

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації  
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем  
(повна назва)

## КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)  
(позначення документа)

Метод підготовки зображення обличчя до розпізнавання в режимі реального часу

(тема)

Виконав:

студент 2 курсу, групи СТМм-22-1

Вадим БСК

(прізвище, ініціали)

Спеціальність 171 Електроніка

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системи, технології і

комп'ютерні засоби мультимедіа

(повна назва освітньої програми)

Керівник проф. Олег СЕРГІЄНКО

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри

(підпис)

Володимир КАРТАШОВ

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформації  
Кафедра Медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 171 Електроніка

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма " Системи, технології і комп'ютерні засоби мультимедіа "

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ р.

**ЗАВДАННЯ**

**НА КВАЛІФІКАЦІЙНУ РОБОТУ**

Студентові Беку Вадиму Романовичу  
(прізвище, ім'я, по батькові)

1. Тема роботи Метод підготовки зображення обличчя до розпізнавання в режимі реального часу

Затверджена наказом по університету від " 20 " листопада 2023 р. № 1371 Ст

2. Термін здачі студентом закінченої роботи 08.01.2024

3. Вихідні дані до роботи:

1.Провести дослідження стиснення зображення, методи та алгоритми

2.Провести дослідження існуючих енкодерів, їх особливостей, налаштувань

3.Розробити рекомендації до ПК для проведення прямих трансляцій

4. Провести налаштування OBS для подальшого дослідження

5.Провести дослідження існуючих можливостей для проведення мультитрансляції

6. Визначити методику тестування енкодерів

7. Зробити порівняння якості вихідного відео при різних енкодерах

8. Зробити порівняння якості вихідного відео на різних платформах

4. Зміст розрахунково - пояснювальної записки (перелік питань, які підлягають розробці):

1. Аналітичний огляд методів підготовки зображень осіб у відеопотоці до розпізнавання

2. Підготовка зображень осіб у відеопотоці до розпізнавання

3. Тестування алгоритмів і результати експериментів

5. Перелік графічного матеріалу (з точним визначенням обов'язкових креслень):

1)Приклад автоматичної розмітки обличчя за допомогою алгоритму ASM; 2)Схема технології підготовки зображень облич до розпізнавання; 3)Схема алгоритму формування індивідуальної 3D моделі обличчя; 4)Помилки FAR та FRR залежно від обраного порога розпізнавання, одне зображення обличчя на людину; 5)Помилки FAR та FRR залежно від вибраного порога розпізнавання, 15 зображень на людину; 6)Помилки FAR та FRR залежно від обраного порога розпізнавання, 15 згенерованих алгоритмом зображень на людину.

6. Дата видачі завдання: 20.11.2023

## КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналітичний огляд методів підготовки зображень осіб у відеопотоці до розпізнавання	21.11.23-02.12.23	
2	Підготовка зображень осіб у відеопотоці до розпізнавання	04.12.23-09.12.23	
3	Тестування алгоритмів і результати експериментів	11.12.23-20.12.23	
4	Графічна частина роботи	21.12.23-23.12.23	
5	Перевірка керівником	24.12.23-26.12.23	
6	Перевірка на норм. контроль	27.12.23-28.12.23	
7	Перевірка на академічний плагіат	29.12.23-02.01.24	
8	Перевірка зав. кафедри, рецензування	03.01.24-08.01.24	

Студент



(підпис)

Вадим БСК

(прізвище, ім'я, по батькові)

Керівник роботи

(підпис)

проф. Олег СЕРГІЄНКО

(прізвище, ім'я, по батькові)

«\_\_» \_\_\_\_\_ 20\_\_ р.

## РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи має: 98с, 42 рис, 17 таблиць, 2 додатки, 55 джерела.

КОМП'ЮТЕРНА ГРАФІКА, ГРАФІЧНІ ЕЛЕМЕНТИ, ПЛАНАРНИЙ ТРЕКІНГ, ПІКСЕЛЬНИЙ ТРЕКІНГ, ADOBE AFTER EFFECTS CC, МЕТОДИ ВИКОРИСТАННЯ TRACK-POINTS.

Об'єктом дослідження у цій роботі є технологія розпізнавання осіб із зображенням особи для завдань реального часу.

Предмет дослідження – процес попередньої обробки та аналізу зображень осіб, які використовуються в алгоритмах розпізнавання, а також вплив цього процесу на точність та швидкість отримання кінцевого результату розпізнавання.

Мета роботи – підвищення точності та швидкості розпізнавання особистості у відеопотоці в режимі реального часу за допомогою технології підготовки зображень, що забезпечує компенсацію ракурсу та трекінг осіб.

Така технологія повинна оперувати складними вихідними даними – зображеннями осіб, які можуть мати низьку контрастність, різкість, мати ракурс, відмінний від фронтального, а також мати продуктивність, достатню для її застосування в комплексі з алгоритмом розпізнавання особи за зображенням обличчя у відеопотоці в режимі реального часу.

Для досягнення цих вимог необхідно вирішити такі завдання:

- Провести огляд існуючих методів аналізу зображень: оцінки схожості, контрастності та різкості, нормування освітлення та вибрати оптимальні.
- Дослідити можливість застосування 3D моделей особи людини у розпізнаванні осіб за зображенням обличчя в режимі реального часу та розробити алгоритм синтезування моделей осіб, що підходить для цих умов.
- Розробити алгоритм, що знижує вплив ракурсу на точність розпізнавання

особистості з зображення обличчя шляхом коригування курсу за допомогою синтезованих моделей та застосовний для завдань реального часу.

– Запропонувати та реалізувати алгоритм трекінгу особи у відеопотоці, що забезпечує ефективне стеження за особою та має швидкодію, що допускає його застосування в системах розпізнавання особистості в режимі реального часу.

– Розробити комплексну технологію автоматичного аналізу та обробки зображень осіб з метою їхньої підготовки до розпізнавання особистості у відеопотоці в режимі реального часу та подати її у вигляді комплексного алгоритму.

– Реалізувати розроблені алгоритми у вигляді прототипу програмної системи та провести обчислювальні експерименти з метою аналізу їх якості та ефективності.

Методи досліджень. У роботі використано методи теорії алгоритмів, теорії моделювання, методи планування чисельних експериментів, а також методи теорії ймовірностей та математичної статистики для кількісної обробки даних експериментів.

– Запропоновано комплексну технологію підготовки зображень осіб, що забезпечує функціонування програмних систем розпізнавання особистості у відеопотоці в режимі реального часу в автоматичному режимі, відмінними рисами технології є алгоритми компенсації ракурсу та трекінгу .

– Запропоновано модифікацію методу формування деформованої моделі особи, яка використовує принципи відновлення ландшафтних поверхонь методом інтерполяції Шепарда та синтезує 3D модель обличчя людини в автоматичному режимі.

– Розроблено алгоритм формування набору зображень осіб з одного фронтального зображення обличчя, що ґрунтується на застосуванні запропонованої модифікації методу DFM та синтезуванні проєкцій цієї моделі на площину.

– Розроблено алгоритм трекінгу осіб у відеопотоці, що комбінує методи background subtraction та Віоли-Джонса, що підвищує швидкість пошуку та трекінгу осіб у порівнянні з алгоритмом, який використовує лише метод Віоли-Джонса.

## ABSTRACT

The explanatory note of the qualification work must: 98with, 42Fig, 17table, 2addition, 55sources

COMPUTER GRAPHICS, GRAPHIC ELEMENTS, PLAN TRACKING, PIXEL TRACKING, ADOBE AFTER EFFECTS CC, METHODS OF USING TRACK-POINTS.

The object of research in this work is face recognition technology with a face image for real-time tasks.

The subject of the study is the process of preprocessing and analysis of images of persons used in recognition algorithms, as well as the impact of this process on the accuracy and speed of obtaining the final result of recognition.

The goal of the work is to increase the accuracy and speed of recognition of a person in a video stream in real time using image preparation technology that provides angle compensation and person tracking.

Such a technology must operate with complex raw data - images of persons that may have low contrast, sharpness, have an angle different from the frontal one, and also have performance sufficient for its use in combination with an algorithm for recognizing a person based on a face image in a real-time video stream .

To achieve these requirements, it is necessary to solve the following tasks:

- Conduct an overview of existing methods of image analysis: evaluation of similarity, contrast and sharpness, normalization of lighting and choose the optimal ones.
- To investigate the possibility of using 3D models of the human face in real-time face recognition and to develop an algorithm for synthesizing face models suitable for these conditions.
- To develop an algorithm that reduces the influence of perspective on the accuracy of recognition of a person from a face image by course correction using synthesized

models and is applicable for real-time tasks.

– To propose and implement an algorithm for tracking a person in a video stream, which provides effective tracking of a person and has a speed that allows it to be used in real-time personality recognition systems.

– To develop a complex technology of automatic analysis and processing of images of persons in order to prepare them for real-time recognition of a person in a video stream and to present it in the form of a complex algorithm.

– Implement the developed algorithms in the form of a software system prototype and conduct computational experiments to analyze their quality and efficiency.

Research methods. The work uses methods of algorithm theory, modeling theory, methods of planning numerical experiments, as well as methods of probability theory and mathematical statistics for quantitative processing of experimental data.

– A complex technology for the preparation of images of persons is proposed, which ensures the functioning of software systems for personality recognition in the video stream in real time in automatic mode, the distinguishing features of the technology are angle compensation and tracking algorithms.

– A modification of the method of forming a deformed face model is proposed, which uses the principles of restoration of landscape surfaces by the Shepard interpolation method and synthesizes a 3D model of a person's face in automatic mode.

– An algorithm for forming a set of face images from one frontal image of a face has been developed, which is based on the application of the proposed modification of the DFM method and the synthesis of projections of this model on the plane.

– An algorithm for tracking persons in a video stream has been developed, which combines the methods of background subtraction and Viola-Jones, which increases the speed of searching and tracking persons in comparison with the algorithm that uses only the method of Viola-Jones.

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

- ADAS – Advanced driver-assistance systems – Удосконалена система допомоги водієві ;
- ЛЕП – Лінії електропередач;
- ПЗ – Програмне забезпечення;
- ІТ – Information Technology – Інформаційні технології;
- МП – Мова програмування;
- ОС – Операційна система;
- ПК – Персональний комп’ютер;
- ASCII – American standard code for information interchange – Таблиця кодування символів;
- IEEE – Institute of Electrical and Electronics Engineers – Некомерційна інженерна асоціація в США;
- OpenCV – Open Source Computer Vision Library – Відкрита бібліотека для роботи з алгоритмами комп’ютерного зору, машинним навчанням та обробкою зображення;
- MRT – Магнітно-резонансна томографія;
- КТ – Комп’ютерна томографія;
- ПЕТ – Позитронно-емісійна томографія;
- RGB – red, green, blue – Адитивна кольорова модель;
- LAB – l, a, b – Система координат з трьох осей;
- HOG – Histogram of oriented gradients – Гістограма напрямлених градієнтів;
- SIFT – Scale-Invariant Feature Transform – Алгоритм комп’ютерного бачення для виявлення, описування та зіставлення локальних ознак;
- SVM – Support Vector Machines – Набір подібних алгоритмів виду «навчання з учителем», що використовуються для завдань класифікації та регресійного аналізу;
- YOLOv4 – Нейронна мережа для виявлення об’єктів;
- НМ – Нейронна мережа.

## ЗМІСТ

ВСТУП.....	11
1 АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ ПІДГОТОВКИ ЗОБРАЖЕНЬ ОСІБ У ВІДЕОПОТОЦІ ДО РОЗПІЗНАВАННЯ.....	15
1.1 Актуальність дослідження, огляд та аналіз стану проблеми .....	15
1.1.1. Трекінг .....	15
1.1.2. Порівняння зображень .....	20
1.1.3. Оцінка контрастності .....	21
1.1.4. Оцінка різкості.....	23
1.1.5. Компенсація освітлення .....	28
1.1.6. Контроль ракурсу.....	31
1.2 Аналіз .....	37
2 ТЕХНОЛОГІЯ ПІДГОТОВКИ ЗОБРАЖЕНЬ ОСІБ У ВІДЕОПОТОЦІ ДО РОЗПІЗНАВАННЯ .....	40
2.1 Вибір методів, що використовуються у складі технології.....	40
2.1.1. Трекінг .....	40
2.1.2. Порівняння зображень .....	45
2.1.3. Методи оцінки контрастності та різкості .....	46
2.1.4. Метод корекції освітлення .....	47
2.1.5 Вибір засобів розробки .....	47
2.1.6 Загальний вид алгоритму, що описує технологію підготовки та обробки зображень .....	50
2.2 Програмна реалізація алгоритмів .....	58
3 ТЕСТУВАННЯ АЛГОРИТМІВ І РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТУ ТОВ..	62
3.1 Підготовка експериментів.....	62
3.2 Тестування програмної реалізації алгоритмів.....	63
3.2.1 База для тестування алгоритмів .....	63
3.2.2 Алгоритм побудова набору зображень облич .....	64

3.2.3 Трекінг .....	66
3.2.4. Фільтр по різкості та контрастності зображень осіб .....	67
3.2.5 Алгоритм формування індивідуальної тривимірної моделі обличчя .....	68
3.2.6 Вирівнювання освітлення та методи отримання ознак .....	69
3.2.7 Результати тестування комплексної технології підготовки зображень облич до розпізнавання .....	72
3.3 Порівняння результатів тестування запропонованої технології з існуючими системами розпізнавання .....	87
ВИСНОВОК .....	91
ПЕРЕДІК ДЖЕРЕЛ ПОСИЛАННЯ .....	92
ДОДАТКИ .....	98
ДОДАТОК А .....	99
ДОДАТОК Б .....	105

## ВСТУП

Роботи з вирішення завдання ідентифікації особистості на основі аналізу зображень особи ведуться з ранніх етапів розвитку комп'ютерного зору. Завдання розпізнавання осіб є одним із перших практичних завдань, яке послужило стимулом для розвитку теорії розпізнавання об'єктів. Цей напрямок з'явився на початку 1980-х років, але його розвиток розпочався у 1990-х роках у процесі створення інформаційно-пошукових систем розпізнавання для ідентифікації особистості. Останнім часом у різних сферах діяльності зростає потреба швидкої та правильної некооперативної ідентифікації особистості людини у відеопотоці в режимі реального часу.

Загалом завдання ідентифікації особистості відеопотоці може бути розділена на кілька етапів:

- Пошук області обличчя на зображенні;
- Спостереження за положенням виявленої особи на послідовності кадрів (трекінг);
- Попереднє оброблення зображення обличчя з метою компенсації умов зйомки;
- Порівняння зображення обличчя із зображеннями в базі зображень.

Завдання пошуку осіб на зображенні на сьогоднішній день успішно вирішено та застосовується у багатьох технічних пристроях. Наприклад, у фототехніці пошук осіб використовується для автоматичного вибору точки фокусування.

Трекінг є необов'язковим етапом, але дозволяє співвідносити зображення осіб людей на різних кадрах і робити розпізнавання за найліпшим зображенням з послідовності кадрів, побічно підвищуючи точність ідентифікації. Завдання трекінгу широко освітлене в літературі і має безліч варіантів застосування в задачах комп'ютерного зору, крім спостереження за особами людей.

Завдання обробки зображень та автоматизованого розпізнавання (ідентифікації) осіб є більш складними, і на даний момент існуючі алгоритми наближаються до точності розпізнавання осіб людиною тільки в заданих умовах, або вимагають обчислювальні ресурси, недоступні при обробці відеопотоку в режимі реального часу.

Розвиток теорії та практики розпізнавання об'єктів багато в чому було обумовлено досягненнями зарубіжних учених: Дж. Філіпса (NIST, США) та Г. Векслера (університет Джорджа Мейсона), що працюють у цій галузі вже понад 15 років, а також Такео Канаде, М. Турка, П. Віоли, М. Джонса, Б. Лукаса, Дж. Сяо, С. Бейкера, Д. Команіцу, Г. Брадські та інших.

В останні роки було запропоновано низку різних підходів до обробки, локалізації та розпізнавання об'єктів, таких як метод головних компонент (Principal Component Analysis, PCA), нейронні мережі, еволюційні алгоритми, алгоритм AdaBoost (Йоав Фройнд, Роберт Шапіре), метод опорних векторів та інші. Точність, надійність та швидкість цих підходів серйозно залежить від шумів на відеопослідовності та різних умов зйомки. Методи, які застосовуються для вирішення задачі розпізнавання осіб у режимі реального часу, повинні забезпечувати прийнятну точність розпізнавання та високу швидкість обробки відеопослідовностей.

Число помилок ідентифікації особистості у всіх сучасних систем розпізнавання осіб у режимі реального часу має явну залежність від наступних факторів:

- якість зображення особи, яка перебуває як у базі даних, так і реєстрованого камерами відеоспостереження;
- перекриття особи сторонніми об'єктами та інші перешкоди на зображенні;
- ракурс зйомки особи, що ідентифікується;
- умови висвітлення.

Зображення аналізованих осіб повинні мати мінімально необхідну якість, найчастіше вимоги до зображення включають такі пункти, як:

- відсутність ефекту змазаного зображення, викликаного рухом об'єкта, що ідентифікується;
- зображення аналізованих осіб має бути менше мінімально необхідного розміру (визначається з відривом між зіницями).

Для вирішення цих проблем існує два шляхи: розширення бази зображень з різними умовами зйомки (що не завжди можливо зробити), а також моделювання умов зйомки на еталонних зображеннях при їх порівнянні з аналізованими.

Прикладом другого підходу є використання тривимірних (3D) моделей осіб та порівняння окремо рельєфу та текстури одержуваних 3D об'єктів. Цей підхід надає широкі можливості для моделювання ракурсу та освітлення, але вимагає створення 3D моделі кожної особи в базі даних, що є ресурсномісткими та нетривіальним завданням.

У багатьох сферах діяльності потрібне некооперативне розпізнавання особи людини, наприклад пошук підозрілої людини за коротким відеозаписом поганої якості, на якому він не дивиться прямо в об'єктив камери, присутні шуми, бічне засвітлення. Це ускладнює застосування як класичних двовимірних (2D) методів розпізнавання, так і методів, що використовують 3D моделі.

Існуючі швидкі 2D методи та алгоритми досягають гарної точності тільки при дотриманні жорстких вимог щодо ракурсу зображень обличчя та умов зйомки, що обумовлено механізмами кластеризації та машинного навчання.

Якщо вплив умов зйомки, таких як освітлення, яскравість і контрастність зображення, в більшості випадків можна компенсувати за допомогою існуючих методів та алгоритмів обробки зображень, то завдання компенсації ракурсу вирішується неефективно, і часто рішення зводяться до використання обладнання, що дозволяє отримувати 3D моделі осіб для подальшого використання у розпізнаванні. Методи, розраховані на побудову 3D моделей з відеопотоку, зазвичай не розглядаються для використання в системах реального часу, оскільки для

створення 3D моделі високої якості в режимі реального часу потрібні значні обчислювальні ресурси.

Умови освітлення повинні дозволяти реєструвати зображення осіб, на яких мають бути помітні основні риси обличчя (достатній контраст). Для всіх сучасних систем ідентифікації важливою є наявність прямого або близького до нього кута падіння променів світла на фронтальну площину обличчя, а також наявність розсіяного джерела світла (умови наближені до умов отримання фотографій для офіційних документів). Вкрай негативний вплив на якість ідентифікації має наявність бічного засвітлення обличчя.

Враховуючи перераховані вище проблеми методів ідентифікації особи по растровому зображенню особи в режимі реального часу, можна зробити висновок про необхідність дослідження існуючих методів попередньої обробки зображень, що надходять, як єдиної технології. Така технологія повинна містити в собі деяку сукупність методів та алгоритмів за попередньою обробкою та оцінкою зображень осіб, підібраних під оптимальне співвідношення якості та швидкості обробки зображень для конкретної задачі, що розв'язується.

Попередня обробка може проводитися за кількома критеріями, такими як контрастність, різкість, освітлення, ракурс, схожість зображень. Технологію попередньої обробки можна представити у вигляді комплексного алгоритму для завдання обробки відеопотоку, в тому числі і в режимі реального часу, який повинен враховувати всі критерії та мати достатню швидкодію.

Проблема попередньої обробки зображень найчастіше розглядається як набір приватних завдань щодо покращення якості, компенсації освітлення та ракурсу зйомки, але не як складна система взаємозалежних алгоритмів.

Таким чином, необхідність оптимізації всіх кроків обробки та аналізу зображень осіб під конкретне завдання розпізнавання, чи то фотографії, чи відеопотік, створюють передумови для створення комплексної системи алгоритмів. Саме тому тема роботи є актуальною.

# 1 АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ ПІДГОТОВКИ ЗОБРАЖЕНЬ ОСІБ У ВІДЕОПОТОЦІ ДО РОЗПІЗНАВАННЯ

## 1.1 Актуальність дослідження, огляд та аналіз стану проблеми

Практично всі приватні завдання комп'ютерного зору, пов'язані з попередньою обробкою зображень, такі як трекінг, порівняння зображень, оцінка контрастності і різкості, компенсація освітлення, контроль курсу, в тому чи іншому вигляді розглядалися в багатьох роботах, однак їх - ліз з погляду застосування у системах розпізнавання особистості режимі реального часу практично ніким не піднімався. Це обумовлює необхідність проведення аналітичного огляду існуючих методів по кожному з цих напрямів та визначення тих областей, у яких доцільно скористатися раніше запропонованими напрацюваннями, а також тих, у яких потрібний перегляд підходу з точки зору автоматизації алгоритмів та підвищення швидкості їх роботи.

