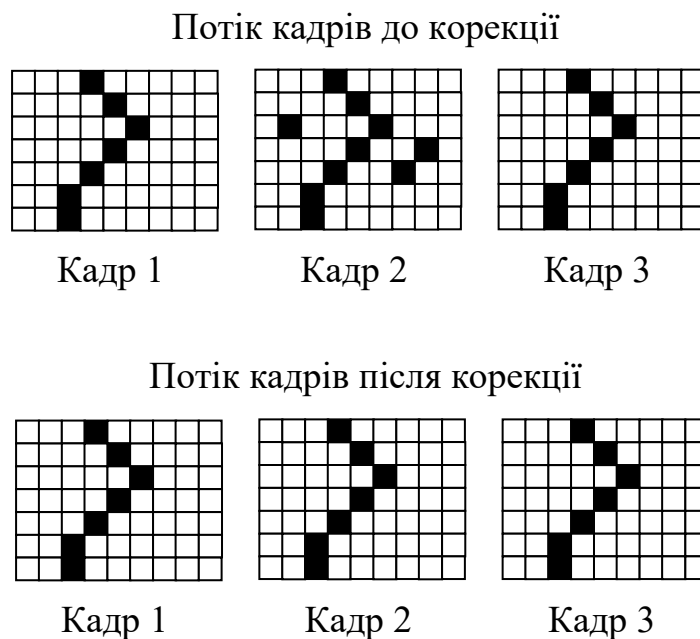


Рисунок 3.4 – Видалення шуму з потоку кадрів

### 3.5 Висновки

1. Алгоритми сегментації, як правило, ґрунтуються на двох фундаментальних принципах: розривності й подібності.
2. Алгоритми виявлення контурів супроводжуються процедурами побудови границь об'єктів з відповідних послідовностей пікселів.
3. Обчислювальна ефективність перетворення Хоуга полягає в поділі простору параметрів на так називані елементи, що збирають.
4. Глобальні пороги застосовуються в ситуаціях, коли є явне розходження між об'єктами й фоном і де освітленість досить однорідна.

Методи зворотної й структурованої освітленості, звичайно дають зображення, які можуть бути сегментовані шляхом застосування глобальних порогів. Але, як правило, довільне висвітлення робочого простору приводить до зображень, які, якщо виходити з визначення граничного рівня, вимагають локального аналізу для компенсації таких ефектів, як неоднорідність висвітлення, тіні й відбиття.

5. Можливість вибору оптимального граничного рівня може бути істотно збільшена у випадку, якщо піки гістограм є високими, вузькими, симетричними й розділені глибокими провалами.

6. Використання пікселів, які задовольняють деяким простим критеріям, заснованим на операторах градієнта й Лапласа, приводить до збільшення провалів між піками гістограми.

7. Двома очевидними проблемами розширення області при областно-орієнтованій сегментації є: вибір початкових вузлів для правильного подання областей, що представляють інтерес, і визначення підходящих властивостей, для включення крапок у різні області в процесі розширення. Вибір безлічі, що складає з однієї або декількох початкових крапок, треба з постановки завдання.

8. Запропонований новий метод ідентифікації контурів, який полягає у порівнянні еталонного зображення з об'єкту і дозволяє знизити часову складність процесу ідентифікації.

9. Запропонований новий метод видалення шуму з потоку кадрів в процесі руху об'єкту ідентифікації.