### 1.1.1. Трекінг

Існує кілька методів розв'язання цього завдання, проте найбільшого поширення набув алгоритм П. Віоли та М. Джонса, який був опублікований у 2001 році [44]. Алгоритм не спеціалізується на виділенні лише осіб, він може бути використаний для знаходження та інших класів об'єктів, для чого має бути використаний відповідний класифікатор. Тим не менш, найбільш широке застосування він отримав саме для вирішення задачі виявлення осіб на растровом зображенні[45].

Детектування осіб використовує три ключові методи:

- згортка функцій Хаара з вихідним зображенням;
- машинне навчання AdaBoost;

– каскадний класифікатор для ефективного поєднання множинних функцій.

Алгоритм Віоли-Джонса використовує вейвлети Хаара, які є прямокутними хвилями однакової довжини (один високий і один низький інтервал). Для вирішення поставленої задачі використовується ядра згортки, деякі з яких схематично показані на (рис. 1).

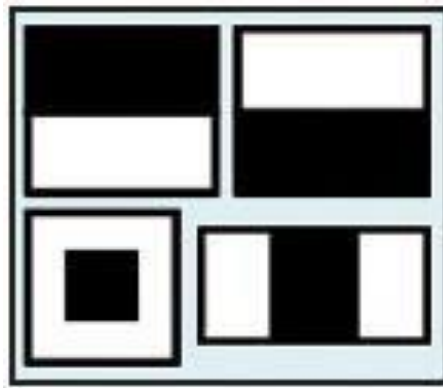


Рисунок 1.1 - Приклади вейвлетів Хаара, білого кольору відповідає 1, а чорному -1 [2]

Якщо результат згортки перевищує заданий поріг, який визначається в процесі навчання класифікатора, вважається, що на даній ділянці зображення є функція Хаара. На (рис. 1.2) показано приклади розміщення відповідних вейвлетів Хаара на зображенні обличчя.

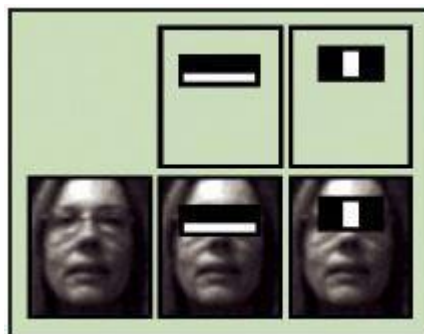


Рисунок 1.2 - Приклади розміщення відповідності перших двох вейвлетів Хаара на зображенні особи [2]

На основі методу машинного навчання AdaBoost проводиться підбір вейвлетів Хаара та зіставлення ним порогових величин, поєднання яких характеризує наявність на зображенні особи. Розмір вейвлетів Хаара визначає розміри шуканих осіб, для пошуку всіх можливих розмірів осіб використовують різні масштаби вейвлетів Хаара - піраміда.

У разі використання алгоритму Віюлі-Джонса для поставленого завдання на кожному кадрі зображення буде знайдено набір прямокутних областей, які відповідають знайденим особам.

Стеження за становищем особи.

Як правило, одна і та ж особа присутня в полі зору камери відеоспостереження протягом деякого часового інтервалу, тому її зображення може бути знайдено в послідовності кадрів. Це дозволяє здійснити трекінг шуканої особи.

Трекінг (tracking) – це визначення розташування об'єкта, що рухається, у відеопотоці. Алгоритм аналізує кадри відеопотоку і видає положення цільових об'єктів, що рухаються щодо кадру, дозволяючи скласти послідовність зображень конкретного об'єкта, яка називається треком (track). Приклад трекінгу об'єктів, що рухаються, представлений на (рис. 1.3).

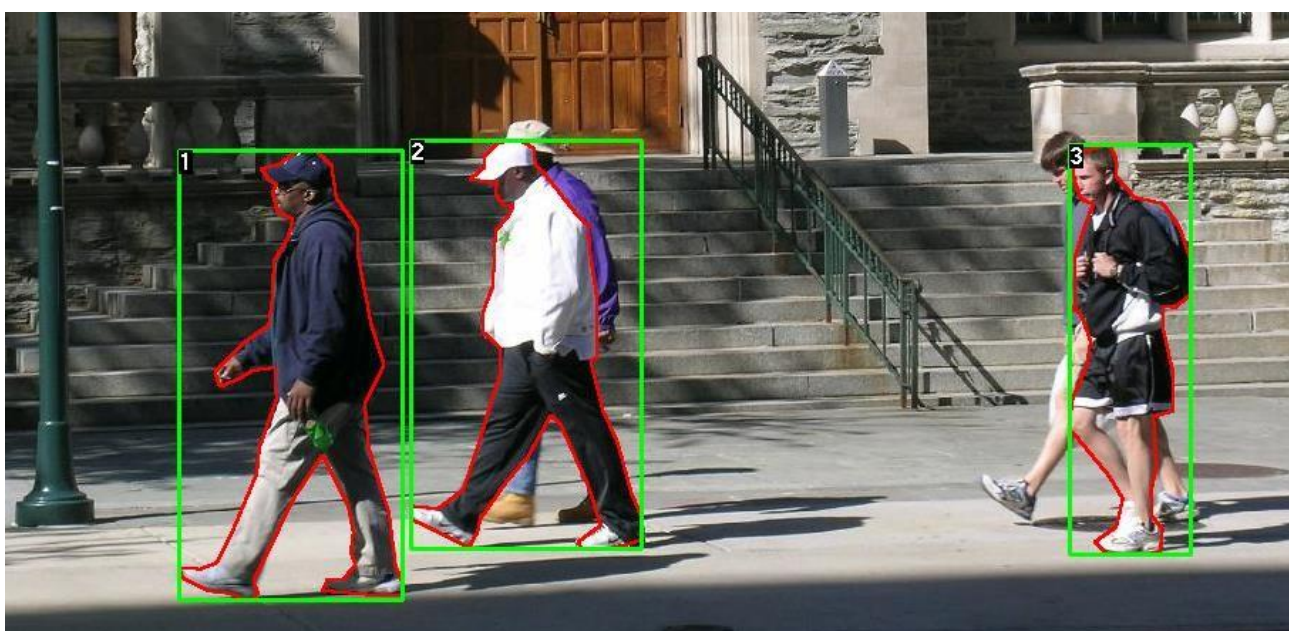


Рисунок 1.3 - Приклад трекінгу об'єктів, що рухаються [3]

Рух об'єкта на відео класифікується на регулярний (рух об'єкта по безперервній траєкторії) та нерегулярний (розриви траєкторій внаслідок заслонення об'єкта стеження іншими об'єктами, стрибка об'єкта та інших причин). У разі регулярного руху об'єкта трекер розглядається як динамічна система [4].

Алгоритми стеження (трекінгу), що оцінюють положення одного або кількох цільових об'єктів (об'єктів стеження) на кадрах відеопослідовності, входять до численних програм комп'ютерного зору в робототехніці та відеоспостереженні [7,8,9]. Знаходження та відстеження рухомих об'єктів у відеопотоці є важливим завданням для таких додатків, як охоронні системи, системи аналізу потоку машин на автомагістралі, smart room тощо. [5,6].

Характеристики цільового об'єкта в реальному відео з часом, як правило, змінюються через природні умови відеозйомки, коли, наприклад, під час зйомки змінюється освітленість об'єкта, змінюються відносні розміри об'єкта в послідовності кадрів при видаленні/наближенні об'єкта щодо камери, об'єкт стеження заступається іншими об'єктами тощо.

Часто першим кроком вирішення цих завдань є виділення переднього плану. Наприклад, у роботах [7,8,9] розглядають методи віднімання фону (background subtraction), що застосовуються для вирішення цього завдання. Методи віднімання фону - це методи, які будують модель заднього плану сцени і на кожному кадрі роблять попіксельне порівняння поточного значення в кожному пікселі з моделлю і відповідно до результатів порівняння приймають рішення, належить піксель до заднього плану або переднього. Основні проблеми, з якими стикаються подібні методи – це змінний задній план, зміна освітлення, тіні, що рухаються, і шум камери (рис. 1.4).



Рисунок 1.4 - Приклад трекінгу методом віднімання фону з придушенням тіней [5]

Також існують алгоритми, засновані на імовірнісних моделях, які покликані вирішувати частину цих проблем, зокрема проблему виділення тіней об'єктів, що рухаються.

У розв'язуваній задачі стеження обличчя людини більшість описаних проблем не є актуальною через те, що пошук об'єкта у разі здійснюється з допомогою добре вивченого і перевіреного алгоритму Віоли-Джонса, котрого розроблено безліч алгоритмів якісного трекінгу [5,6].

Так як трекінг використовується виключно з метою знизити обсяг обчислень за рахунок поєднання послідовностей зображень осіб у відео-потоці (або з метою отримати набір з кількох зображень однієї особи), то до нього не слід висувати

вимоги в абсолютній точності та застосування до безлічі різних об'єктів. У цьому випадку є сенс використовувати алгоритм, що оперує з інформацією про положення зображень осіб у кадрі з додатковим попереднім аналізом сусідніх кадрів та фону зображення, що дозволить знизити обчислювальне навантаження на систему розпізнавання детектором осіб та співвідносити більшу частину зображень осіб на двох сусідніх. кадрів.

### 1.1.2. Порівняння зображень

У більшості систем відеоспостереження використовується стандартна частота надходження кадрів, що дорівнює двадцяти п'яти кадрам в секунду. Очевидно, що зображення обличчя однієї людини на двох сусідніх кадрах практично не повинні відрізнятися один від одного. Біометричний аналіз однакових зображень не виправданий з погляду завантаження обчислювальних ресурсів. У зв'язку з цим доцільно проводити відбір кадрів, у яких зображення однієї особи істотно відрізняються друг від друга.

Існує кілька основних методів порівняння зображень.

- Застосування хеш-функцій до зменшеної копії зображення та їх порівняння. Даний метод має гарну швидкодію, але дає дуже грубу оцінку, яка безпосередньо залежить від зменшення зображення, тому може застосовуватися тільки для пошуку віддаленої візуальної схожості об'єктів, наприклад, пошуку схожих картинок у колекції. Метод відноситься до пошуку зображень по вмісту (англ. Content-based image retrieval (CBIR)) [7,8] – розділ комп'ютерного зору, який вирішує завдання пошуку зображень, які мають необхідний зміст, у великому наборі цифрових зображень.

- Обчислення коефіцієнта кореляції по одній із формул. Цей метод досить чутливий до невеликих змін зображення,

проте вимогливий до обчислювальних ресурсів. Його застосування обмежено специфічними областями роботи із зображеннями [5,6].

- Порівняння зображень через побудову та аналіз SURF-дескрипторів, які

описують деякі невеликі області зображення. Дослідження показують, що даний метод показує найкращі результати при порівнянні зображень об'єктів, знятих під різними кутами та в різних умовах освітлення, проте він вимагає у кілька разів більше обчислювальних ресурсів, ніж порівняно трудомісткий метод обчислення коефіцієнта кореляції [61]. Існує модифікація даного методу (BRISK), яка має високу швидкодію [62], проте цей метод складно застосувати для пошуку малопомітних відмінностей у кадрах.

### 1.1.3. Оцінка контрастності

Перед тим, як покращувати зображення, необхідно оцінити його якість. Людина, кинувши погляд на зображення, може сказати яскраве воно чи темне, контрастне чи ні, чітке чи розмите тощо. Алгоритми працюють детально, аналізуючи зображення по піксельно або невеликими групами пікселів. Тому, виходячи з роботи алгоритму, важко дати загальну оцінку якості зображення.

Контраст (фр. *contraste*) – в оптиці (сенситометрії та фотометрії) різниця в характеристиках різних ділянок зображення, здатність фотографічного матеріалу або оптичної системи відтворювати цю різницю, а також характеристика чутливості ока (зорової системи) щодо яскравості та кольору .

Контрастність (також, у різних контекстах вживається і саме слово контраст і коефіцієнт контрасту) – ступінь контрасту, найчастіше виражається безрозмірною величиною, ставленням або логарифмом відносин.

Мірою контрастності є величина

$$k = \frac{B_b - B_s}{B_s},$$

де,  $B_b$  – яскравість фотона.

де  $B_s$  - яскравість фото

Контрастність зображення є показником його інформативності, оскільки цей показник враховує освітленість та помітність об'єктів на зображенні. Цей параметр

недостатній, щоб можна було визнати зображення якісним, оскільки бічна засвітка об'єктів дає високу оцінку контрасту, але ускладнює виділення деталей зображення.

Проте, якщо зображення обличчя вдається нормалізувати, контрастність стає більш корисною оцінкою.

Контрастність зображення буває яскрава та тонова.

Яскрава контрастність є різницею між фізичною або видимою яскравістю окремих ділянок зображення. Взагалі, обчислення фізичної або видимої яскравості можна розглядати як конвертацію кольорового зображення в ахроматичні кольори. Тому яскравість контрастності – це порівняння двох ділянок зображення, наведених до ахроматичних кольорів.

Якщо проаналізувати RGB-гістограми, можна зробити висновок, що у контрастного зображення кількість темних і світлих пікселів має бути приблизно однаковою, різниця в їхній яскравості – значна, а основне місце зосередження пікселів – біля меж діапазону.

Тонова контрастність враховує як яскравісні переходи, а й переходи кольору. Так як при аналізі осіб використовуються чорно-білі зображення (у градаціях сірого), то використання тонової контрастності не має сенсу.

Хороший критерій оцінки яскравості контрастності – дисперсія яскравості пікселів зображення [6]:

$$\sigma^2 = \frac{1}{N} \sum_{p=1}^N (Y_p - Y)^2, \quad (1)$$

де  $Y$  – середнє значення яскравості контрастності всього зображення,  $Y_p$  – значення яскравості контрастності в точці  $p$ ,  $N$  – загальна кількість точок зображення.

Оцінка контрастності (1) може бути нормована шляхом обчислення відношення середньоквадратичного відхилення до максимально можливого значення яскравості:

$$C = \frac{2\sigma}{Y_{\max}}. \quad (2)$$

Значення  $C$  змінюється у діапазоні  $[0;1]$ . Значення 0 відповідає однотонному зображенню, значення 1 — максимально контрастному. Оптимальне значення контрастності (2) залежить від типу об'єкта, представленого на  $H_a$  (рис. 1.5) наведено приклади зображень та їх оцінка контрастності.



Рисунок 1.5 - Приклади оцінки контрастності зображень:

а)  $C = 0,42$ ; б)  $C = 0,28$ ; в)  $C = 0,57$ ; г)  $C = 0,52$ .

Як можна побачити на (рис. 1.5), контрастність зображення не може бути точним показником якості зображення, наприклад, зображення а) та г) розмиті та менш інформативні, але при цьому мають більш високу оцінку контрастності, ніж зображення б), через велику кількість точок максимальної та мінімальної яскравості.

#### 1.1.4. Оцінка різкості

Різкість найчастіше сприймається як якість відтворення на фото досить великих фрагментів. У цьому, умовно, її відмінність від чіткості зображення, що характеризує помітність дрібних деталей.

Підвищення різкості – один із найпоширеніших способів покращення фотографій. Зазвичай завжди потрібно тією чи іншою мірою посилювати різкість початкового зображення.

Об'єкти на зображенні відрізняються один від одного за яскравістю або оптичною щільністю, що є причиною їхньої помітності. Кордон між об'єктами завжди має той чи інший ступінь розмитості.

Ширина розмитості теоретично різкої межі між світлими (яскравими) та темними ділянками (об'єктами) зображення характеризує різкість зображення. Причому це притаманно будь-яких зображень, як кольорових, і чорно-білих.

Чіткість зображення – це параметр, який характеризує якість отриманого зображення та показує, наскільки чітко помітні дрібні об'єкти на зображенні та визначається мінімальною відстанню між ними, так, щоб ці об'єкти сприймалися як окремі.

Іншими словами, чіткість показує повноту відтворення дрібних об'єктів на зображенні та наскільки у них різкі контури.

Отже, чіткість характеризує ступінь розмиття кордонів навколо об'єктів В

Зображення високої чіткості має різкі, чітко видимі межі. Різкість може виступати ще одним критерієм інформативності вина-

У загальному випадку може бути визначена наявністю контурів контрастних переходів. Однозначного методу визначення різкості не існує, але є кілька алгоритмів, які дозволяють обчислити оцінку різкості з різних характеристик зображення.

Головна відмінність всіх запропонованих методів полягає у формулах, які визначають кількісну оцінку контрастності зображення. Загалом усі методи можна поділити на 3 групи:

– Методи, що обчислюють середнє значення диференціальних операторів зображення;

– Методи, що обчислюють ексцес амплітудного спектра зображення;

– Методи, що обчислюють відношення значущих і незначних точок амплітудного спектра зображення;

Перша група методів обчислює різницеву характеристику безпосередньо з вихідного зображення, наприклад, за допомогою оператора Лапласа або оператора Робертса, друга і третя групи обчислюють частотні характеристики зображення (перетворення Фур'є, або Wavelet-перетворення), і обчислює кількість контрастних переходів із спектру зображення .

Ідея методу, що використовує диференціальні оператори, полягає у обчисленні середнього значення яскравості від перетвореного оператором зображення. Такий метод чутливий до якості зображення, а також характеру зображених об'єктів, тобто він може дати високу оцінку різкості для поганого зображення зі складним об'єктом, і меншу – для хорошого зображення з малим числом контрастних переходів.

Оператор Лапласа – диференціальний оператор, сума других приватних похідних за координатами зображення, еквівалентний операціям градієнта та дивергенції:

$$Laplace(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

Оператор Робертса – пакунок зображення з ядрами

$$\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \text{ и } \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}.$$

Оператор Робертса порівняно з оператором Лапласа вимагає менше обчислень, точніше виділяє тонкі контури, але дуже чутливий до шуму (рис. 1.6).

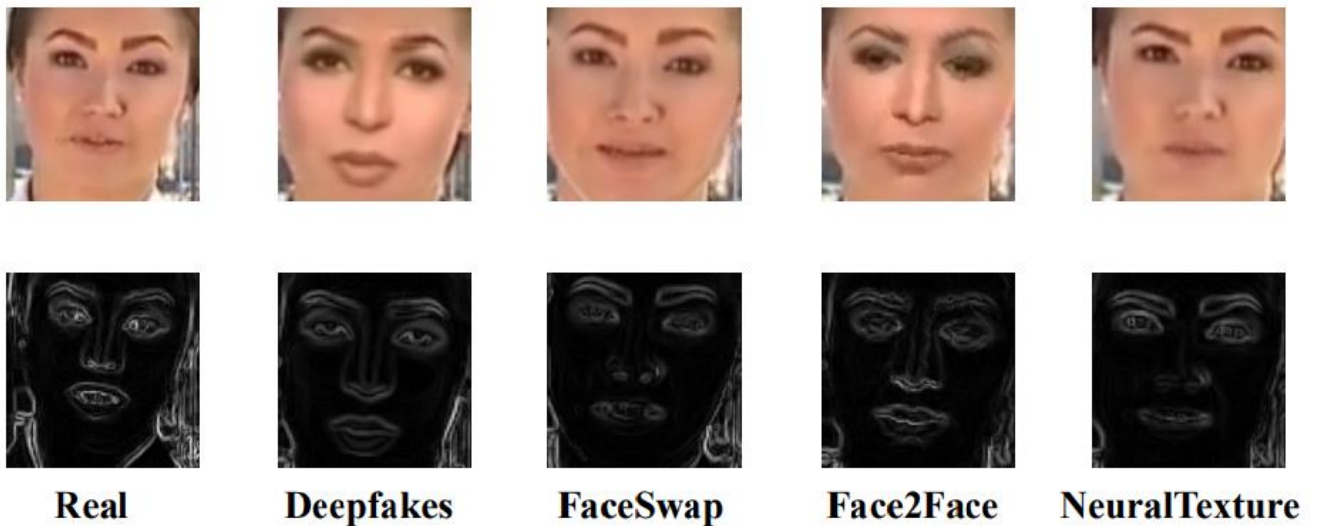


Рисунок 1.6 - а) вихідне зображення; б) виділення контурів оператором Лапласа; в) виділення контурів оператором Робертса [7]

В роботі [64] розглянуто метод оцінки різкості на основі перетворення Фур'є, суть якого полягає в обчисленні ексцесу зображення в частотній області:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4},$$

де  $\mu_2$  – другий центральний момент зображення,  $\mu_4$  – четвертий центральний момент,  $\sigma$  – стандартне відхилення.

В роботі [65] також використовується перетворення зображення в частотну область, але як критерій оцінки різкості пропонується відношення

$$FM = \frac{T_H}{M \cdot N},$$

де  $T_H$  – число значущих точок амплітудного спектра,  $M \cdot N$  – розмір зображення амплітудного спектра. Значні точки визначаються як точки, значення яких перевищує деяку частку максимуму амплітуди спектра.

Дослідження методів показало, що оцінка різкості з використанням оператора Лапласа дає помітний розкид результатів за схожих умов висвітлення та ракурсу, що пов'язано з відсутністю нормування результату та способом виділення області обличчя. Оцінка (3) показує стабільніші результати, але практично не реагує на фактичне зниження різкості, отримане розмиттям за Гаусом. Оцінка (4) найбільш точно відповідає значній задачі виділення пошуку різких зображень, вона коректно реагує на зміни різкості при русі та відсутність фокусу, і крім цього має строге нормування, що дозволяє за заданим рівнем відсіювати зображення незадовільної якості.

На (рис. 1.7) наведено приклади зображень, їх амплітудних спектрів.

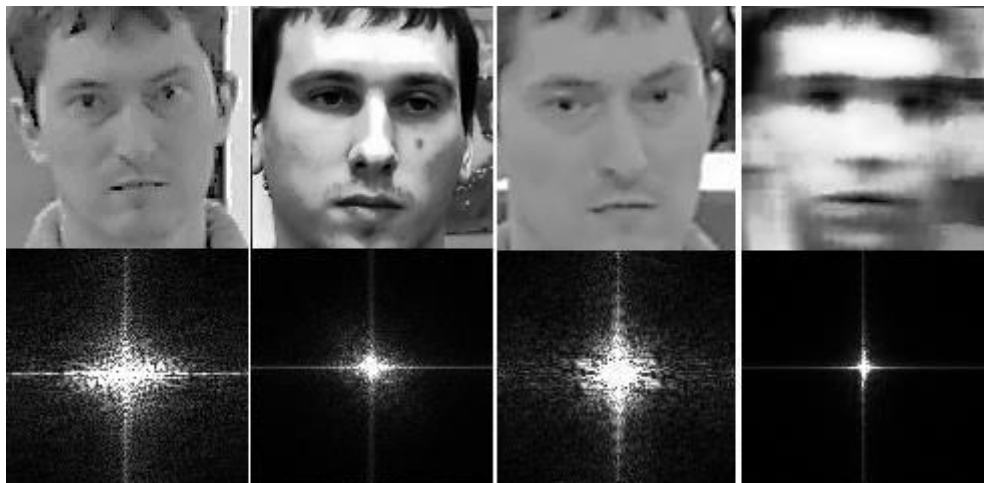


Рисунок 1.7 - Приклади оцінки різкості зображень: а)  $\gamma^2=0.0035$ , FM=0.1301; б)  $\gamma^2=0.0025$ , FM=0.0388; в)  $\gamma^2=0.004$ , FM=0.1203; г)  $\gamma^2=0.0062$ , FM=0.0111, [10]

Незважаючи на переваги оцінки (4), вона частково схильна до тих же проблем, що й оцінка з використанням Лапласіана. Основна причина цього в тому, що різкість зображення оцінюється кількісно, і тому пов'язана із загальним числом контрастних переходів, а це залежить від об'єкта на зображенні. Виходячи з цього, можна говорити лише про відносну різкість зображення, яка оцінюється для одного об'єкта в треку (серії зображень). При цьому можна припустити, що метод дасть більш точну оцінку, якщо при оцінці використовуються зображення тих ділянок

особи, які в переважній більшості випадків мають схожі показники різкості для всіх людей, наприклад, очі, ніс, губи.

#### 1.1.5. Компенсація освітлення

У разі завдання підготовки зображення до розпізнавання облич проблема різних умов зйомки може бути вирішена за допомогою об'єднання трекінгу особи у відеопотоці та обчислення оцінок для серії зображень однієї й тієї ж особи, з наступним виділенням кадру з найкращими значеннями. Підвищити застосовність методів (1)-(4) також може компенсація зображення особи за умовами освітлення та ракурсу.

Для нормування вихідного зображення за умовами освітлення існує кілька методів, наприклад, DoG-фільтрація [6] і частотна фільтрація смуговим фільтром [67], що прибирає низькі та високі частоти. Обидва методи дозволяють досить ефективно вирівнювати зображення при дисбалансі яскравості освітлення, наприклад у випадку, коли є бічна засвітка обличчя.

DoG виявляє краї шляхом застосування розмиття зображення за Гаусом (Gaussian blur) із зазначеним значенням стандартного відхилення. Результати зображення є розмитою версією вихідного зображення. Далі застосовується інше розмиття з меншим значенням стандартного відхилення, що дає різкіше зображення, ніж у попередньому випадку. Фінальне зображення обчислюється шляхом заміни кожного пікселя різницею між двома розмитими зображеннями.

Для нормалізації яскравості зображення з цим методом часто використовується гамма-корекція.

Частотний смуговий фільтр при нормалізації освітлення пригнічує високочастотні шуми та згладжує нерівномірне освітлення зображень, що кодується у зображенні низькими частотами. Приклад смугового фільтра частотної області представлений на (рис. 1.8).

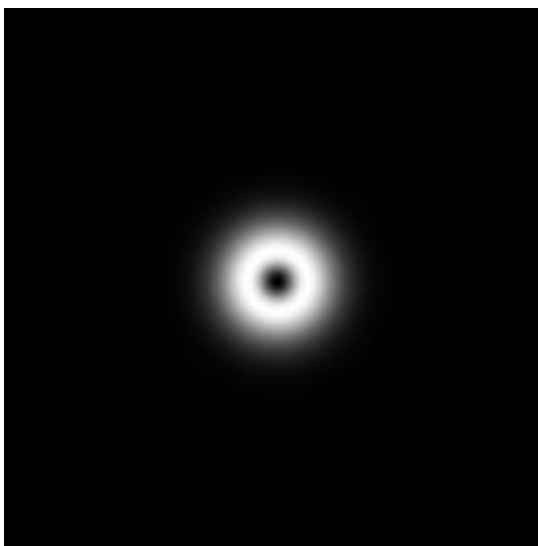


Рисунок 1.8 - Приклад амплітудного спектру смугового фільтра для зображення [10]

Порівняння цих методів нормування висвітлення представлено на (рис. 1.9) Обидва методи підвищують помітність основних рис особи, але не усувають повністю вплив нерівномірного освітлення.



Рисунок 1.9 - а) вихідне зображення; б) DoG-фільтрація; в) фільтрація смуговим частотним фільтром [10]

Також для підвищення контрасту зображення застосовується адаптивне вирівнювання гистограми з обмеженням (CLAHE) [8,9]. Спочатку зображення розбивається прямокутні області (блоки). У кожному блоці обчислюється гистограма розподілу яскравості. Щоб уникнути перенасичення монотонних

ділянок зображення, обмежується нахил функції розподілу яскравості шляхом введення обмеження на максимальне значення

частоти яскравості. Нехай  $M$  – число рядків у блоці,  $N$  – число стовпців у блоці,  $\alpha$

– параметр, який визначає обмеження. Граничне значення гистограми обчислюється відповідно до формули:

$$hist_{lim} = hist_{min} + \alpha(MN - hist_{min}),$$

де  $hist_{min} = MN / (L + 1)$ ,  $L$  – максимальне значення яскравості.

Після обмеження кожної області формується своя функція перетворення яскравості вирівнювання гистограми розподілу яскравості. Для вхідного зображення обчислюється значення яскравостей вихідного зображення відповідно до формули:

$$LUT(k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n},$$

де  $LUT(k)$  – значення яскравості вихідного зображення, що відповідає яскравості  $r_k$  вхідного зображення,  $k \in [0, L]$ ;  $r_j$  – яскравість елемента вхідного зображення,  $n_j$  – кількість елементів вхідного зображення зі значенням  $r_j$ ,  $n$  – загальна кількість елементів зображення;  $p_r(r_j)$  – ймовірність появи елемента вхідного зображення з яскравістю  $r_j$ .

Таке перетворення дозволяє рівномірніше розподілити рівні яскравості за частотою, в основі перетворення лежить припущення, що в контрастних чорно-білих зображеннях значення яскравості рівномірно розподілені.

Щоб зруйнувати блочну структуру отриманого зображення, вихідне зображення формується шляхом білінійної інтерполяції за чотирма найближчими LUT. На (рис. 1.10) наведено приклад обробки зображення за допомогою методу CLAHE.



Рисунок 1.10 – Зліва – вихідне зображення, праворуч – зображення, оброблене методом CLAHE [12]

#### 1.1.6. Контроль ракурсу

Контроль положення особи щодо оптичної осі камери є важливою проблемою у зв'язку з чутливістю алгоритмів розпізнавання до ракурсу. Найбільш відомим методом оцінки ракурсу є POSIT (Pose from Orthography and Scaling with Iterations), він відрізняється високою продуктивністю та швидкою збіжністю.

Алгоритм POSIT обчислює матрицю повороту та вектор паралельного перенесення так, щоб помилка між проекцією характерних точок тривимірної моделі на площину зображення та відповідними ними розміченими точками на зображенні була мінімальною. Метод POSIT комбінує два алгоритми. Перший, POS (Pose from Orthography and Scaling), апроксимує перспективну проекцію за допомогою масштабованої ортографічної проекції та обчислює матрицю повороту та вектор паралельного перенесення об'єкта за допомогою рішення СЛАУ. Другий, POSIT (POS with Iterations), у циклі використовує результати роботи POS для знаходження більш точної масштабованої ортографічної проекції характерних точок, а потім застосовує POS до отриманих точок, а не до вихідних. У результаті алгоритм POSIT сходить до точного рішення за кілька ітерацій [7].

Проблема використання алгоритму практично у тому, що для його роботи необхідно зображення, на якому мають бути відзначені характерні точки особи,

що у випадку не є автоматизованою процедурою. Існуючі методи автоматичного розміщення характерних точок допускають серйозні похибки, але на малих кутах повороту дозволяють оцінити ракурс особи з достатньою точністю [7].

При правильному розміщенні характерних точок обличчя алгоритм оцінки курсу обличчя дозволяє отримати коефіцієнти для завдання корекції ракурсу зображень осіб та обчислити нормоване зображення.

Існує кілька методів автоматичного розташування точок, всі вони засновані на активних моделях форм (Active Shape Models, ASM).

Активні моделі форми - це статистичні моделі форми об'єктів, які можуть неодноразово піддаватися деформаціям для найкращої відповідності об'єкту, представленому на зображенні. Дані моделі були запропоновані Крісом Тейлором і Тімом Кутесом в 1995 [7]. Модель розподілу точок обмежує форму об'єкта, тому зміна доступна лише в межах, заданих розміченими прикладами з навчальної вибірки. Форма об'єкта представлена безліччю точок, що контролюються через форму моделі. Метою алгоритму ASM є зіставлення моделі з новим зображенням. В основі алгоритму лежать дві дії, що чергуються:

- Пошук кращої позиції для цієї точки на зображенні навколо кожної точки;
- Оновлення параметрів ASM, які найкраще відповідають новим знайденим позиціям.

Для пошуку кращої позиції в кожній точці можна визначати чіткі межі на зображенні, а можна поєднати очікування цієї точки зі статистичною моделлю. При обчисленні кращої позиції для кожного орієнтира точки оригінальний метод передбачає використання відстань Махаланобіса.

Метод активних моделей форм знаходить широке застосування у аналізі механічних вузлів, зображень осіб, а також медичних знімків (як двовимірних, так і тривимірних).

#### Розвиток активних моделей форм

Деякий час математичний апарат активних моделей форм активно розвивався. На даний момент виділяються два основних підходи до формування таких моделей:

класичний, запропонований Кутесом і Тейлором спочатку, і підхід, заснований на зворотній композиції (запропонований Бейкером і Метьюсом у 2003 році [7]).

Ці два підходи мають багато спільного. Активні моделі форм використовують два типи параметрів: пов'язані з формою об'єкта (параметри форми) та пов'язані з текстурою або статистичною моделлю зображення (параметри зовнішнього вигляду). Перед застосуванням модель навчається на безлічі зображень із заздалегідь нанесеною розміткою крапок. Попередня розмітка зображень найчастіше здійснюється або повністю вручну, або в напівавтоматичному режимі, коли спочатку за допомогою іншого простішого алгоритму знаходиться приблизне розташування міток, а потім воно уточнюється експертом. Кожна мітка має свій порядковий номер і визначає характерну точку, яку модель повинна знаходити при адаптації до нового зображення. Приклад такої розмітки показаний (рис. 1.11).

На цьому зображенні особи поставлено 68 міток (особливих точок обличчя), які утворюють форму ASM. Така форма визначає зовнішній контур (овал) обличчя, контури очей, рота, брів та носа. Характер розмітки дозволяє визначити різні параметри обличчя на його зображенні, які можуть бути використані надалі, у тому числі й для подальшої обробки за допомогою інших алгоритмів. Наприклад, ASM добре сумісний з алгоритмами ідентифікації особистості, визначення емоційного стану суб'єкта і навіть аудіовізуального розпізнавання мови.

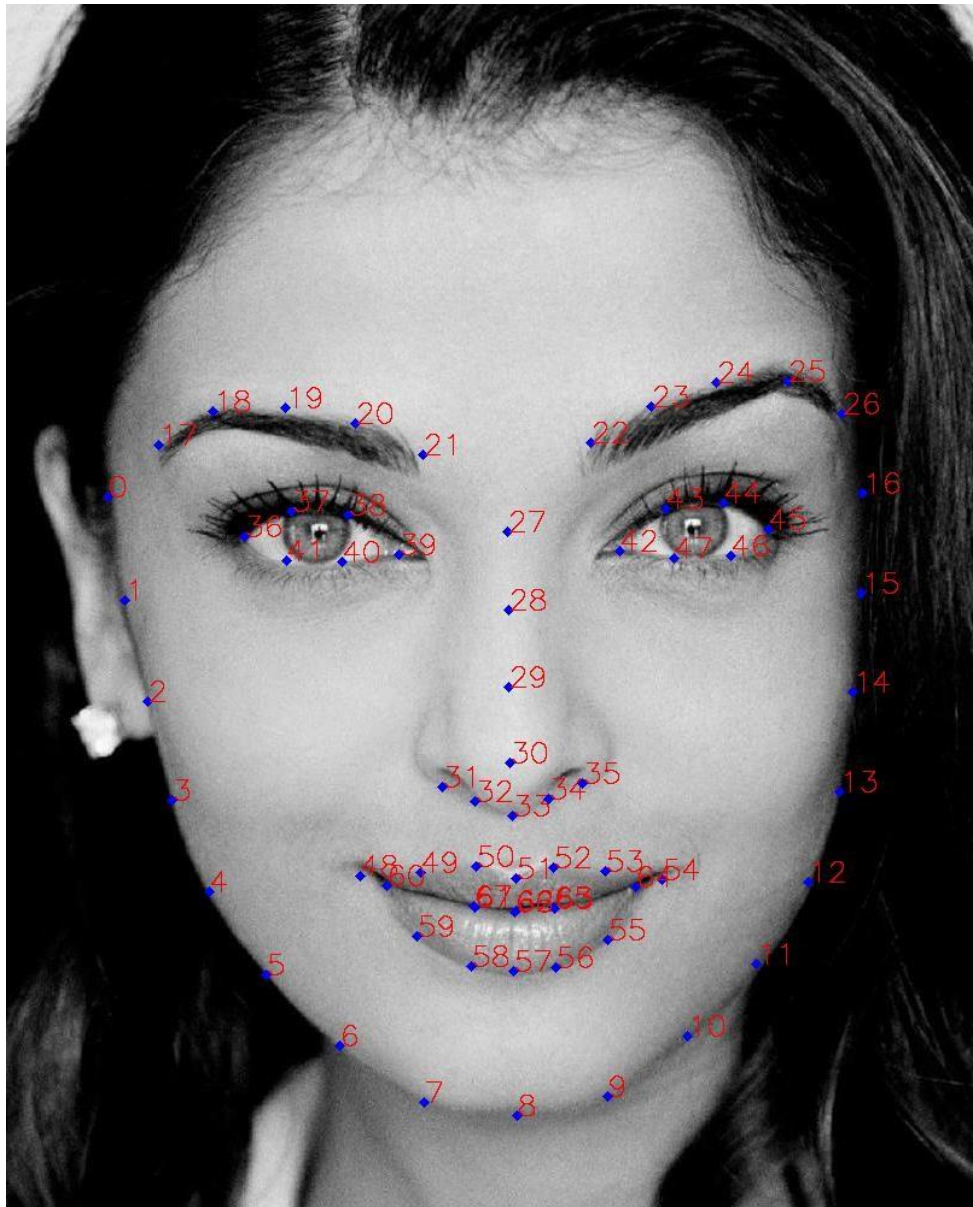


Рисунок 1.11 - Приклад автоматичної розмітки обличчя за допомогою алгоритму ASM

Процес навчання активних моделей форм вимагає, перш за все, компенсувати відмінності в масштабі, зміщенні та нахилі, тому він починається з нормалізації становища всіх форм. Основний метод, застосований для нормалізації становища форм - узагальнений Прокрустів аналіз [74]. На малюнку 12 показано, як виглядає безліч особливих точок до і після нормалізації (відповідно до [75]).

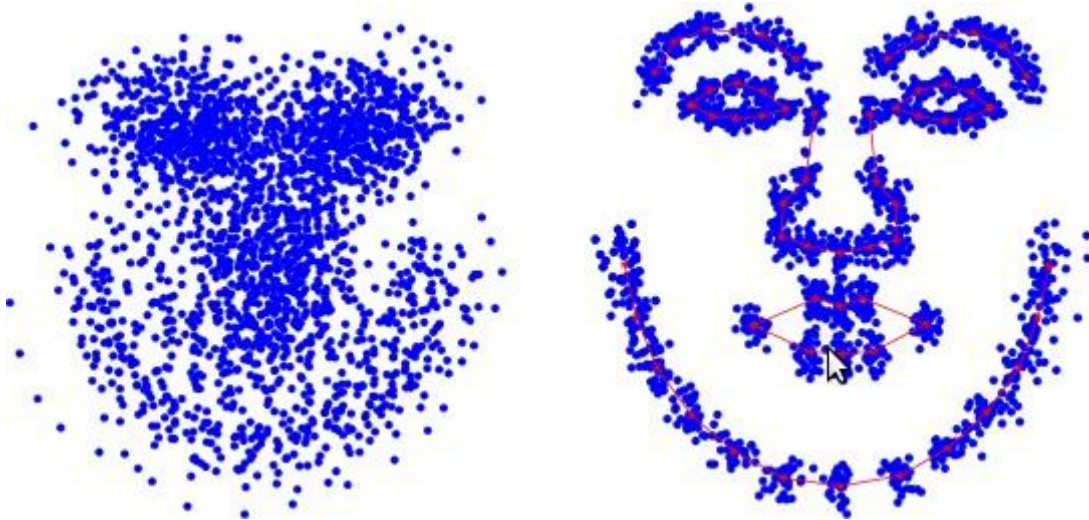


Рисунок 12. Безліч особливих точок до та після нормалізації

Після нормування всіх форм з точок, що їх складають, проводиться формування матриці  $S = [s_1, s_2, \dots, s_M]$ , где  $s_M = [x_1^m, \dots, x, y_1^m, \dots, y_N^m]^T$ .

Вираз для сконструйованої форми можна отримати шляхом виділення головних компонентів матриці  $S$ :

$$s = s_0 + \Phi_s b_s,$$

де  $s_0$  - базова форма, тобто форма, усереднена за кожною з реалізацій навчальної вибірки,  $\Phi_s$  - матриця основних векторів,  $b_s$  - параметри форми. Поданий вище вираз показує, що форма  $s$  може бути виражена у вигляді суми базової форми  $s_0$  з лінійною комбінацією, що містяться в матриці  $\Phi$  власних форм. Для припасування форми під реальне зображення можна отримувати її деформації різного роду шляхом зміни вектора параметрів  $b_s$ . На (рис. 1,13) наведено приклад подібної форми [7]. Напрямки головних компонентів показані червоними та синіми стрілками.