## РОЗДІЛ 4

### ТЕХНІЧНА РЕАЛІЗАЦІЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ІДЕНТИФІКАЦІЇ КОНТУРІВ ОБ'ЄКТІВ ЗА ДОПОМОГОЮ ЗАСОБІВ ТЕХНІЧНОГО ЗОРУ

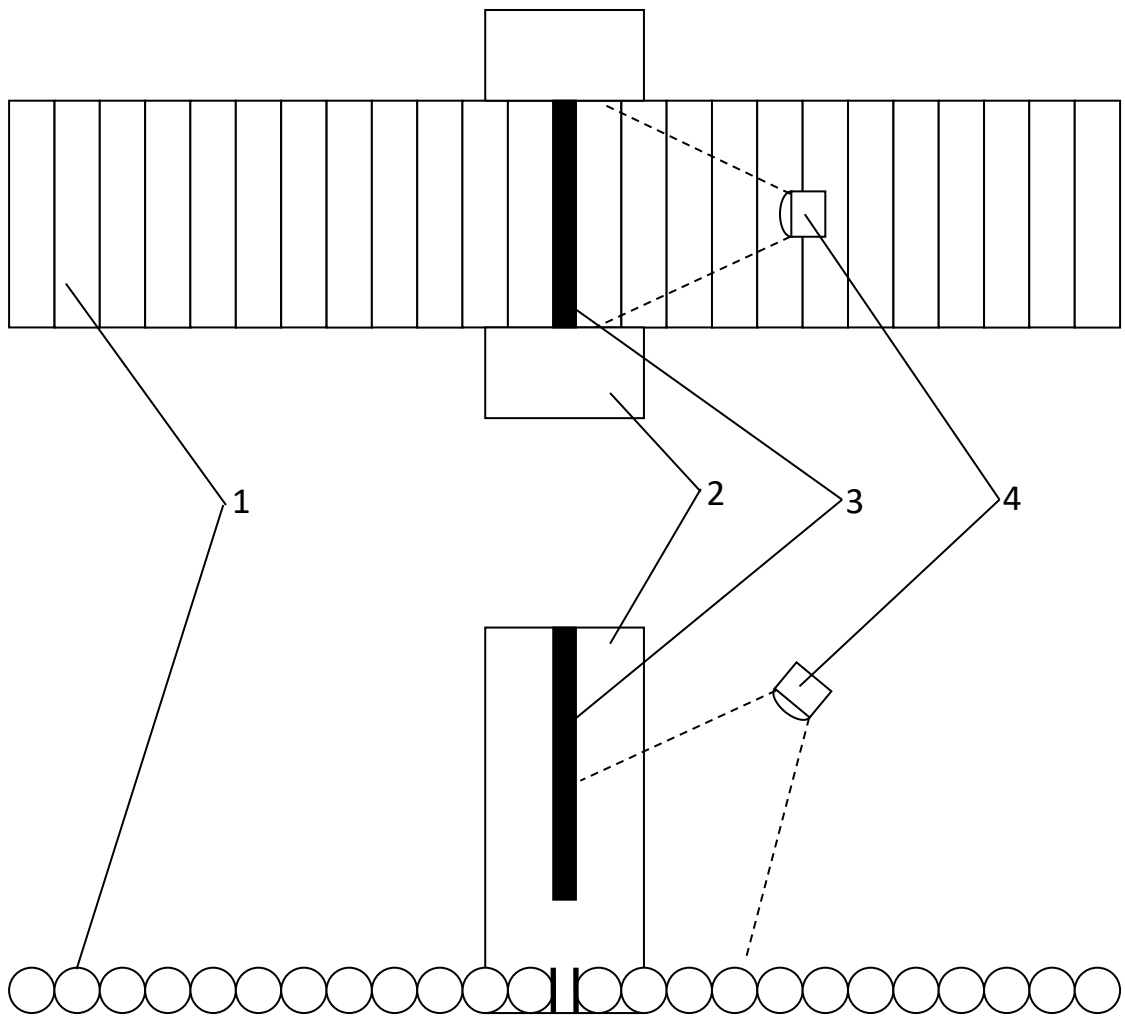
#### 4.1 Схеми розміщення прийомного відеопристрою автоматизованої системи ідентифікації контурів об'єктів

Виходячи з вищеописаного, для створення автоматизованої системи керування гильотинними ножицями металургійного виробництва, необхідно безконтактне (оптичне) визначення положення гарячого металу в режимі реального часу. Для реалізації із цієї мети системи технічного зору пропонується розміщення камери у відповідності зі схемою, наведеної на рисунку 4.1.