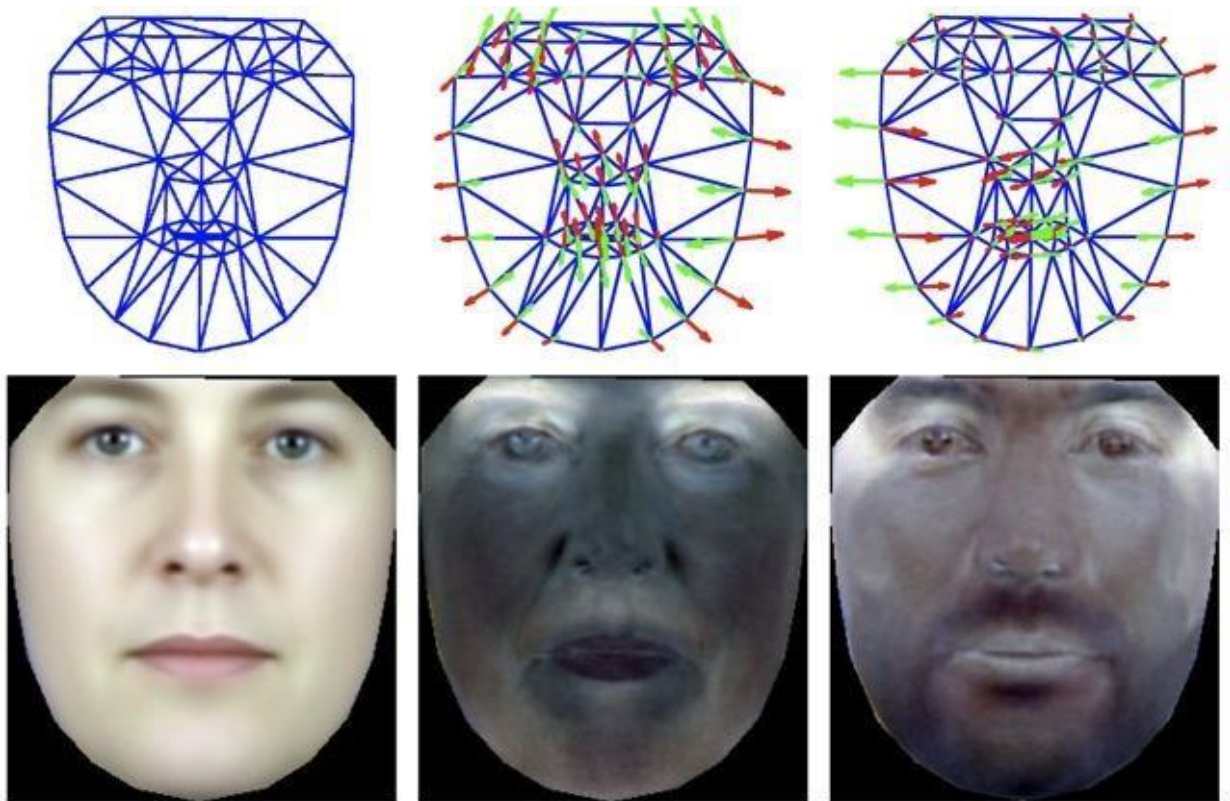


Рисунок 1.13 - Набуття нової форми особи шляхом модифікації базової форми [10]

Також усі моделі активних форм поділяються на моделі з жорсткою деформацією та з не жорсткою деформацією. До моделей з жорсткою деформацією допустимо застосовувати лише афінні перетворення (зсув, поворот, масштабування), а з моделями з не жорсткою деформацією допустимо використовувати й інші види деформацій. На практиці найчастіше використовується деяка комбінація жорстких та не жорстких деформацій. При такій комбінаційній моделі як параметри форми додаються також параметри розташування (зміщення, кут повороту, масштаб або коефіцієнти афінного перетворення).

Після обчислення компонентів форми виконується процедура навчання компонентів зовнішнього вигляду (матриця головних компонент і базова форма). Процес навчання поділяється на три етапи:

- Витяг із навчальних зображень тих текстур, які відповідають базовій формі якнайкраще. Для цього виконується тріангуляція особливих точок базової

форми та форми, що складається з особливих точок

навчального зображення. Далі отримані в результаті тріангуляції регіони навчального зображення відображаються у відповідні регіони текстури, що формується, за допомогою шматкової інтерполяції.

## 1.2 Аналіз

Існуючі методи трекінгу об'єктів, що рухаються, що оперують змінами в зображенні в часі, вимагають багато ресурсів і залежать від багатьох параметрів, таких як фон (рухливий або нерухомий), характер руху [9,10,11,12].

У зв'язку з тим, що в процесі розпізнавання особи людини першим етапом завжди є пошук особи на зображенні (наприклад, методом Віюлі-Джонса), доцільно скористатися цим у процесі трекінгу. Знаючи координати зафіксованої особи на зображенні, стає можливим оперувати не масивом пікселів, а лише його відносною зміною положення у послідовності кадрів. Сам метод детектування області особи на зображенні досить ресурсоємний і може займати більше половини процесорного часу при обробці відеопотоку високої якості на слабких ЕОМ.

У такій постановці завдання може бути запропонований новий ефективніший алгоритм трекінгу особи, який використовує не тільки інформацію про положення особи на зображенні, а й швидко попередню обробку кадрів відеопотоку з метою зменшити область, яка підходить для використання детектора.

### Порівняння зображень

Методи порівняння зображень, засновані на SURF-дескрипторах або на обчисленні хеш-функцій, мають хорошу швидкодію, але погано підходять до завдання виділення слабких відмінностей на зображення, в той час як метод обчислення різниці двох зображень дуже чутливий до шумів, що є невід'ємною властивістю зображень відеопотоку.

Єдиним методом, який підходить для виконання поставленого завдання, є метод обчислення коефіцієнта кореляції, але його швидкодія сильно залежить від

вихідного розміру зображень, тому його можна застосовувати лише до області зображення зі знайденою особою, тобто необхідно порівняти лише треки осіб.

#### Оцінка контрастності та різкості

Для оцінки контрастності цілком підходить оцінка яскравості контрастності, а для оцінки різкості – оцінка на основі аналізу амплітудного спектру зображення. Однак обидві ці оцінки дають лише наближене уявлення про якість зображення, вони можуть бути використані для виключення з обробки свідомо поганих зображень або для оцінки застосування методів відновлення контрастності та компенсації освітлення.

#### Компенсація освітлення

Кожен із методів компенсації освітлення має свої переваги, і хоча візуально кращу картинку дає метод CLANE, для різних методів ідентифікації особистості різна обробка знімків може давати зовсім інші результати. Щоб зробити висновки про кінцеву ефективність ности методів, необхідно провести тестування кожного з методів у сукупності з вибраним методом розпізнавання осіб.

#### Контроль ракурсу

Для завдання розпізнавання особи може бути застосований як підхід відновлення фронтального ракурсу особи (у цьому випадку немає суворих вимог до початкового набору зображень особи в базі) або більш загальний підхід одержання множини зображень особи з різними кутами повороту щодо оптичної осі камери.

У першому підході достатньо використовувати алгоритм автоматичного розміщення точок обличчя ASM та алгоритм POSIT для оцінки кута повороту. Далі застосовується алгоритм побудови невідомої тривимірної моделі обличчя та виконується поворот.

У другому підході необхідно рухатися від зворотного завдання, тобто виконувати розстановку точок на відомому фронтальному обличчі, і для нього будувати тривимірну модель і формувати зображення поворотів. У цьому випадку вихідне зображення має бути фронтальним або близьким до нього, але з процесу побудови моделі виключається потенційно ненадійний алгоритм оцінки ракурсу

POSIT, що залежить від точності розміщення точок та їх взаємного розташування, якості зображення тощо [10].

Перший підхід має ще один недолік, а саме погану застосовність до існуючих алгоритмів ідентифікації особистості зображення обличчя. Найбільш ефективно такі системи працюють в умовах, коли є набір зображень у різних умовах зйомки, за якими будуються біовектори, що відповідають конкретній людині. Другий підхід дозволяє збільшити кількість зображень у кілька разів, підвищивши точність ідентифікації.

Виходячи з цих міркувань, поворот фронтального зображення для генерації набору біовекторів є кращим.

Для формування тривимірної моделі оптимальним є використання узагальненої або деформованої моделі особи, оскільки не накладає додаткових вимог до обладнання та процесу підготовки системи розпізнавання та надає при цьому прийнятну точність.

Для вибору між узагальненою та деформованою моделями необхідно провести додаткове тестування якості розпізнавання при їх використанні.

Висновки за розділом: розглянуто різні методи оцінки та обробки зображень, що дозволяють підвищити розрізнення об'єктів, виділити ослідовність кадрів, порівняти два зображення.

Вибрано найбільш оптимальні методи оцінки контрастності та різкості, методи компенсації освітлення та порівняння зображень.

Проаналізовано існуючі методи трекінгу об'єктів у відеопотоці, сформульовано умови та вимоги до трекінгу осіб у відеопотоці.

Зроблено критичний аналіз методів та алгоритмів розміщення особливих точок обличчя, методу оцінки кута повороту POSIT, а також технології формування тривимірних моделей обличчя, на основі якого обґрунтовано розробку алгоритму побудови тривимірної моделі обличчя з одного фронтального растрового зображення та створення набору зображень на основі побудованої моделі.

## 2 ПІДГОТОВКА ЗОБРАЖЕНЬ ОСІБ У ВІДЕОПОТОЦІ ДО РОЗПІЗНАВАННЯ

### 2.1 Вибір методів, що використовуються у складі технології

Проведений у розділі 1 огляд та аналіз існуючих методів обробки зображень показав, що для низки завдань попередньої обробки зображень осіб стосовно розпізнавання особистості не існує ефективних рішень, спрямованих на реалізацію програмно-апаратних систем реального часу.

Алгоритми трекінгу, оцінки та компенсації ракурсу зображень осіб рідко розглядаються як компоненти системи розпізнавання, що ускладнює їх адаптацію до розв'язання задачі. Розробка методів, оптимальних для автоматичного застосування в системах реального часу, є критично важливою для формування комплексної технології попередньої обробки зображень осіб у відеопотоці.

#### 2.1.1. Трекінг

Для розробки алгоритму трекінгу за основу взято найпростіший метод, що оперує лише даними про становище виявлених осіб у кадрі.

Відеопотік можна подати у вигляді послідовності кадрів (зображень)

$$V_s = [f_1, f_2 \dots f_N],$$

де  $N$  – загальна кількість кадрів,  $f_i$  – поточний кадр. Будь-яка особа, знайдена у кадрі, описується як набір

$$F_{c_{ij}} = \{I_{f_{ij}}, R_{f_{ij}}\},$$

де  $I_{f_{ij}}$  – зображення обличчя,  $R_{f_{ij}}$  – його становище у кадрі,  $i$  – номер кадру,  $j$  – номер особи у кадрі.

При цьому  $R_{f_{ij}}$  можна подати як набір з двох точок

$$R_{f_{ij}} = \{p_1=(x_1; y_1), p_2=(x_2; y_2)\},$$

де  $p_1$  – верхня ліва точка прямокутної області,  $p_2$  – нижня права точка,  $x$  та  $y$  – їх координати.

Тоді трек обличчя  $k$  у відеопотоці – це послідовність осіб  $F_{c_{ik}}$ , що відповідає одній людині:

$$Tr_k = [F_{c_{1k}}, F_{c_{2k}} \dots F_{c_{Nk}}].$$

Основне завдання у трекінгу полягає у зіставленні положень цільового об'єкта на послідовності кадрів. Таким чином, системи трекінгу зазвичай використовують модель руху, яка описує, як може змінюватися зображення цільового об'єкта при різних його рухах [2]. У разі спостереження за становищем особи передбачається, що об'єкти у наступному кадрі змінюють свою позицію незначно. У зв'язку з цим запропоновано такий метод трекінгу: знайдене зображення особи належить треку  $Tr_k$ , якщо точка-центр прямокутника належить області прямокутника  $R_f$  особи  $F_{c_{i-1,j}}$ , знайденої на попередньому кадрі та що належить треку  $Tr_k$  (рис. 2.1).

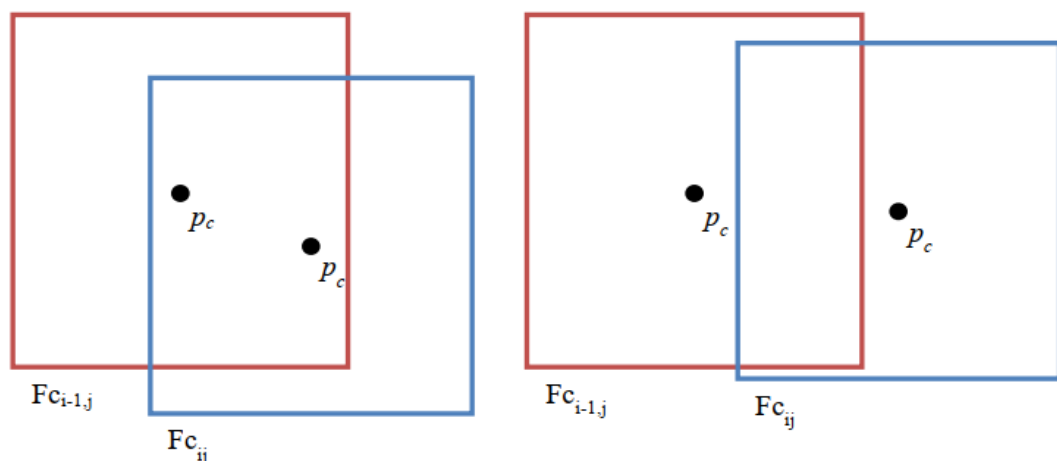


Рисунок 2.1 - Функціонування найпростішого методу трекінгу: зліва – області виявлених осіб належать одному треку, праворуч – області осіб будуть вважатися такими, що належать різним людям [11]

Розглянутий метод не може гарантувати абсолютну точність розпізнавання треку, але дуже простий з обчислювальної точки зору і часто досить ефективний для вирішення завдання об'єднання зображень осіб на відео у послідовності при уточненні результуючої оцінки належності зображень обличчя людині у системах розпізнавання. Приклад треку особи у відеопотоці, отриманого таким методом, наведено на (рис. 2.2).

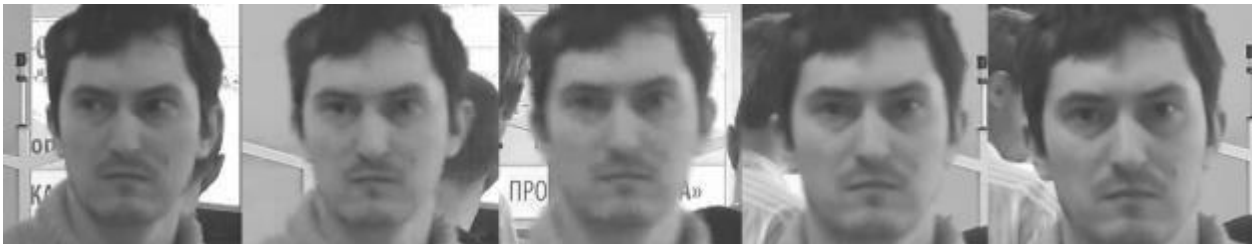


Рисунок 2.2 - Приклад треку обличчя у відеопотоці [15]

Даний метод може бути покращений у двох напрямках – у напрямку підвищення якості стеження, тобто усунення розривів у русі через неспрацьовування детектора, та у напрямку підвищення швидкості обчислень.

Для компенсації нестачі точності цього методу існує низка рішень, спрямованих на утримання фокусу інтересу на областях кадру, в яких детектору не вдалося виявити особу, з метою пов'язати його з попереднім на підставі деяких відмінних рис, таких як взаємне розташування точок, їх зміщення (наприклад, методи SURF, mean shift, KLT) [10]. Проте, з погляду застосування у системах розпізнавання, подібний трекінг необхідний який завжди. Для завдання розпізнавання аналіз рис особи за зображенням, яке спотворене в мірі, що не допускає його виявлення детектором, не має практичного сенсу: це може призвести до зниження точності, так і швидкодії системи розпізнавання. Варто зазначити, що у разі застосування алгоритмів розпізнавання, заснованих на аналізі 3D образів осіб або приватних рис особи, не прив'язаних до цілісного об'єкта, методи покращення якості трекінгу мають більше значення через більш високу ймовірність розпізнати обличчя за зображенням, що містить сильні перешкоди. .

У розглянутій задачі, пов'язаній з підвищенням ефективності та швидкості обробки відео послідовностей системами розпізнавання особи, більше значення має швидкодія трекінгу. У контексті функціонування системи розпізнавання реального часу необхідність безперервного сканування всієї області зображення детектором осіб нівелює переваги трекінгу осіб існуючими методами. Для того, щоб знизити обчислювальне навантаження, яке може займати більше половини обчислювальних ресурсів залежно від архітектури ЕОМ та параметрів ПЗ, має сенс проводити попередній аналіз кадрів відео для виділення областей, що потенційно містять особи. У цьому плані найперспективнішими є швидкі методи віднімання фону та виявлення руху, що складаються з найпростіших операцій обчислення різниці між двома кадрами відеопотоку та є легкими з обчислювальної точки зору [10].

Пропонується алгоритм прискорення базового методу трекінгу осіб, що ґрунтується на положенні осіб у кадрі, що складається з наступних дій:

– Обчислення різниці між поточним кадром та попереднім, що визначає наявність руху в потоці відео:

–

$$f_M = f_i - f_{i-1},$$

де  $f_i$  – поточний кадр відеопотоку,  $f_M$  – матриця, кожен з елементів якої визначає різницю відповідних пікселів  $f_i$  та  $f_{i-1}$  кадрів відеопотоку.

– Обчислення різниці між поточним кадром і першим, що визначає наявність у потоці відео об'єктів, що не є тлом:

–

$$f_T = f_i - f_0,$$

де  $f_i$  – поточний кадр відеопотоку,  $f_T$  – матриця, кожен з елементів якої визначає різницю відповідних пікселів  $f_i$  та  $f_0$  кадрів відеопотоку.

– Обчислення бінарних зображень  $f_T$  і  $f_M$  з порогом, що залежить від якості відеопотоку та шумів.

– Оцінка розподілу білих пікселів бінарного зображення для  $fT$  і  $fM$  за допомогою ковзного вікна зі стороною, що дорівнює половині мінімального вікна детектора осіб. Ті області зображення, в яких сума білих пікселів перевищує обраний поріг, відзначаються як відповідні для подальшого аналізу. Результуючі області для зображень  $fT$  і  $fM$  об'єднуються, що допомагає включити в загальну область як об'єкти, що рухаються, так і об'єкти, що знаходяться в кадрі нерухомо.

– Детектор осіб застосовується у вибраній області зображення, дозволяючи виділити серед рухомих об'єктів обличчя людей.

Такий алгоритм дозволить уникнути обчислення вейвлетів Хаара для всього кадру, значно знижуючи обчислювальну складність трекінгу осіб за координатами області зображення обличчя.

Слід зазначити, що запропонований алгоритм здатний ефективно працювати лише за обмежених умов: за наявності нерухомого, переважно статичного фону, тобто з відеопотоками, отриманими від камер відеоспостереження на стаціонарних об'єктах. У разі великої динаміки сцен такий спосіб прискорення обробки кадрів виявиться неефективним і навіть призведе до уповільнення алгоритму пошуку облич на зображенні. Повільні або незначні зміни фону можуть бути компенсовані регулярною заміною опорного кадру в моменти, коли сумарна різниця сусідніх кадрів у цілому менша за деякий критичний поріг [10].

Для відеопотоку, аналізованого в кожному колірному каналі окремо, є можливість ще більше підвищити точність виділення областей, що призведе до часткового зниження загальної продуктивності трекінгу.

У загальному випадку виділена область з об'єктами буде меншою за кадр повністю, що веде до зниження кількості операцій з детектором осіб Віоли-Джонса. В ідеальному випадку, коли в кадрі немає руху та об'єктів, детектор зовсім не виконує своїх обчислень, що вивільняє додаткові ресурси на обробку інших кадрів відеопотоку, що знаходяться в буфері вводу-виводу.

### 2.1.2. Порівняння зображень

З трьох розглянутих підходів до порівняння зображень для вирішення поставленої задачі найбільш підходящим є метод обчислення коефіцієнта кореляції. За допомогою цього методу можна виключати з треку Trk зображення з високим рівнем кореляції. У загальному випадку кореляція двох різномасштабних зображень може бути обчислена за такою формулою:

$$R(x, y) = \frac{\sum_{x', y'} (I_{ik}(x', y') \cdot I_{i+1, k}(x + x', y + y'))}{\sqrt{\sum_{x', y'} I_{ik}(x', y')^2 \cdot \sum_{x', y'} I_{i+1, k}(x + x', y + y')^2}}$$

Де  $I_{ik}$  - вхідне зображення,  $I_{i+1, k}$  - зображення для порівняння. Оцінка, одержувана за формулою, лежить у діапазоні від 0 до 1, де 0 відповідає максимально різним зображенням, а 1 максимально схожим.

Застосовуючи операцію обчислення  $R(x, y)$  за формулою (5) для зразка та ділянки зображення, можна отримати зображення в градаціях сірого, в якому найбільш яскрава пляма відповідає найбільш схожій на зразок ділянці зображення (рис. 2.3). Для двох зображень однакового розміру матриця  $R(x, y)$  складатиметься лише з одного елемента, що також лежить у діапазоні  $[0; 1]$  і характеризує ступінь схожості цих зображень.



Рисунок 2.3 - Результат обчислення коефіцієнта кореляції [12]

Цей метод досить коректно відображає ступінь схожості двох зображень, але погано застосовується для оцінки кадру цілком, так як вимогливий до обчислювальних ресурсів. Однак його можна застосовувати для пошуку серії осіб на зображенні, що належать одній людині, і вибору з цієї серії зображень, що тільки відрізняються один від одного.

### 2.1.3. Методи оцінки контрастності та різкості

Для оцінки контрастності застосуємо метод обчислення нормованої контрастності яскравості зображення. Застосування цієї оцінки для кадру повністю або області обличчя на зображенні може дати непередбачуваний результат обчислень. Як оптимальний варіант для обчислення можна використовувати лише ті точки зображення, які з високою ймовірністю належать області обличчя, тобто точки, що лежать усередині фігури, що утворюється

контуром овалу обличчя та брів, отриманим однією з модифікацій алгоритму ASM.

Найпростіший алгоритм, що дозволяє обчислити такі точки, полягає в наступному:

- Обчислення особливих точок особи;
- Збільшення яскравості кожної точки зображення на одиницю, таким чином, на зображенні не залишається нульових значень яскравості;
- Розбиття області обличчя на трикутники, утворені найближчими особливими точками, та додавання зображень трикутників із зображенням нульової яскравості.