На підставі наведеної схеми, автоматизована система керування буде формувати сигнал керування приводом ножиців по настанню події перекривання площиною листа ріжучої крайки. У такий спосіб система ідентифікації контуру зображення буде відслідковувати момент повного перекриття крайки. Схема переміщення листа наведена на малюнку 4.2.

На рисунку 4.2 цифрами позначене: 1 - положення крайки листа в різні моменти часу при русі по прокатному стані; 2 - положення крайки листа, при якому система ідентифікації передає сигнал системі керування станом на зупинку; 3 - прокатний стан; 4 - ріжуча крайка.

Таким чином, робота системи ідентифікації контурів зображень повинна бути заснована на виділенні координат країв виробів розглянутими в попередній главі методами систем технічного зору. Зображення виробу повинне формуватися на позиційному фотоприймачі, оцифровуватися і оброблятися на промислової ЕОМ (можна типу РС-104). Результат обробки передається в систему керування прокатним станом по мережі. (Наприклад Ethernet).



1 - прокатні валки; 2 - опори ножиців; 3 - різальний інструмент; 4 - камера

Рисунок 4.1 - Схема розміщення камери для контролю торцевих крайок листа

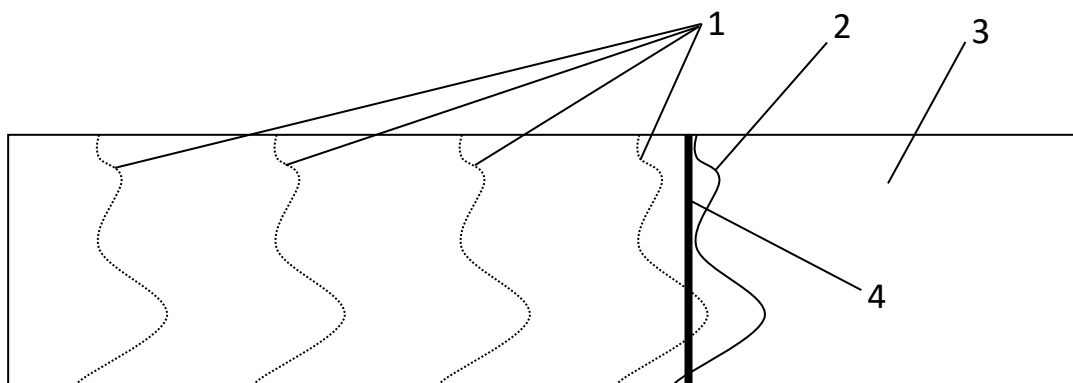


Рисунок 4.2 - Схема переміщення листа при прокатці

## 4.2 Структура автоматизованої системи ідентифікації контурів об'єктів

На основі вищесказаного, пропонується структура автоматизованої системи ідентифікації контурів об'єктів, яка наведена на рисунку 4.3.

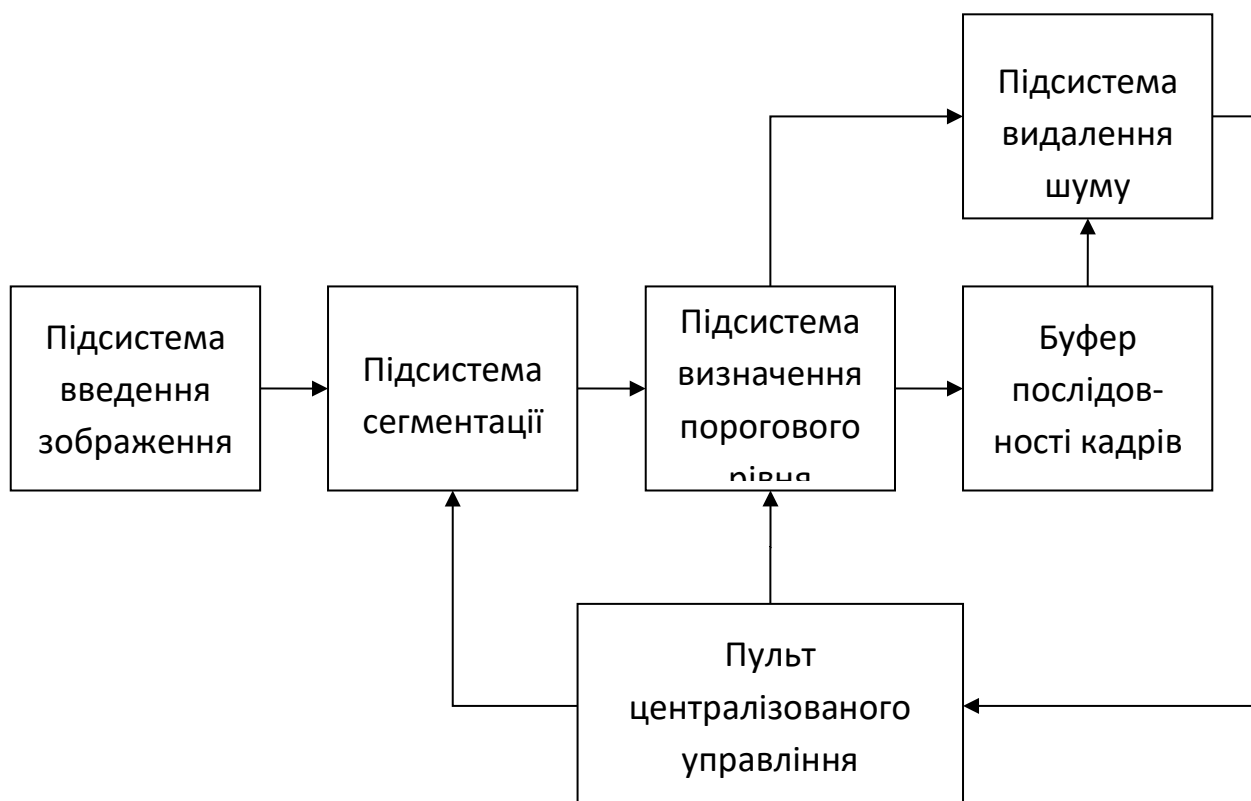


Рисунок 4.3 – Структура автоматизованої системи ідентифікації контурів об'єктів за допомогою засобів технічного зору

Управління усією системою здійснюється оператором через пульт централізованого управління, яка формує управляючу інформацію до Підсистеми сегментації, та підсистеми визначення порогового рівня. Послідовність кадрів, зображення об'єктів що ідентифікується фіксується підсистемою введення зображення. Надалі інформація надходить до підсистеми сегментації де здійснюється розбивка зображення на мінімальні визначені області. Далі сегментована послідовність кадрів надходить до підсистеми визначення порогового рівня, де здійснюється визначення атрибутів для кожного сегмента. З підсистеми визначення порогового рівня



інформація заноситься у буфер послідовності кадрів, а також до підсистеми видалення шуму, яка здійснює обробку потоку кадрів, відповідно до запропонованого в розділі 3 методу. Надалі оброблене зображення поступає до пульта централізованого управління в якому закладаються необхідні алгоритми обробки ідентифікованих об'єктів.

### **4.3 Структура програмного забезпечення автоматизованої системи ідентифікації контурів об'єктів**

Програмне забезпечення, яке потрібно встановити в приладі, повинне функціонувати на основі операційної системи реального часу й виконувати наступні функції:

- Ініціалізація всіх апаратних компонентів датчика;
- Сканування багатоелементного фотодіодного приймача із заданою частотою циклу обробки й одержання “кадру” зображення;
- Цифрова обробка отриманого кадру з виділенням центра об'єкта (положення петлі металу) і розміру об'єкта в одиницях дискретності приймача помножених на 1000 (тобто з поданням до 0.001 дискрети).