В результаті цього вийде зображення, де всі точки, що не належать особі, мають нульову яскравість, і за цим критерієм можна виключати їх з обчислення оцінки контрастності зображення.

Аналогічно цей підхід можна застосувати і для оцінки різкості, але навіть з огляду на це кількісна оцінка різкості зображення скрутна. Для обчислення відносної оцінки різкості можна скористатися формулою (4).

#### 2.1.4. Метод корекції освітлення

Серед методів корекції освітлення неможливо вибрати оптимальний без аналізу того, як кожен із методів впливає на кінцевий результат розпізнавання. Тому для алгоритму підготовки зображень вирішено протестувати всі три методи та вибрати той, з використанням якого ефективність розпізнавання буде максимальною.

#### 2.1.5 Вибір засобів розробки

До програмної реалізації алгоритмів, що розробляються, пред'явлено кілька вимог:

- Реалізація у вигляді програми, доступної для використання в

операційній системі сімейства Windows;

- Видача результатів роботи у форматі, придатному для подальшого використання у процесі розпізнавання;

- Можливість вибору методів обробки зображень в інтерфейсі програми.

Як основну мову програмування було вирішено використовувати C++ як найбільш поширену і реалізацію компіляторів для більшості популярних операційних систем. У процесі розробки використовувалося середовище розробки MS Visual Studio 2013 разом з Qt (кроссплатформний інструментарій розробки ПЗ мовою програмування C++), а також бібліотека OpenCV для роботи з графічними об'єктами та Dlib для пошуку особливих точок обличчя на зображеннях.

Qt дозволяє запускати програми в більшості сучасних операційних систем шляхом простої їх компіляції для кожної ОС без зміни вихідного коду. Включає всі основні класи, які можуть знадобитися при розробці прикладного програмного забезпечення, починаючи від елементів графічного інтерфейсу і закінчуючи класами для роботи з мережею, базами даних і XML. Qt є повністю об'єктно-орієнтованим, що легко розширюється і підтримує техніку компонентного програмування.

Відмінна риса Qt від інших бібліотек – використання Meta Object Compiler (MOC) – попередньої системи обробки вихідного коду. MOC дозволяє багато разів збільшити потужність бібліотек, вводячи такі поняття, як слоти та сигнали. Крім того, це дозволяє зробити код лаконічнішим.

Qt комплектується візуальним середовищем розробки графічного інтерфейсу Qt Designer, що дозволяє створювати діалоги та форми в режимі WYSIWYG. У поставці Qt є Qt Linguist — графічна утиліта, що дозволяє спростити локалізацію та переклад програми на багато мов; і «Qt Assistant» — довідкова система Qt, що спрощує роботу з документацією по бібліотеці, а також дозволяє створювати крос-платформну довідку для програмного забезпечення, що розробляється на основі Qt. У комплект Qt включено середовище розробки

«Qt Creator», яка включає редактор коду, довідку, графічні засоби «Qt Designer» і можливість налагодження додатків. Qt Creator може використовувати GCC або Microsoft VC++ як компілятор і GDB як відладчик. Для Windows версій бібліотека комплектується компілятором, заголовками та об'єктними файлами MinGW.

Qt поширюється на 3 ліцензії: Qt Commercial, GNU GPL, GNU LGPL. Всі вони дозволяють розробляти програмне забезпечення з власною ліцензією.

За допомогою Qt реалізовано безліч додатків, таких як Autodesk Maya, Skype, Telegram, Медіапрогравач VLC, VirtualBox, а також використовується такими фірмами, як DreamWorks, Google, HP, Lucasfilm, Panasonic, Philips, Samsung, Siemens, Volvo, Walt Disney Animation Studios та інші.

OpenCV (англ. Open Source Computer Vision Library, бібліотека комп'ютерного зору з відкритим кодом) — бібліотека алгоритмів комп'ютерного зору, обробки зображень та чисельних алгоритмів загального призначення з відкритим кодом. Реалізована C/C++, також розробляється інших мов програмування. Може вільно використовуватися в академічних та комерційних цілях — поширюється за умов ліцензії BSD. Основне призначення бібліотеки — затвердження загального стандартного інтерфейсу комп'ютерного зору для додатків у цій галузі, сприяння зростанню числа таких додатків та створення нових моделей використання РС.

Dlib – крос-платформна бібліотека алгоритмів загального призначення з відкритим вихідним кодом, написана мовою програмування C++. Бібліотека є сукупністю незалежних програмних компонентів.

Dlib включає різноманітні інструменти, такі як програмні компоненти для роботи з мережею, потоками, графічними користувальницькими інтерфейсами, структурами даних, лінійною алгеброю, машинним навчанням, обробкою зображень, аналізом даних, XML і текстовим розбором, чисельною оптимізацією, байєсовськими мережами та багатьма іншими завданнями. Значну увагу приділено створенню широкого набору статистичних інструментів машинного навчання.

## 2.1.6 Загальний вид алгоритму, що описує технологію підготовки та обробки зображень

Загальний вигляд розробленої технології підготовки та обробки зображень осіб у відеопотоці представлений схемою на (рис. 2.4).

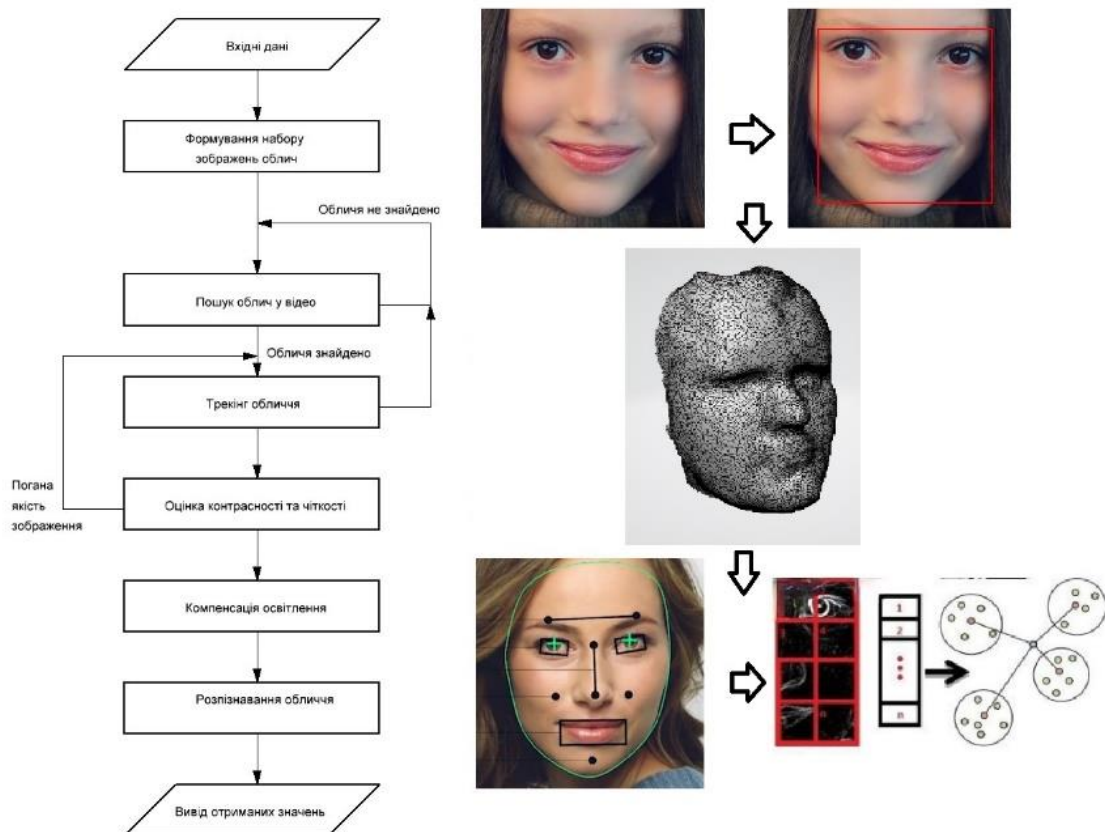


Рисунок 2.4 - Схема технології підготовки зображень облич до розпізнавання

Вихідні дані: растрові фронтальні зображення для кожної людини, яку потрібно розпізнавати; набір 3D-моделі обличчя, відеопотік.

Блок 1. Формування набору зображень осіб: для кожного зображення особи проводиться пошук особливих точок та модифікованим методом Шепард обчислюється індивідуальна 3D модель обличчя. На основі отриманої 3D моделі створюється набір зображень з різними кутами повороту обличчя, який використовується для формування бази даних кластерів.

Блок 2.Пошук обличчя у відеопотоці: за допомогою методу Віоли-Джонса здійснюється пошук обличчя у кадрі з відеопотоку, якщо обличчя знайдено, воно передається на обробку, інакше пошук триває в наступному кадрі.

Блок 3.Трекінг особи: координати положення зображення особи у двох сусідніх кадрах порівнюються відповідно до запропонованого алгоритму, якщо особа не знайдена, трек переривається, якщо знайдено, вона розглядається в сукупності з попереднім зображенням обличчя у тому ж треку.

Блок 4.Оцінка контрастності та різкості: за результатами цієї оцінки зображення може бути виключене з обробки, якщо воно недостатньо різке або чітке.

Блок 5.Компенсація освітлення: зображення приводиться до світлового діапазону, де знаходяться зображення в базі даних.

Блок 6.Розпізнавання особи: нормоване за висвітленням зображення обличчя передається на аналіз алгоритму, який витягує з нього біовектор ознак і порівнює його з кластерами біовекторів, сформованих з бази даних осіб.

Отримані дані: описана технологія повертає набір зображень обличчя у різних ракурсах, а також вибирає кадри з відеопотоку та нормує їх за висвітленням.

Для деяких блоків технології, запропонованих у рамках даної роботи, потрібен детальніший опис. На (рис. 2.5) представлено схема алгоритму формування індивідуальної 3D моделі обличчя.



Рисунок 2.5 - Схема алгоритму формування індивідуальної 3D моделі обличчя

Вихідні дані: фронтальне зображення обличчя, набір  $N$  відомих 3D моделей осіб.

Блок 1. Пошук особливих точок обличчя: особливі точки обличчя на зображенні та для відомих 3D моделей шукаються за допомогою модифікації методу ASM, що використовується в бібліотеці Dlib.

Блок 2. Нормування 3D моделей за особливими точками: для кожної 3D моделі обчислюється коефіцієнт масштабування з відстані між особливими точками-центрами очей і точка-початок координат.

Блок 3.Обчислення радіусу та вагових коефіцієнтів інтерполяції для кожної особливої точки: задається радіус інтерполяції для кожної особливої точки як число, що належить діапазону

$R \in [\min \text{Dist}; \max \text{Dist}]$ ,

де  $\min \text{Dist}$  – мінімальна, а  $\max \text{Dist}$  – максимальна відстань між точкою зображення та однією з 3D моделей. Обчислюються вагові коефіцієнти за формулою інтерполяції Шепарда.

Блок 4.Обчислення середньої ваги кожної моделі та виділення значних ваг: обчислюється середня для кожної 3D моделі вага по всіх спеціальних точках. Задається рівень, яким вплив 3D моделі вважається значним під час формування нової. Всі ваги нижче за рівень обнуляються.

Блок 5.Побудова нової 3D моделі за ваговими коефіцієнтами: точки нової 3D моделі обчислюються як комбінація точок відомих 3D моделей із відповідними ваговими коефіцієнтами.

Отримані дані:3D модель, що відповідає вихідному фронтальному зображенню.

На (рис. 2.6) представлено блок-схему алгоритму формування бази зображень осіб.

Вихідні дані:3D модель обличчя з відомими особливими точками, фронтальне зображення обличчя,  $N$  – кількість різних варіантів повороту,  $R$  – безліч значень кутів повороту.

Блок 1.Триангуляція за особливими точками: розбиття фронтального зображення обличчя на трикутники. В результаті цієї операції буде отримано набір трикутників, що описує область обличчя на зображенні

Блок 2.Розбиття трикутників: дроблення трикутників за вибраним критерієм, що дозволяє знизити розміри трикутників, що використовуються, і наблизити трикутну сітку до рівномірної.

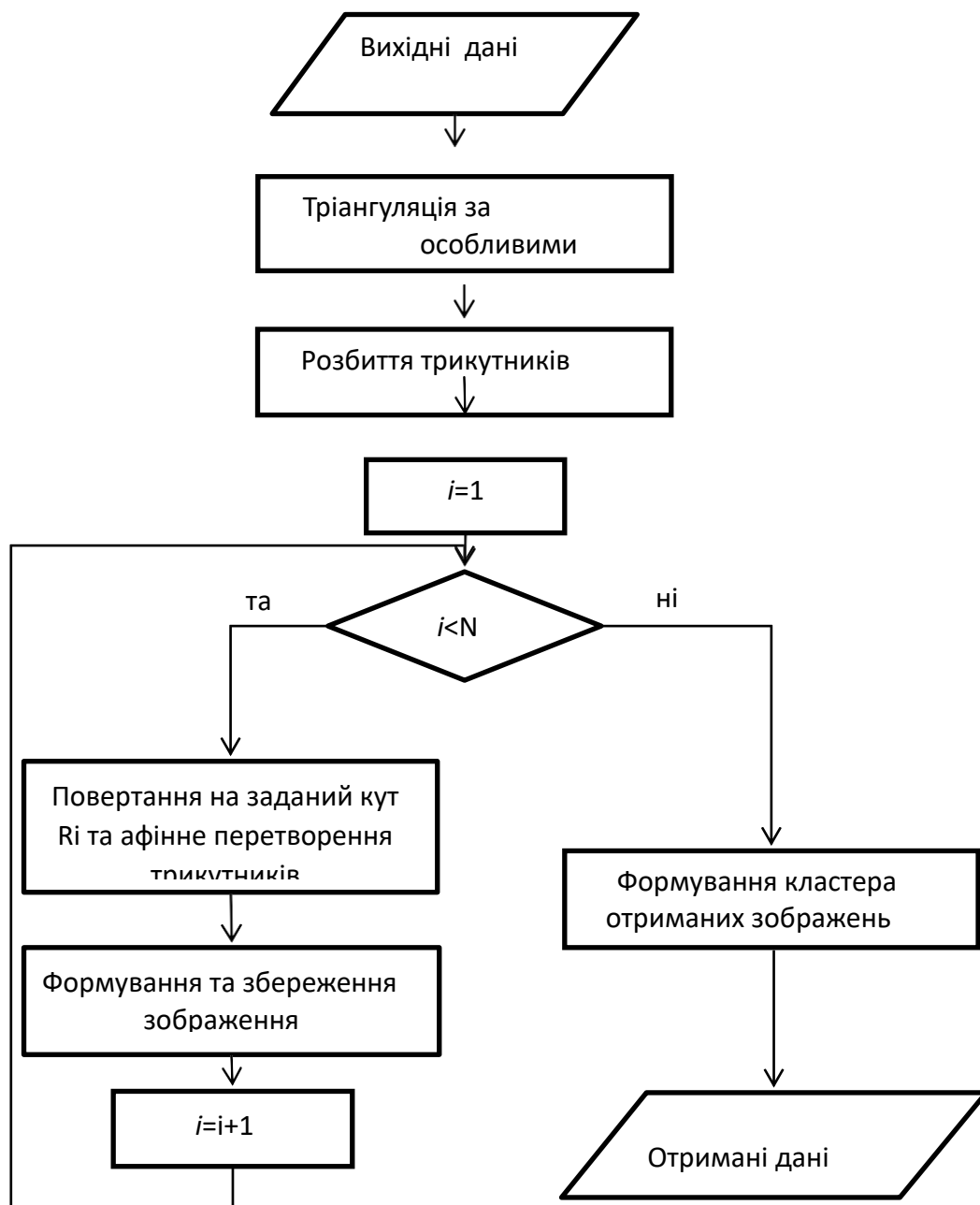


Рисунок 2.6 - Схема алгоритму формування бази даних осіб

Блок 3. Повертання на заданий кут  $R_i$  та афінне перетворення трикутників: кожен із трикутників за допомогою операцій обертання щодо заданого центру та афінного перетворення зміщується та деформується відповідно до величини кута  $R_i$ , заданої умовою.

Блок 4.Формування та збереження зображення: зображення обличчя, отримане після перетворення кожного з трикутників, зберігається як растрове для подальшого використання у формуванні бази даних.

Блок 5.Формування кластера отриманих зображень: зображення об'єднуються в єдиний масив, що відповідає конкретній людині та використовується для вилучення біовекторів у процесі навчання системи розпізнавання.

Отримані дані: набір зображень обличчя людини з різними курсами.  
На малюнку 30 представлена схема алгоритму трекінгу осіб.

Вихідні дані: набір кадрів відеопотоку,  $N$  – кількість кадрів.

Блок 1.Обчислення різниці фону та поточного кадру  $fT=f_i-f_0$ .

Блок 2.Обчислення різниці попереднього та поточного кадру  $fM=f_i-f_{i-1}$ .

Блок 3.Обчислення бінарних зображень  $V(fT)$  та  $V(fM)$  із заданим порогом, що залежить від шумів на зображенні.

Блок 4.Визначення областей для об'єкта та руху ковзним вікном:  $V(fT)>C$ ,  $V(fM)>C$ , де  $C$  – заданий поріг за кількістю білих пікселів у ковзному вікні.

Блок 5.Додавання певних областей бінарних зображень  $V(fT)$  та  $V(fM)$ .

Блок 6.Пошук осіб у певній області на вихідному зображенні

Методом Віоли-Джонса, всі знайдені особи утворюють початкові треки  $Trk$ .

Блок 7.Цикл кадрів відеопотоку.

Блок 8.Якщо особа на попередньому кадрі знайдена, ведеться пошук осіб на новому кадрі.

Блок 9.Для кожної нової особи проводиться перевірка, чи належить центр знайденого прямокутника області прямокутника на попередньому кадрі. Якщо так, нова особа додається треку, якщо ні, формується новий трек.

Блок 10.Усі знайдені треки передаються для подальшої обробки.

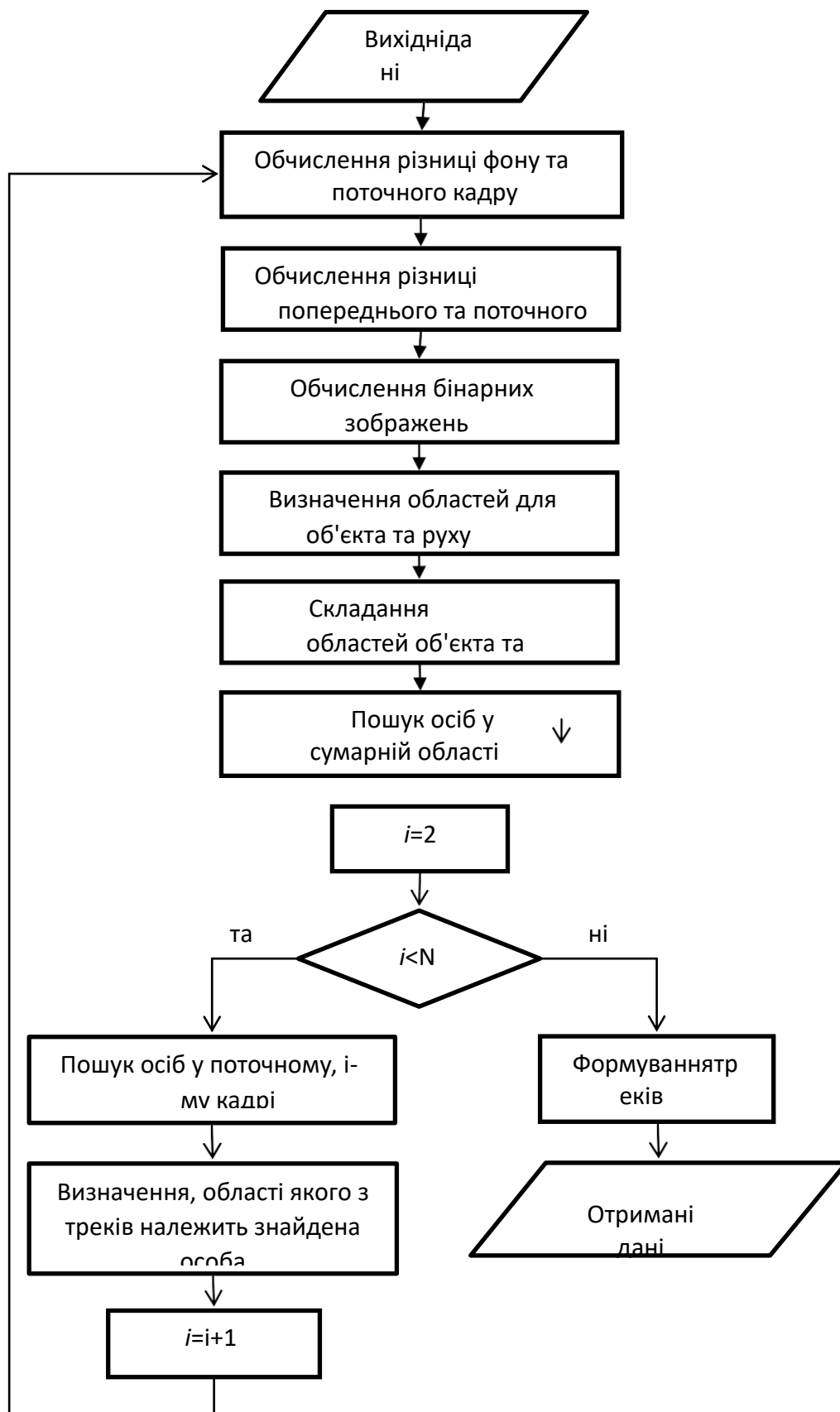


Рисунок 2.7 - Схема алгоритму трекінгу осіб

Отримані дані: набір треків осіб та окремих зображень, що не входять до треків (поодинокі розпізнавання).

На (рис. 2.8) представлена схема алгоритму порівняння зображень осіб, якою можна доповнити запропонований алгоритм трекінгу у разі аналізу зображень осіб, що повільно рухаються

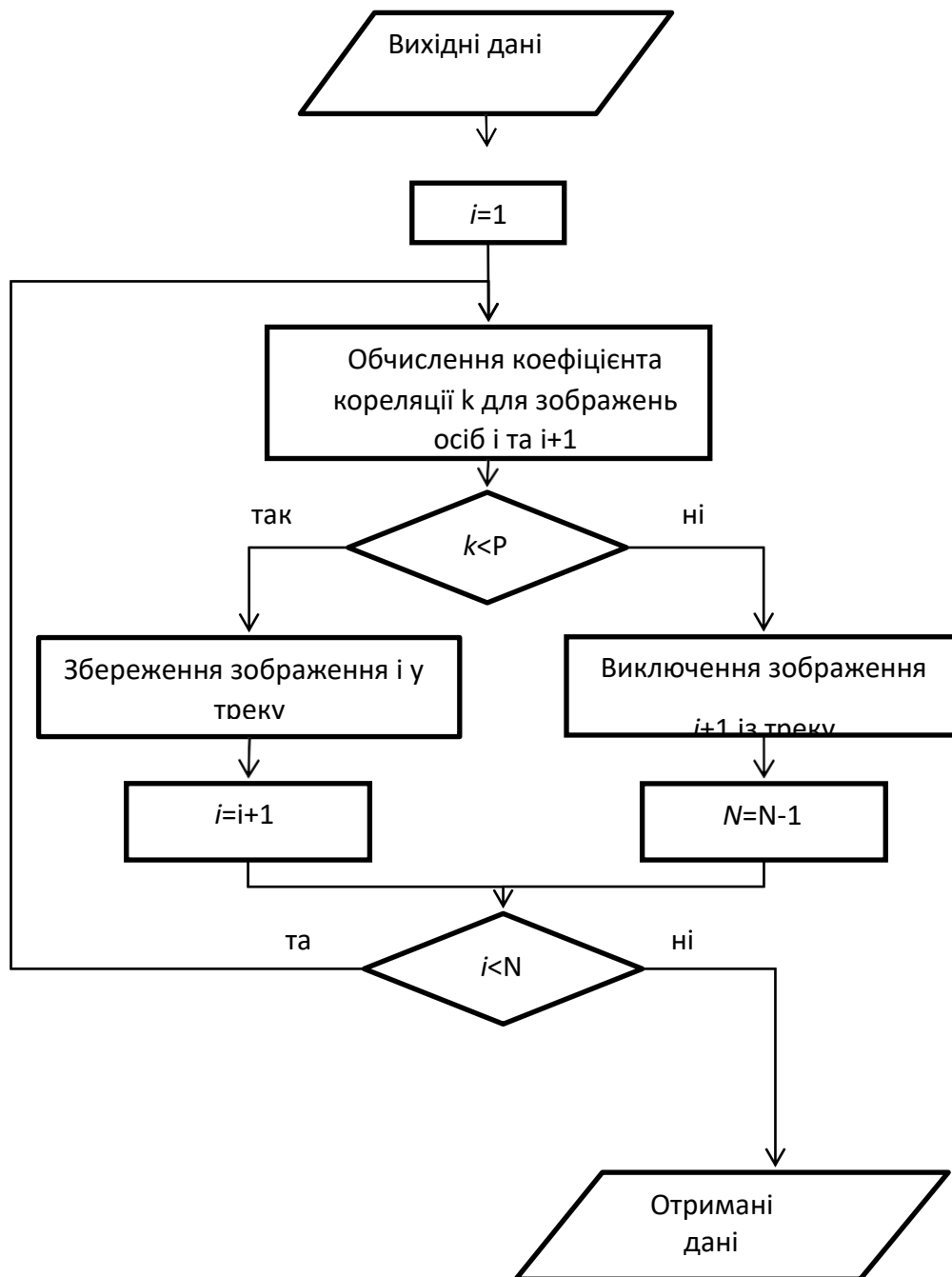


Рисунок 2.8 - Схема алгоритму порівняння зображень облич

Вихідні дані:трек зображень обличчя,  $N$  – число зображень,  $P$  – граничний коефіцієнт кореляції.

Блок 1.Для перших двох осіб обчислюється коефіцієнт взаємної кореляції.

Блок 2.Якщо обчислений коефіцієнт  $k$  більший, ніж заданий пороговий коефіцієнт  $P$ , що визначає ступінь схожості зображень обличчя, друге порівнюване зображення виключається з треку.

Блок 3. Якщо обчислений коефіцієнт  $k$  менший, то перше зображення зберігається в треку, а друге порівнюється з наступним.

Блок 4.Цикл продовжується доти, доки в треку не залишаться лише різні зображення обличчя, або одне зображення.

Отримані дані:трек різних зображень однієї особи.

## 2.2 Програмна реалізація алгоритмів

Розроблені алгоритми були впроваджені у програмний комплекс, що реалізовує захоплення відеопотоку з цифрової камери, обробку кадрів, розпізнавання особи з зображення обличчя та трекінг об'єктів. В рамках даного програмного комплексу було проведено тестування як на базах зображень осіб, так і на відеофайлах.

Створений програмний комплекс складається з кількох модулів, що здійснюють основні функції та алгоритми, що застосовуються при розпізнаванні: модуль налаштування, збереження та завантаження конфігурацій (рис. 2.9), модуль запуску та відображення результатів розпізнавання у відеопотоці (рис. 2.10), модуль формування біометричних векторів із бази зображень осіб (рис. 2.11) Кожен із модулів має інтерфейс користувача, написаний з використанням конструктора форм QT Creator.

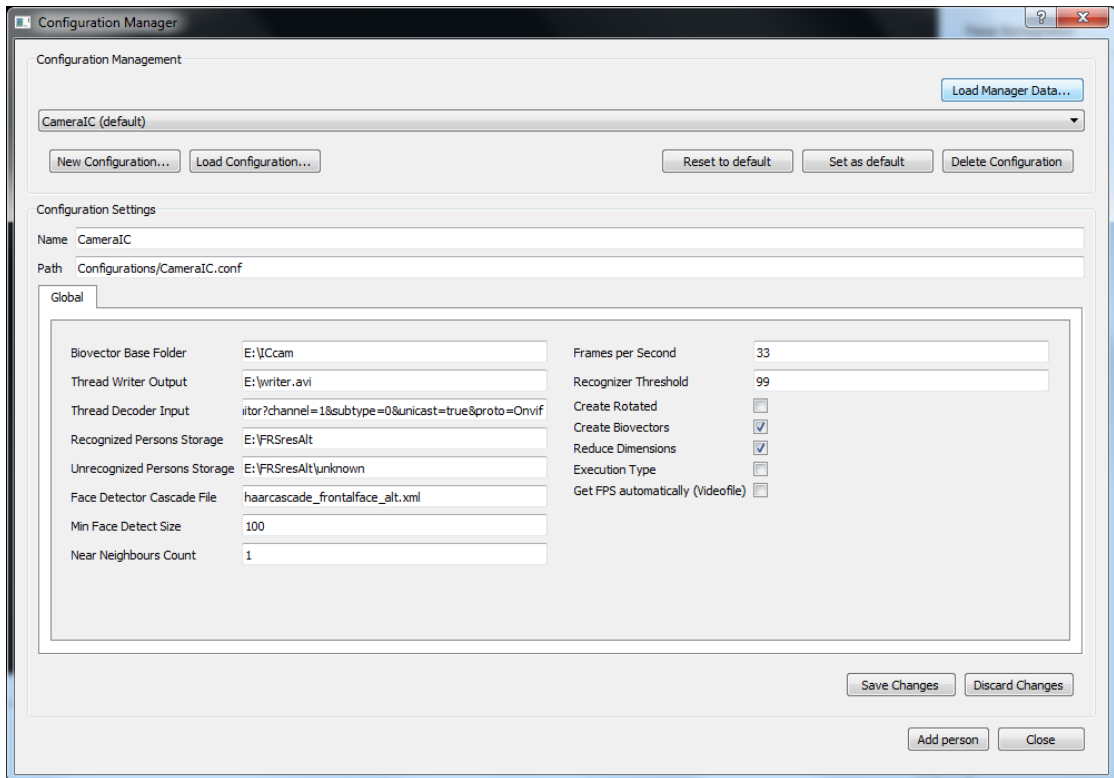


Рисунок 2.9 - Модуль менеджера конфігурацій

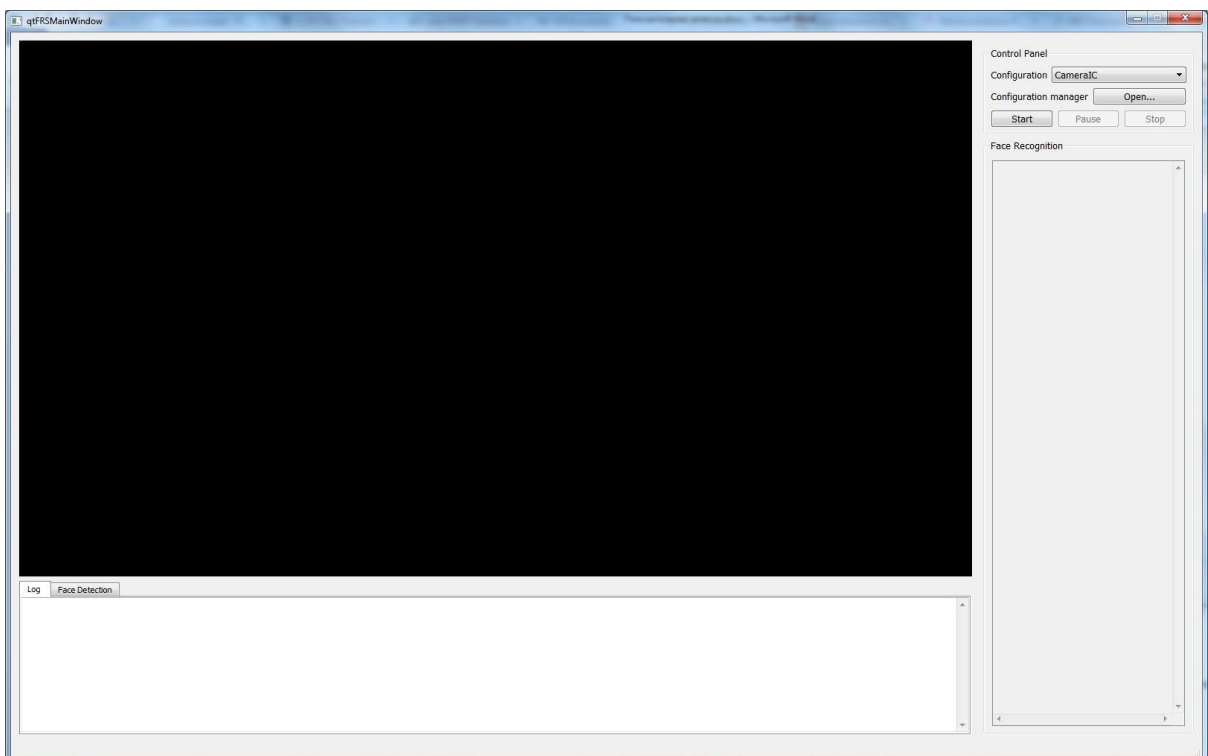


Рисунок 2.10 - Модуль основного вікна розпізнавання

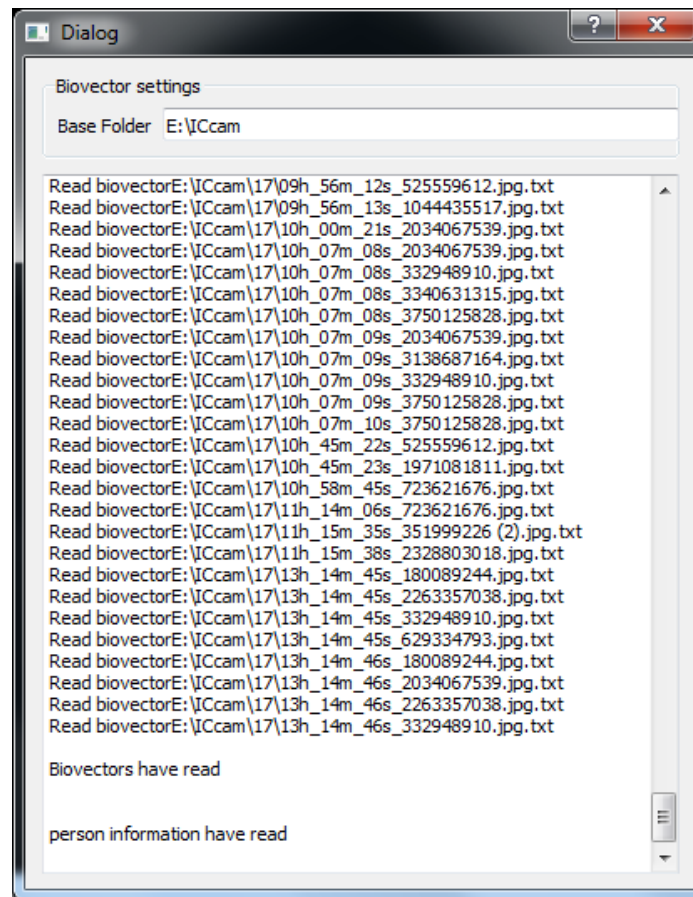


Рисунок 2.11 - Модуль формування бази зображень та біометричних векторів

Застосування всіх розроблених алгоритмів регулюється за допомогою модуля конфігурацій, після чого стає можливим запуск розпізнавання осіб у реальному часі як на відео, так і з довільної камери, підключеної до персонального комп'ютера. Робота в режимі реального часу забезпечується завдяки паралелізації алгоритму обробки кадрів відео з використанням функцій сучасних процесорів.

Для вибору алгоритму обчислення біометричних векторів було проведено тестування кількох існуючих методів спільно з методами нормування освітлення, в результаті якого було обрано алгоритм очищення стандартного відхилення (STD). Як алгоритм зниження розмірності біометричних векторів було обрано метод головних компонент (Princial Component Analysis, PCA). Як основний алгоритм класифікації обрано метод  $k$  найближчих сусідів ( $k$  Nearest Neighbours,  $k$ NN).

Висновки за розділом: Запропоновано алгоритм трекінгу осіб у відеопотоці, заснований на методі віднімання фону background subtraction та детекторі осіб Віоли-Джонса. Вибрано методи оцінки різкості, контрастності, порівняння зображень, які підходять для застосування в аналізі осіб на відео.

Запропоновано модифікацію методу побудови 3D моделі обличчя з використанням інтерполяції Шепарда стосовно узагальненої моделі та деформованої моделі.

Запропоновано алгоритм формування бази зображень облич для методів розпізнавання з використанням 3D моделей облич. Запропоновано модифікацію технології підготовки зображень осіб у відеопотоці до розпізнавання та її алгоритм.

Для всіх запропонованих алгоритмів та технології підготовки зображень осіб представлені схеми, алгоритми впроваджено у програмний комплекс, що реалізує розпізнавання особи з зображення обличчя у відеопотоці в режимі реального часу за допомогою методу класифікації зображень k Nearest Neighbours.

### 3 ТЕСТУВАННЯ АЛГОРИТМІВ І РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТУТОВ

#### 3.1 Підготовка експериментів.

Запропонована комплексна технологія попередньої обробки образів осіб для підготовки до розпізнавання в режимі реального часу, а також алгоритми компенсації умов зйомки (ракурсу та освітлення) і трекінгу, що входять до неї, вимагають експериментальної перевірки ефективності з точки зору їх застосовності до завдань розпізнавання особистості. Женню особи, а також порівняння з існуючими алгоритмами [11].

Для реалізації експерименту, що підтверджує ефективність розробленого алгоритму, необхідно здійснити виконання алгоритму розпізнавання з підключенням методів попередньої обробки зображень на базах даних зображень осіб, які знаходяться у вільному доступі та широко використовуються для тестування та порівняння методів ідентифікації особистості. Як такі бази даних вибрані бази Caltech [11] та FERET [11]. Також як експериментальні дані отримані зображення осіб на реальних відеозаписах та фотографіях.

База даних FERET містить 14051 8-бітове зображення людських осіб у відтінках сірого з різними ракурсами.

База даних Caltech містить 450 кольорових передніх зображень обличчя 27 людей з різним освітленням, тлом, виразом обличчя.

Реальні відеозаписи були отримані на IP-камері відеоспостереження RVi-IPC21WDN з варіофокальним об'єктивом мегапіксельним з ПЧ-корекцією RVi-1240AIR, а також за допомогою веб-камери Microsoft "LifeCam Studio For Business" 5WH-00002.

Експеримент проведено у кілька етапів:

- Деформована модель особи: візуальна перевірка розрізнення тривимірних моделей, що отримуються, для осіб з різною геометрією обличчя.
- Формування бази даних осіб: коректність триангуляції та операцій

перенесення зображення.

Попередня обробка зображень осіб:

- Оцінка ефективності розпізнавання з фільтрацією зображень за контрастністю та різкістю;
- Оцінка ефективності розпізнавання з вирівнюванням освітлення у кадрі різними методами;
- Порівняння числа аналізованих об'єктів при формуванні треків та при прямій обробці;
- Порівняння ефективності розпізнавання за різних способів формування бази осіб (одне зображення, набір зображень, зображення, повернені з індивідуальною моделлю, що деформується).

### 3.2 Тестування програмної реалізації алгоритмів.

#### 3.2.1 База для тестування алгоритмів

Початкове тестування працездатності та ефективності методів та алгоритмів проведено на записаних відрізках відео, всього записано 263 короткі ролики, в записі яких брало участь 12 осіб (рис. 3.1).

Всього на даних відео за допомогою стандартного детектора осіб було знайдено 21735 зображень осіб, включаючи зображення з недостатньою різкістю, ефектом змащеного зображення та з різними ракурсами щодо оптичної осі камери.



Рисунок 3.1 - Приклад фронтальних зображень осіб із відеозаписів

### 3.2.2 Алгоритм побудова набору зображень облич

Для кожного з 12 фронтальних осіб обчислено деформовану тривимірну модель особи згідно з алгоритмом, представленим на малюнку 28. За допомогою цих моделей отримано набори зображень осіб, повернутих на кут від  $-20$  до  $+20$  градусів по горизонталі щодо оптичної осі камери та від  $-15$  до  $+15$  градусів по вертикалі з кроком 5 градусів по кожній осі згідно з алгоритмом, представленим на малюнку 29. В результаті були отримані набори з 15 зображень кожної особи. Приклади сформованих зображень представлені (рис.3.2).