### **4.4 Висновки**

1. Запропонований метод контролю положення листа прокату при його позиціонуванні на гільйотинних ножицях, сутність якого полягає у використанні сучасних інформаційних технологій і засобів технічного зору.

2. Запропонований спосіб розміщення відеопристрою для аналізу контуру листа при прокаті, який дозволяє знизити кількість оптичних систем і знизити часову і алгоритмічну складність процесу ідентифікації контурів об'єктів.

3. Запропонований метод контролю процесу позиціонування листа прокату для обрізання торцевих кромek, який дозволяє знизити часову складність аналізу зображень.

4. Запропонована структура підсистеми введення зображення, яка полягає у обробці відео зображення у цифровому форматі з використанням запропонованих технічних рішень.

5. Запропонована мікропроцесорна підсистема управління і обробки інформації, яка дозволяє провести ідентифікацію контурів об'єктів з використанням запропонованих технічних рішень.

6. Запропонована структура автоматизованої системи ідентифікації контурів зображень за допомогою засобів технічного зору.

## ВИСНОВКИ

В результаті проведеної роботи було зроблено такі основні висновки:

1. На основі аналізу об'єкта автоматизації – гільйотинних ножиців для металургійного виробництва, можна зробити висновок, що необхідним є впровадження нових методів аналізу положення кромки листа при обрізанні.

2. Для якісного визначення моменту зупинки прокатного стану для обрізання листа необхідне використання методів, що дозволяють контролювати усю довжину кромки листа.

3. Використання засобів технічного зору дозволяє забезпечити необхідну якість контролю положення кромки листа при обрізанні.

4. Аналіз методів і способів формування зображення об'єктів показав, що кожен з них відрізняється максимальною точністю оцифрування і часовою складністю.

5. Для забезпечення максимальної точності і зниження часової складності аналізу зображення для ідентифікації контурів об'єктів доцільне використання ортогональної або слабо перспективної проекції.

6. Дескриптори не повинні залежати від розмірів, розташування й орієнтації об'єкта, але повинні містити достатню кількість інформації для надійної ідентифікації об'єктів.

7. Дослідження методів визначення дескрипторів границі показало, що нормування є точним тільки в тому випадку, коли самі границі інваріантні до повороту й зміни масштабу.

8. Точність нормування може бути збільшена, якщо вибирати довжини елементів ланцюга більшими, ніж відстані між пікселями дискретного образу, або ж вибирати орієнтацію ґрат уздовж головних осей кодуємого об'єкта.

9. Дискретну границю з довільною точністю можна апроксимувати багатокутниками. Для замкнутої кривої апроксимація є точною, коли число сегментів у багатокутнику дорівнює числу крапок границі, так що кожна

пара сусідніх крапок визначає сегмент багатокутника.

10. Існуючі системи технічного зору ґрунтуються на досить простих дескрипторах області, що робить їх більше привабливими з обчислювальної точки зору. Як варто очікувати, застосування цих дескрипторів обмежено ситуаціями, у яких об'єкти, що представляють інтерес, розрізняються настільки, що для їхньої ідентифікації досить кілька основних дескрипторів.

11. Як правило, складно виділити край, що має певне семантичне значення, оскільки для цього необхідна інформація більше високого рівня. Проте, досвід створення зорових систем підказує, що вивчення країв необхідно й корисно, тому варто знати, де ці краї проходять.

12. Основна проблема завдання визначення країв - це шуми зображення. Детектори країв побудовані таким чином, що дають великий вихід при різких змінах, а одна із причин виникнення різких змін - це додаток до пікселів сторонніх значень.