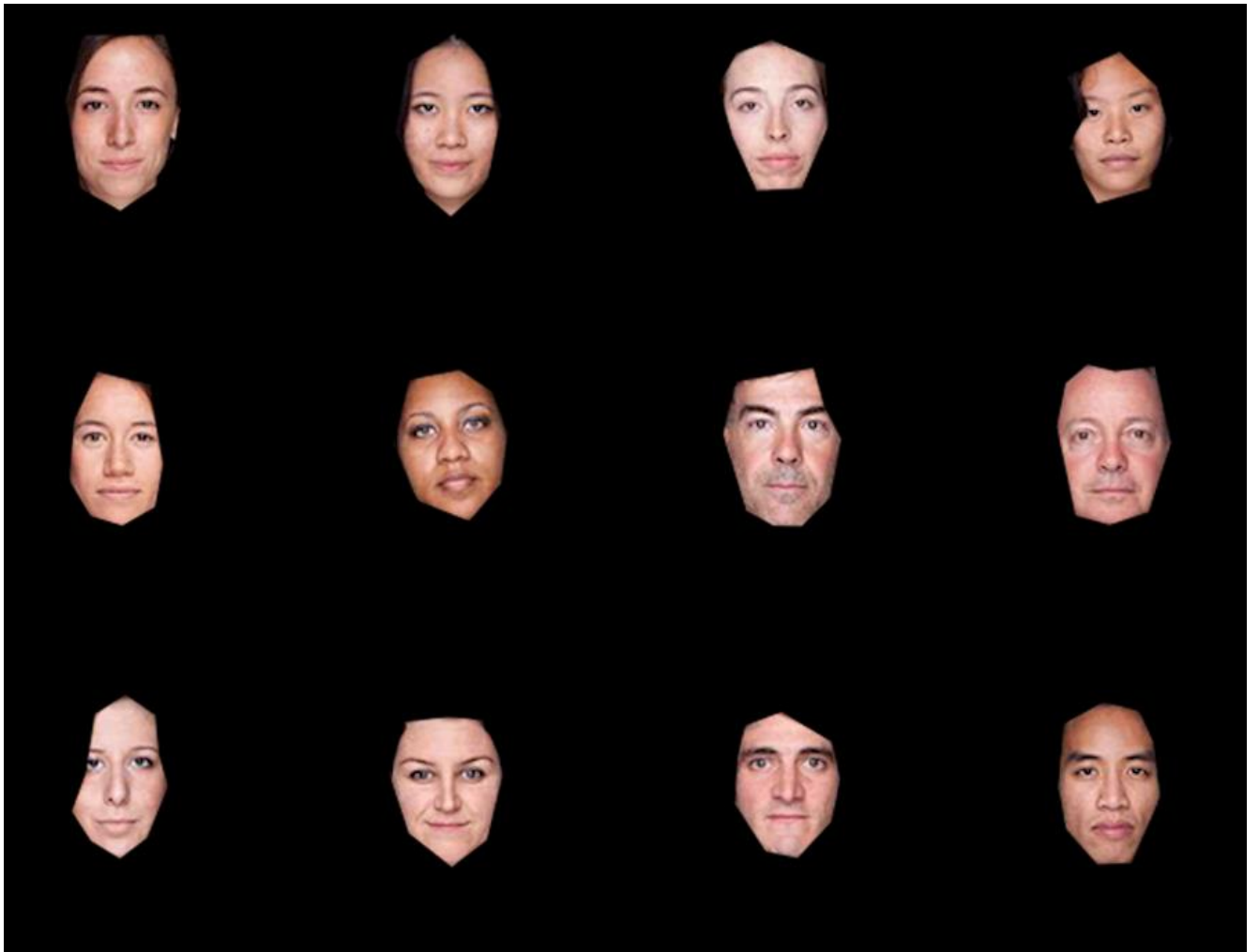


Рисунок 3.2 - Приклад зображень, отриманих за допомогою деформованих моделей осіб

Зображення, представлені на малюнку 36, підтверджують, що модель обличчя, що деформується, коректно зіставляє точки тривимірної моделі точкам плоского растрового зображення і шляхом афінних перетворень дозволяє отримати візуально вірні зображення осіб у різних ракурсах.

Для порівняння ефективності розпізнавання за допомогою методу найближчих сусідів зроблено вибір зображень осіб з різними ракурсами безпосередньо з відеопослідовностей за співвідношенням 15 зображень на людину.

Також порівняння проводилося з базою, складеною з припущення, що є лише одне зображення кожної людини.

### 3.2.3 Трекінг

Для початкового порівняльного тестування було взято два варіанти трекінгу: трекінг лише за допомогою методу Віюли-Джонса та трекінг за запропонованим алгоритмом з використанням віднімання фону та руху в кадрі.

Тестування проводилося наступної конфігурації комп'ютера: процесор Intel Core i7-3770 3,4 GHz, 16 Гб RAM, Windows 8.1 64-bit.

Оскільки тестування алгоритму трекінгу можливе лише на відео, бази Caltech та FERET у тестуванні брати участь не могли.

Власна база відео складається з 263 відеопослідовностей у форматі 1280\*720 (HD), на кожній з яких є одна людина, яка здійснює різноманітні рухи і повертається щодо оптичної осі камери на кути до 90 градусів. Загалом у кадрах детектором осіб на основі методу Віюли-Джонса у реалізації бібліотеки openCV було виявлено 21735 зображень осіб.

Тестування алгоритму, який використовує лише метод Віюли-Джонса, на власній базі відео дало наступні результати: із знайдених 21735 зображень особи вдалося скласти 1846 треків, середній час обробки одного кадру 16,75 мс.

Алгоритм, що використовує попередній аналіз фону та руху, дозволив знизити середній час обробки одного кадру до 2,36 мс, в основному за рахунок того, що кадри, що не містять відмінностей від фону та об'єктів, що рухаються, не допускаються до обробки детектором осіб. При цьому час обробки одного кадру, що містить об'єкти, що рухаються, знизилося незначно - в середньому 12,14 мс. Загальна кількість треків також поменшала – 1834 – завдяки тому, що в результаті звуження області аналізу частина помилкових спрацьовувань детектора осіб була виключена з подальшої обробки.

Середня довжина кожного треку становить 11 кадрів, це дозволяє виключати з обробки частину кадрів, наприклад ті, які мають недостатню різкість або контрастність, але при цьому з високою ймовірністю належать одній людині.

Алгоритм прискорення пошуку осіб застосовується для ситуацій, у яких ведеться стаціонарна зйомка з простим статичним фоном, у випадках

динамічного фону цей алгоритм може призвести до падіння продуктивності системи на час попередньої обробки за допомогою віднімання фону 2-8 мс.

Час обробки кожного кадру залежить від параметрів детектора області обличчя на зображенні, у тестових прикладах близько 10-15 мс, і не потребує додаткових обчислень після виявлення обличчя у кадрі. Запропонований алгоритм прискореного трекінгу об'єктів на статичному фоні не заважає застосовувати інші методи підвищення точності стеження за об'єктами на відео, у тому числі й ті, які забезпечують більш високу точність стеження при суттєвому зниженні швидкості [11]. Таким чином, швидкість запропонованого алгоритму трекінгу для відео зі статичним фоном на 27% вища порівняно з трекінгом на основі виділення області особи, яка не використовує попереднього аналізу фону та руху.

При зйомці в реальних умовах можливе зниження середньої довжини треку, але загальна тенденція збережеться, і трекінг як елемент попередньої обробки зображень є заможним та ефективним методом, дозволяючи не обробляти повторно особи вже розпізнаних у відеопослідовності людей.

#### 3.2.4. Фільтр по різкості та контрастності зображень осіб

Усі тестовані набори зображень осіб мають гарну різкість і контрастність, але все ж таки дозволяють провести порівняння впливу відсікання зображень нижньої межі оцінки цих параметрів на ефективність розпізнавання.

Тим не менш, на тестових наборах не вдалося отримати жодного помітного покращення результатів розпізнавання за допомогою відсіювання зображень за критеріями контрастності та різкості. Це з незначною відмінністю якості зображень в тестових наборах. При установці порога відсіювання зображень за контрастністю та різкістю FAR та FRR зберігалися пропорційно до загальної кількості зображень, що пройшли відбір. Ймовірно, застосування цих методів у запропонованому вигляді не підходить для зображень високої роздільної здатності та високої якості.

### 3.2.5 Алгоритм формування індивідуальної тривимірної моделі обличчя

Для випробування коректності підходу на формування зображень обличчя з різними ракурсами було використано тривимірну модель, отриману модифікованим методом Шепарда.

З одного фронтального зображення (рис. 3.4) за допомогою алгоритму було отримано набір зображень обличчя, поверненого на кут від  $-20$  до  $+20$  градусів щодо оптичної осі камери (рис. 3.3).



Рисунок 3.3 - Зліва – зображення обличчя, отримані поворотом уздовж вертикальної осі на 20 градусів, праворуч – уздовж горизонтальної осі



Рисунок 3.4 - Вихідне фронтальне зображення обличчя

Отримані зображення, використані для формування кластера біовекторів у системі ідентифікації особистості, дозволили обмежитись використанням однієї фотографії людини без зниження ефективності розпізнавання.

### 3.2.6 Вирівнювання освітлення та методи отримання ознак

Ефективність алгоритмів розпізнавання може бути оцінена двома імовірнісними характеристиками:

- FAR (False Accept Rate) – ймовірність помилкового виявлення, система приймає одну людину, яка не має фотографій у базі зображень системи розпізнавання, за іншу, зареєстровану в базі («розпізнали»);

- FRR (False Reject Rate) – ймовірність пропуску мети, система не реєструє збігу за фотографією зареєстрованого в ній користувача («не розпізнали»).

У системах розпізнавання ці ймовірності завжди є взаємозворотними величинами, у різних системах більш важливою вважається або перша, або друга характеристика.

Як середня оцінка ефективності розпізнавання зазвичай береться характеристика EER (Equal Error Rate) – рівний рівень помилок FAR і FRR. Алгоритм з меншою оцінкою EER вважається ефективнішим [12].

Однак така оцінка не завжди відображає реальну ефективність, оскільки вона залежить не тільки від мінімальних або оптимальних значень оцінок FAR і FRR, а й від динаміки їх зміни, сфери застосування алгоритмів. Наприклад, в системах, що контролюють вхід людей на об'єкти обмеженого допуску, можливість пропуску мети не така критична характеристика, як можливість помилкового виявлення. Виходячи з цих міркувань, прийmemo кінцевою оцінкою ефективності системи ідентифікації особи рівень FRR при заданому рівні  $FAR \leq 1\%$  від вибірки.

Описані методи вирівнювання освітлення були скомбіновані у різних варіантах. Тестування проводилося на базі облич Caltech Faces, що містить 450 зображень облич у різних варіаціях освітлення. Отримані вектори ознак порівнювалися за допомогою метрики Евкліда. У таблиці 3.1 наведено результати тестування розглянутих алгоритмів. Алгоритми відсортовані в порядку зменшення параметра FRR. Крім FRR у таблиці наведено оцінки часу формування вектора ознак, а також довжини одержуваного вектора ознак. Тут представлені комбінації методів вирівнювання освітлення разом із методами вилучення ознак (для обчислення біометричних векторів). Так, Difference of Gaussians (DoG), вейвлети Хаара (Haar), фільтр Габора (Gabor) та логарифм від нього (Log-Gabor) тестувалися спільно з обчисленням моментів Ху на основі центральних моментів зображення (HU), гістограмами спрямованих градієнтів (HoG), дискретним косинусним перетворенням (DCT) та стандартним відхиленням (STD).

Таблиця 3.1. Результати тестування алгоритмів вирівнювання освітлення та отримання вектора ознак, відсортовані в порядку зменшення EER

Методи	RR, % (при FAR=1%)	Час обробки зображення, з	Довжина вектор ознак
DoG + HU	23,548	0,055	175
Log-Gabor + HU	19,098	0,075	175
Haar + HoG	18,594	0,205	25205
Gabor + HU	16,79	0,095	175
Haar + HU	16,638	0,06	175
Gabor + HoG	15,632	0,215	25205
Gabor + HoG	14,557	0,235	25205
Haar + STD	12,294	0,135	4900
DoG + HoG	12,253	0,205	25205
DoG + DCT	11,835	0,135	50410
Gabor + STD	11,068	0,165	4900
Gabor + DCT	11,059	0,155	50410
DFT + HU	10,764	0,06	175
Haar + DCT	9,74	0,145	50410
DFT + DCT	9,736	0,155	50410
Gabor + DCT	9,561	0,095	250
DFT + HoG	8,628	0,215	25205
DFT+STD	7,03	0,135	4900
Gabor + STD	6,926	0,145	4900
DoG+STD	6,375	0,135	4900

З результатів, представлених у таблиці 3.1 можна зробити висновок, що найбільш ефективним в рамках даного дослідження виявився алгоритм, що ґрунтується на застосуванні фільтра Difference of Gaussians, фільтрації Log- Gabor

та витягуванні вектора ознак на основі обчислення стандартного відхилення. У порівнянні з іншими алгоритмами, наведеними в таблиці 3.1, даний алгоритм має прийнятний час роботи і довжину одержуваного вектора ознак. Також слід зазначити, що виключення з цього набору фільтра Log-Gabor слабо впливає на точність розпізнавання, але при цьому знижує час обробки на 10 мілісекунд, таким чином цей підхід може бути розглянутий для використання в системах розпізнавання в режимі реального часу разом з складнішою схемою компенсації умов освітлення.

### 3.2.7 Результати тестування комплексної технології підготовки зображень облич до розпізнавання

Найбільш ефективні з описаних методів підготовки зображення були скомбіновані у складі комплексної технології підготовки зображень осіб до розпізнавання. На основі цієї комбінації проводилося порівняння ефективності запропонованого алгоритму компенсації ракурсу на різних базах. Перше тестування проводилося з урахуванням, складеної з відео. Друге тестування проводилося з урахуванням осіб Caltech Faces. Третє – на базі осіб FERET 2001 року. Отримані вектори ознак порівнювалися за допомогою метрики Евкліда.

Перше тестування проводилося на базі, складеній із відео (~21 тисяча зображень осіб, 12 унікальних осіб). У таблицях 3.2 - 3.6 та (Рис. 3.5-3.8) представлені результати тестування системи розпізнавання з базою з 12 фронтальних зображень осіб, з базою з 15 зображень для кожної особи (всього 180) та з базою зображень, створених за допомогою деформованих моделей осіб.

Таблиця 3.2. Результати тестування системи розпізнавання облич на відеофайлах, одне зображення обличчя на людину

Поріг алгоритм-розпізнавання, %	Вірно розпізнані особи	Помилково опозначені особи	Вірно розпізнані особи, %	FAR, %	FRR, %
25	2385	0	10,973	0	89,027
30	3883	14	17,865	0,064	82,135
35	5491	97	25,263	0,446	74,736
40	7177	424	33,02	1,951	66,979
45	8792	1333	40,451	6,133	59,549
50	10249	2916	47,154	13,416	52,846

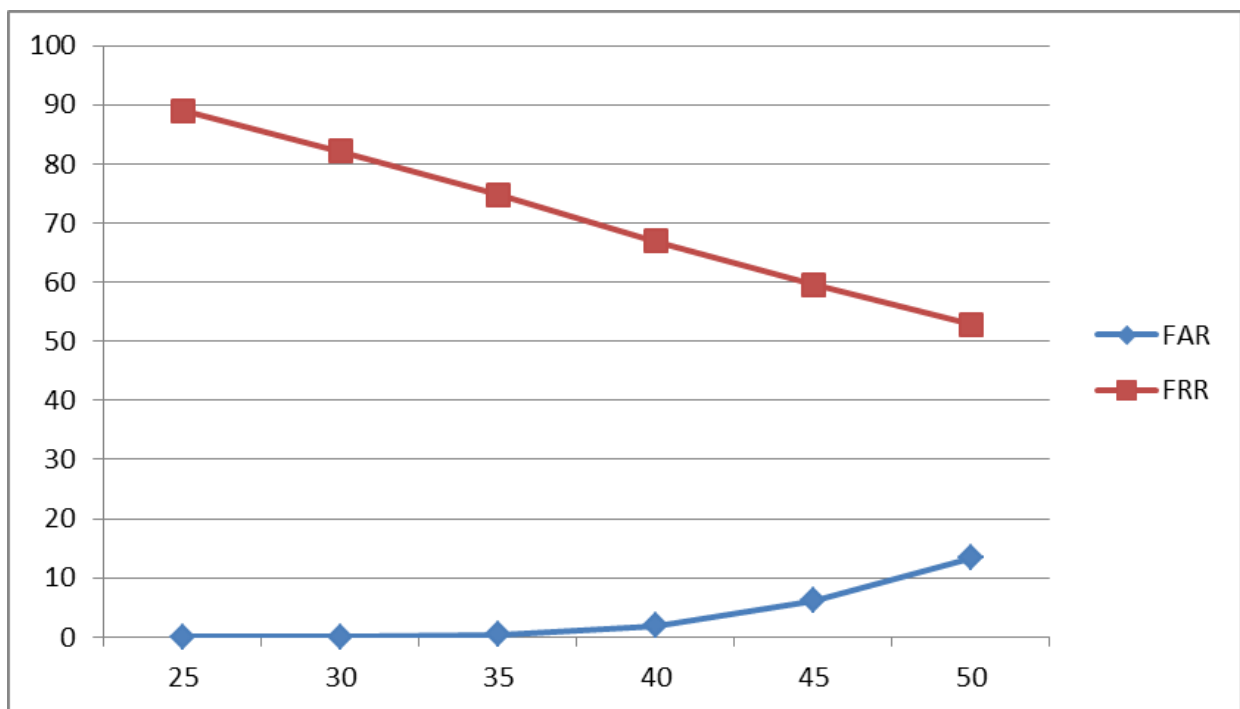


Рисунок 3.5 - Помилки FAR та FRR залежно від обраного порога розпізнавання, одне зображення обличчя на людину

Таблиця 3.3. Результати тестування системи розпізнавання осіб на відеофайлах,  
15 зображень на людину

Поріг алгоритм- розпізнавання, %	Вірно розпізнані особи	Помилково опо- знані особи	Вірно розпізнані особи, %	FAR, %	FRR, %
50	3959	0	18,215	0	81,785
55	5089	0	23,414	0	76,586
60	6156	18	28,323	0,083	71,677
65	7149	76	32,892	0,35	67,108
70	8066	207	37,111	0,952	62,889
75	8957	445	41,21	2,047	58,79

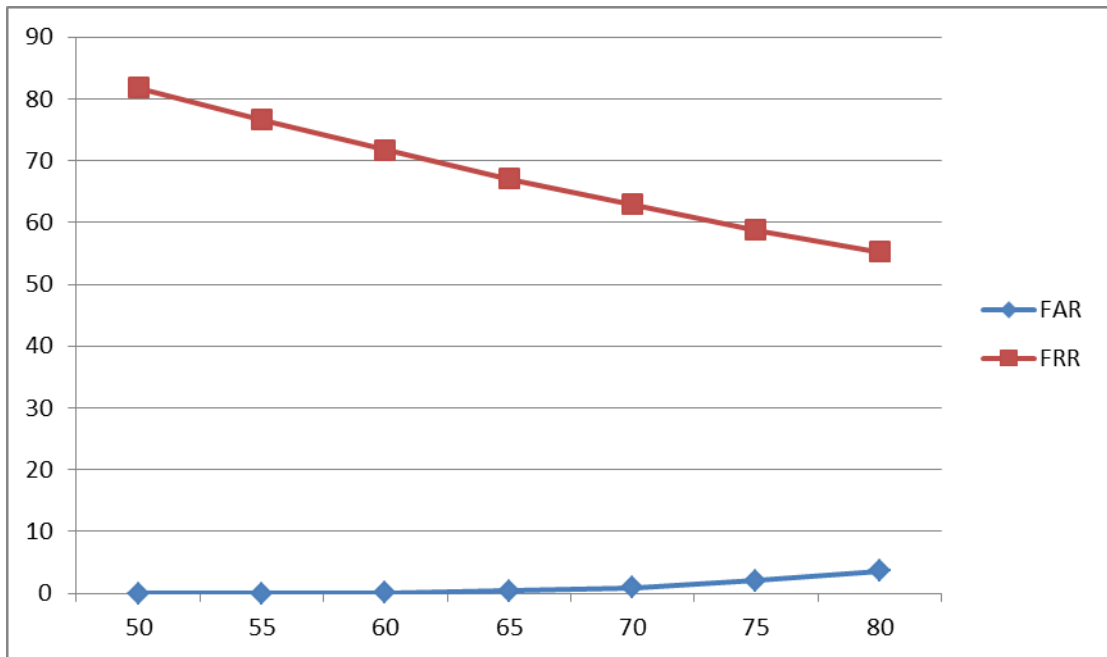


Рисунок 3.6 - Помилки FAR та FRR залежно від вибраного порога розпізнавання, 15 зображень на людину

Таблиця 3.4. Результати тестування системи розпізнавання осіб на відеофайлах, 15 згенерованих алгоритмом зображень на людину

Поріг алгоритму розпізнавання, %	Вірно розпізнані особи	Неправдиво впізнані особи	Вірно розпізнані особи, %	FAR, %	FRR, %
55	3580	0	16,471	0	83,529
60	4924	1	22,655	0,005	77,345
65	6315	10	29,055	0,046	70,945
70	7632	65	35,114	0,299	64,886
75	8767	156	40,336	0,718	59,664

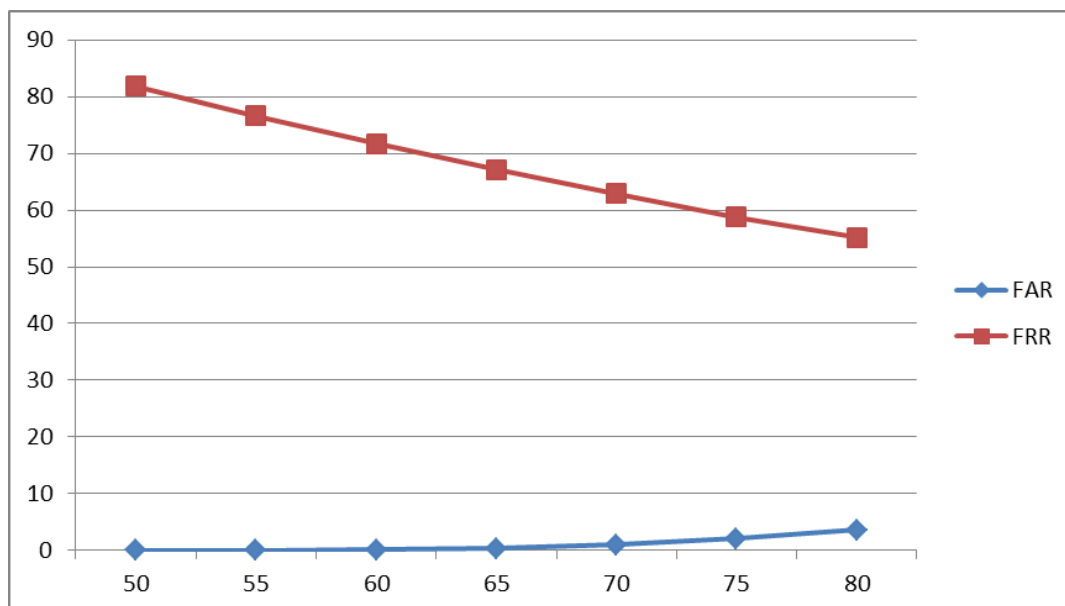


Рисунок 3.7 - Помилки FAR та FRR залежно від обраного порога розпізнавання, 15 згенерованих алгоритмом зображень на людину

Отримані при тестуванні результати об'єднані у вигляді таблиці взаємозалежності рівнів помилок FAR та FRR, а також у вигляді графіка PR-кривої, що показує відношення точності (precision) та повноти (recall). Тут повнота R – це частка вірно розпізнаних зображень осіб щодо загальної кількості прийнятих (вірно розпізнаних і хибно упізнаних), виражена у відсотках, а точність P – це частка вірно розпізнаних осіб щодо всіх шуканих осіб (вірно розпізнаних та хибно відвернутих), також виражена у відсотках.