13. Ядро згладжування знижує дисперсію шуму, оскільки звичайно намагаються застосовувати позитивні ядра згладжування.

14. На деяких етапах визначення країв, особливо при немаксимальному придушенні, відкидається інформація, що надзвичайно складно відновити. Це відбувається внаслідок ухвалення важливого рішення - перевірити результат за допомогою граничного значення.

15. Алгоритми сегментації, як правило, ґрунтуються на двох фундаментальних принципах: розривності й подоби.

16. Алгоритми виявлення контурів супроводжуються процедурами побудови границь об'єктів з відповідних послідовностей пікселів.

17. Обчислювальна ефективність перетворення Хоуга полягає в поділі простору параметрів на так називані елементи, що збирають.

18. Глобальні пороги застосовуються в ситуаціях, коли є явне розходження між об'єктами й фоном і де освітленість досить однорідна. Методи зворотної й структурованої освітленості, звичайно дають зображення, які можуть бути сегментовані шляхом застосування глобальних

порогів. Але, як правило, довільне висвітлення робочого простору приводить до зображень, які, якщо виходити з визначення граничного рівня, вимагають локального аналізу для компенсації таких ефектів, як неоднорідність висвітлення, тіні й відбиття.

19. Можливість вибору оптимального граничного рівня може бути істотно збільшена у випадку, якщо піки гістограм є високими, вузькими, симетричними й розділені глибокими провалами.

20. Використання пікселів, які задовольняють деяким простим критеріям, заснованим на операторах градієнта й Лапласа, приводить до збільшення провалів між піками гістограми.

21. Двома очевидними проблемами розширення області при областно-орієнтованій сегментації є: вибір початкових вузлів для правильного подання областей, що представляють інтерес, і визначення підходящих властивостей, для включення крапок у різні області в процесі розширення. Вибір безлічі, що складає з однієї або декількох початкових крапок, треба з постановки завдання.

8. Запропонований новий метод ідентифікації контурів, який полягає у порівнянні еталонного зображення з об'єкту і дозволяє знизити часову складність процесу ідентифікації.

9. Запропонований новий метод видалення шуму з потоку кадрів в процесі руху об'єкту ідентифікації.

22. Запропонований метод контролю положення листа прокату при його позиціонуванні на гільйотинних ножицях, сутність якого полягає у використанні сучасних інформаційних технологій і засобів технічного зору.

23. Запропонований спосіб розміщення відеопристрою для аналізу контуру листа при прокаті, який дозволяє знизити кількість оптичних систем і знизити часову і алгоритмічну складність процесу ідентифікації контурів об'єктів.

24. Запропонований метод контролю процесу позиціонування листа прокату для обрізання торцевих кромek, який дозволяє знизити часову складність аналізу зображень.

25. Запропонована структура підсистеми введення зображення, яка полягає у обробці відео зображення у цифровому форматі з використанням запропонованих технічних рішень.

26. Запропонована мікропроцесорна підсистема управління і обробки інформації, яка дозволяє провести ідентифікацію контурів об'єктів з використанням запропонованих технічних рішень.

27. Запропонована структура автоматизованої системи ідентифікації контурів зображень за допомогою засобів технічного зору.

## ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Браус Р., Сэком Дж. Компьютрное зрение.: СПб., Вилиамс, 2004 – 1027с.
2. [www.inprise.com](http://www.inprise.com)
3. [www.opti-sytems.com.tw](http://www.opti-sytems.com.tw)
4. Антощук С.Г. Теоретичні і реалізаційні основи створення адаптивно-критеріальних систем побудови інформаційних технологій обробки візуальної інформації в АСУ: дис. на здобуття наукового ступеня доктора технічних наук 05.13.06 – Автоматизовані системи управління та прогресивні інформаційні технології., Одеса. – ОНПУ. – 2004.
5. [www.globalsystems.com](http://www.globalsystems.com).