Таблиця 3.5. F1-міра тестування на відеофайлах

Поріг розпізнавання	Одне зображення обличчя на людину	15 зображень на людину	15 згенерованих зображень
55	19,776	<b>37,944</b>	28,284
60	30,312	<b>44,135</b>	36,94
65	40,3	<b>49,459</b>	45,022
70	49,403	<b>53,992</b>	51,936
75	56,538	<b>58,013</b>	57,366
80	<b>61,057</b>	<b>61,168</b>	<b>61,737</b>

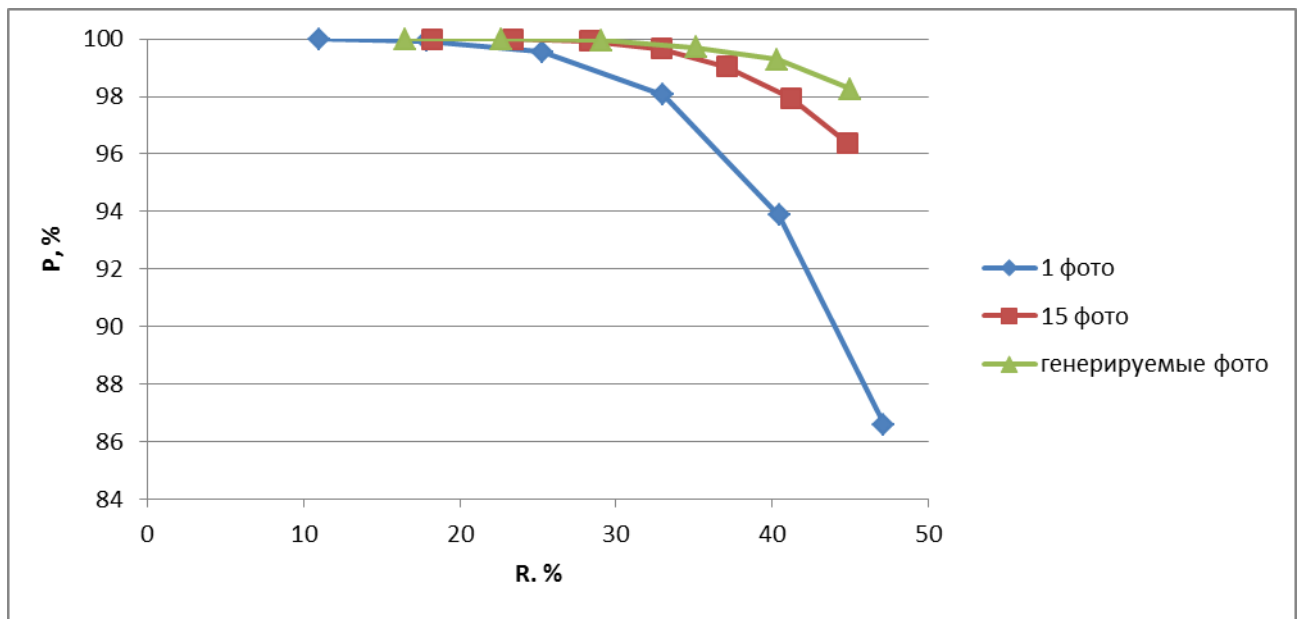


Рисунок 3.8 - PR-крива для тестування алгоритму на відеофайлах

Збалансована F1-мера кожної з баз зображень представлена таблиці 3.5.

Таблиця 3.6. Результати тестування системи розпізнавання осіб на відеофайли, FRR при FAR=1%

База зображень	RR, % (при FAR=1%)
Одне зображення обличчя на людину	8,069
15 зображень на людину	4,82
15 згенерованих алгоритмом зображень на людину	2,179

Як можна бачити з таблиць 3.5- 3.6, використання згенерованих зображень як бази зображень осіб для записаних відеофайлів не тільки не знижує точність розпізнавання, але й виявляється на 2.6% ефективнішим, ніж набір із 15 зображень, що знаходяться в початковій базі.

Друге тестування проводилося з урахуванням зображень обличч Caltech Faces [11]. Ця база містить зображення 26 осіб, знятих у фронтальному ракурсі, але з різною якістю та мімікою, загальна кількість зображень 450. Оскільки кількість

зображень кожної людини в базі різна, при тестуванні використовувалися наступні варіанти: 1 зображення на людину, від 3 до 6 зображень на людину та 15 згенерованих зображень на людину. У таблицях 3.7- 3.11 та на (рис. 3.7-3.10) подано основні результати тестування.

Таблиця 3.7. Результати тестування системи розпізнавання облич на базі зображень облич Caltech Faces, одне зображення обличчя на людину

Поріг алгоритм-розпізнавання, %	Вірно розпізнані особи	Помилково опозначені особи	Вірно розпізнані особи, %	FAR, %	FRR, %
25	295	0	72,839	0	27,16
30	326	0	80,494	0	19,506
35	354	0	87,407	0	12,593
40	374	1	92,346	0,247	7,654
45	383	2	94,568	0,494	5,432

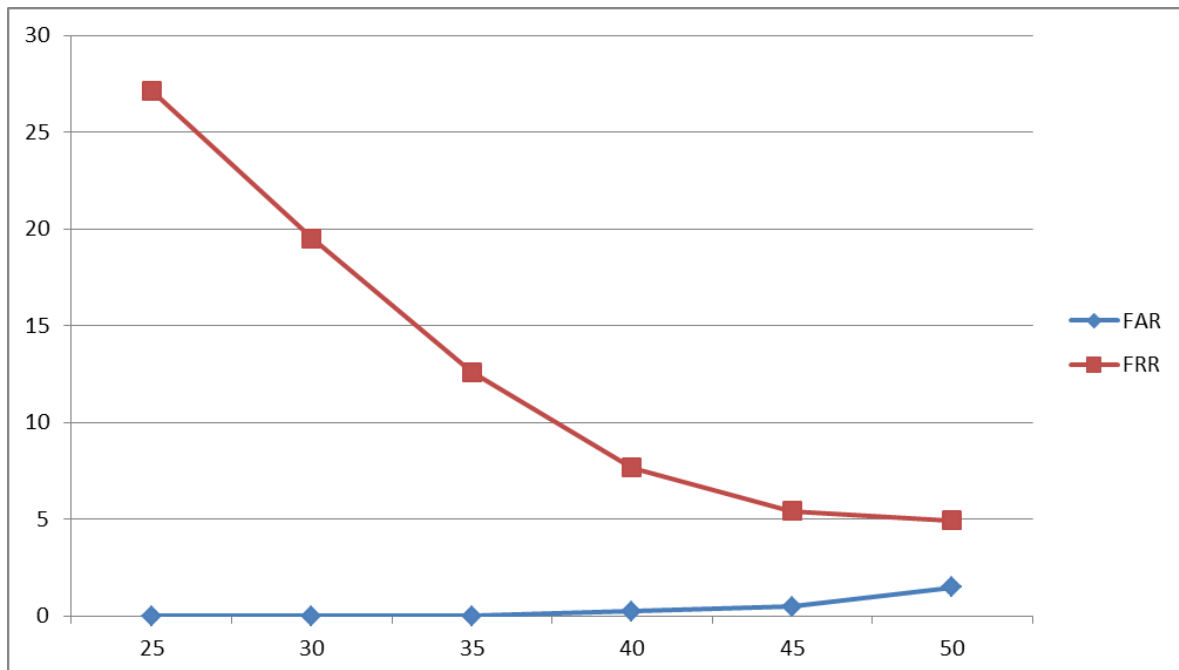


Рисунок 3.9 - Помилки FAR та FRR залежно від обраного порога розпізнавання, одне зображення обличчя на людину

Таблиця 3.8. Результати тестування системи розпізнавання облич на базі зображень облич Caltech Faces, 3-6 зображень на людину

Поріг алгоритма розпізнавання, %	Вірно розпізнані особи	Хибно впізнані чия	Вірно розпізнані особи, %	FAR, %	FRR, %
50	280	0	93,959	0	6,04
55	285	0	95,637	0	4,362
60	287	0	96,309	0	3,691
65	289	0	96,98	0	3,02
70	291	0	97,651	0	2,349
75	291	1	97,651	0,336	2,349
80	291	3	97,651	1,007	2,349

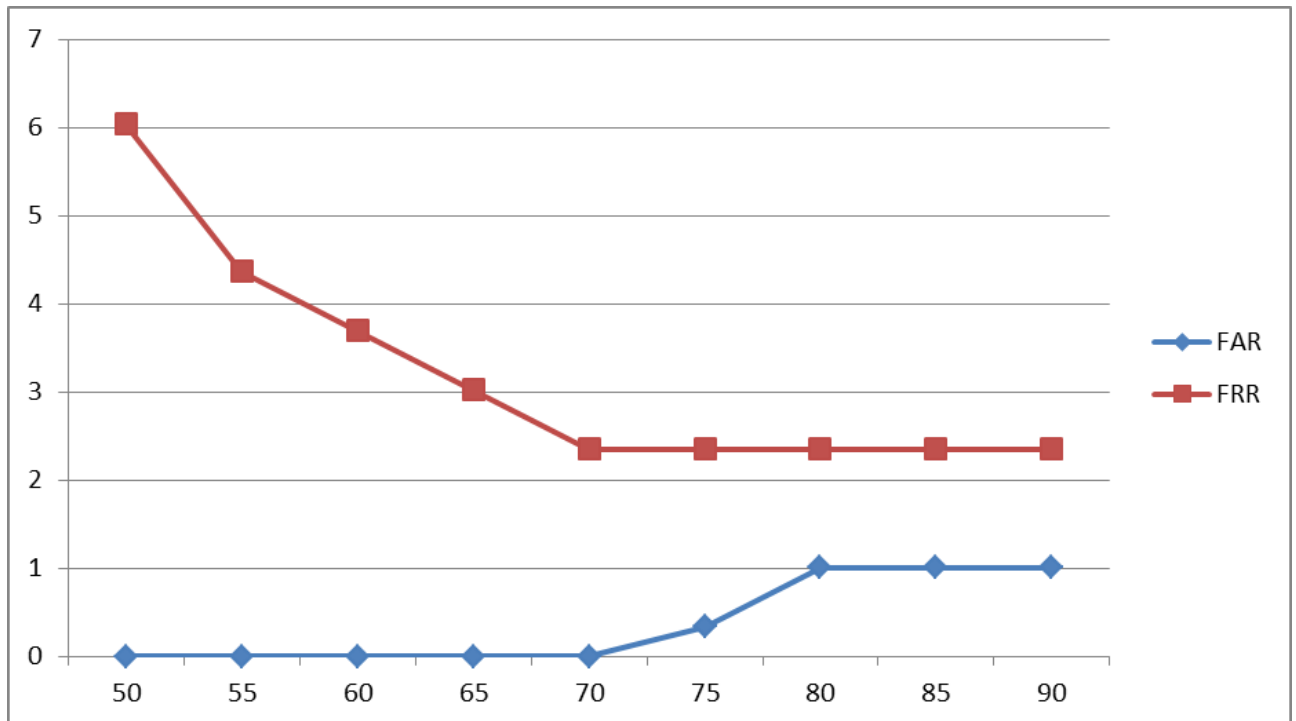


Рисунок 3.10 - Помилки FAR та FRR залежно від вибраного порога розпізнавання, 3-6 зображень на людину

Таблиця 3.9. Результати тестування системи розпізнавання облич на базі зображень облич Caltech Faces, 15 згенерованих алгоритмом зображень на людину

Поріг алгоритм-розпізнавання, %	Вірно розпізнані особи	Помилково опозначені особи	Вірно розпізнані особи, %	FAR, %	FRR, %
50	320	0	79,012	0	20,988
55	333	0	82,222	0	17,778
60	356	0	87,901	0	12,099
65	365	0	90,123	0	9,876
70	376	0	92,839	0	7,16
75	382	1	94,32	0,247	5,679
80	386	2	95,309	0,494	4,691
85	389	3	96,049	0,741	3,951
95	391	4	96,543	0,988	3,457
99	393	4	97,037	0,988	2,963

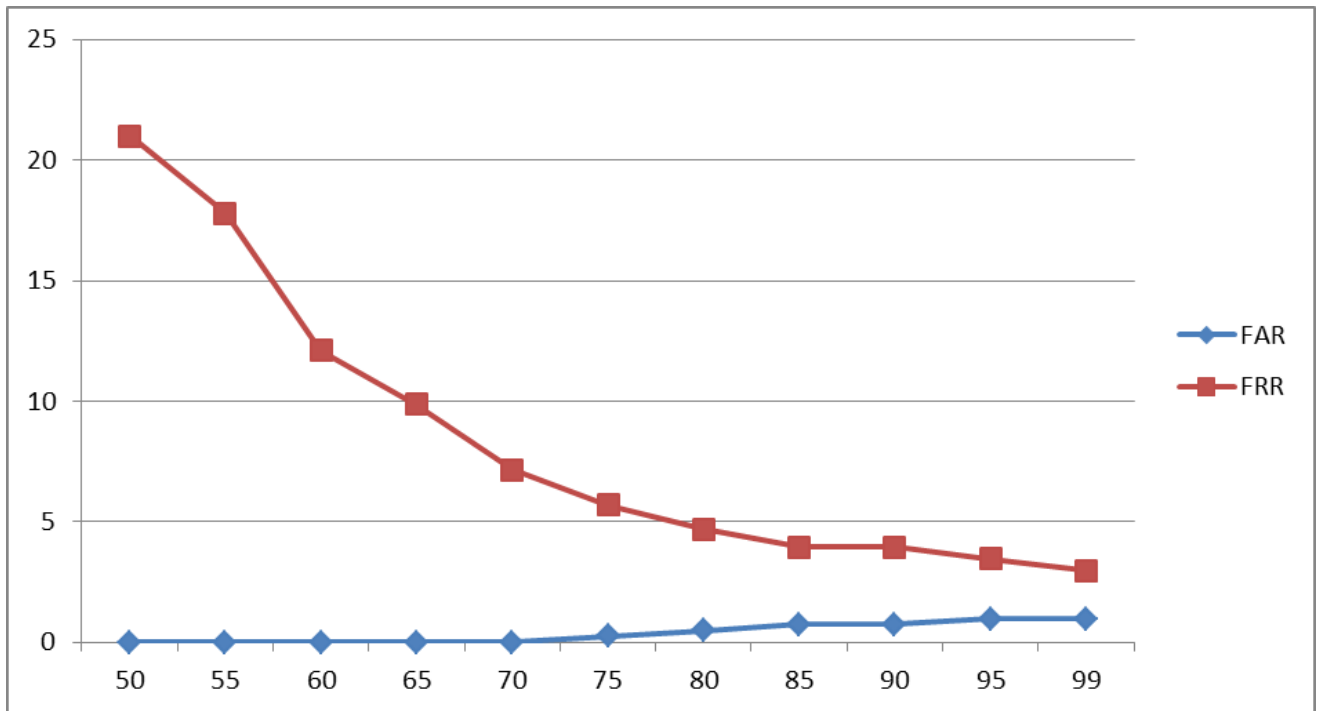


Рисунок 3.11 - Помилки FAR та FRR залежно від обраного порога розпізнавання, 15 згенерованих алгоритмом зображень на людину

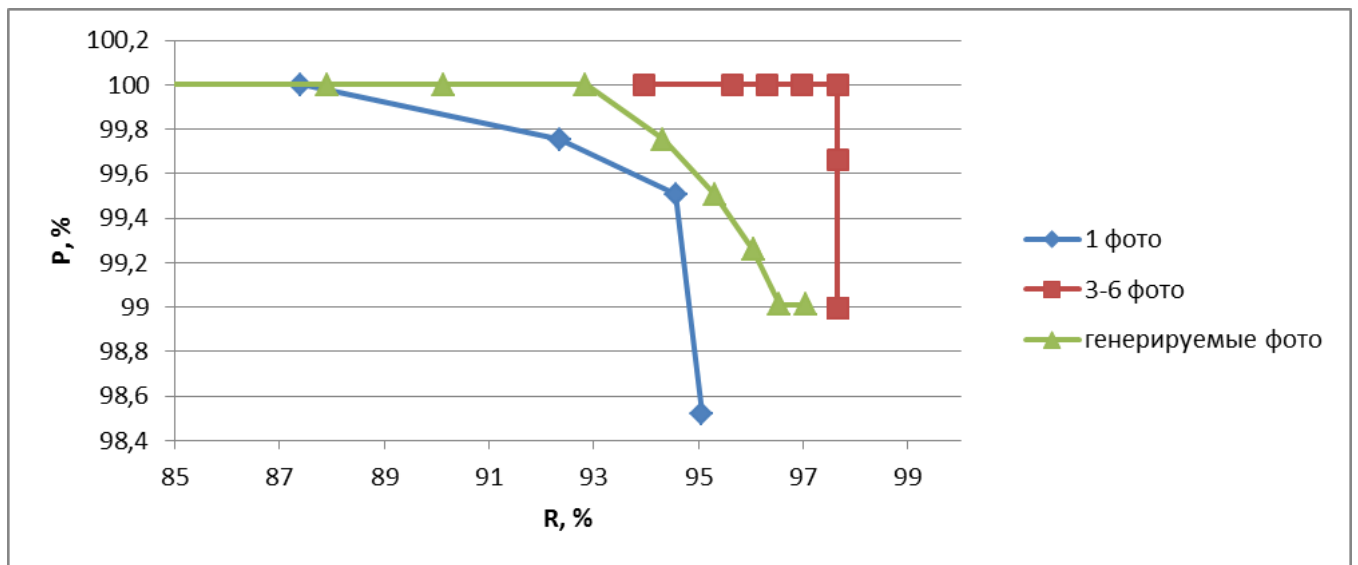


Рисунок 3.12 - PR-крива для тестування алгоритму з урахуванням Caltech Faces

Таблиця 3.10 - F1-мера тестування з урахуванням Caltech Faces

Поріг розпізнавання	Одне зображення обличчя на людину	15 зображень на людину	15 згенерованих зображень
75	84,286	98,467	96,961
80	89,193	98,812	97,362
85	93,281	98,647	97,628
90	95,907	98,318	97,628
95	96,974	98,318	97,762
99	96,759	98,318	98,015

Таблиця 3.11 - Результати тестування системи розпізнавання осіб з урахуванням зображень обличч Caltech Faces, FRR при FAR=1%

База зображень	RR, % (при FAR=1%)
Одне зображення обличчя на людину	5,185
3-6 зображень на людину	2,349
15 згенерованих алгоритмом зображень на людину	3,457

Як можна бачити з таблиць 10-11, використання згенерованих зображень виявляється на 1% менш ефективним для бази, що складається з фронтальних фотографій, ніж використання кількох зображень для кожної людини, але на 1,7% ефективнішим, ніж використання одного зображення.

Третє тестування проводилося з урахуванням зображень осіб FERET [11]. Ця база містить зображення 996 осіб, знятих у різних ракурсах (повороти на 0, 45 і 90 градусів по горизонталі щодо оптичної осі камери), з різною якістю, освітленням та мімікою, загальна кількість зображень 11297, з них успішно упізнаються детектором 5335. Оскільки кількість зображень кожної людини в базі різна, при тестуванні використовувалися наступні варіанти: 1 зображення на людину, від 2 до 6 зображень на людину та 15 згенерованих зображень на людину.

Для розпізнавання була складена база з 30 осіб, загальна кількість зображень 181. У таблицях 12-16 та на рисунках 47-50 представлені основні результати тестування.

Таблиця 3.12 - Результати тестування системи розпізнавання облич на базі зображень облич FERET, одне зображення обличчя на людину

Поріг алгоритму розпізнавання, %	Вірно розпізнані особи	Помилково опознані особи	Вірно розпізнані особи, %	FAR, %	FRR, %
25	58	6	51,381	0	48,619
30	67	26	53,039	0,056	46,961
35	75	216	54,696	0,131	45,304
40	82	865	58,564	0,356	41,436
45	86	1999	61,326	1,818	38,674
50	94	3243	64,641	6,542	35,359
55	99	4206	66,851	15,933	33,149

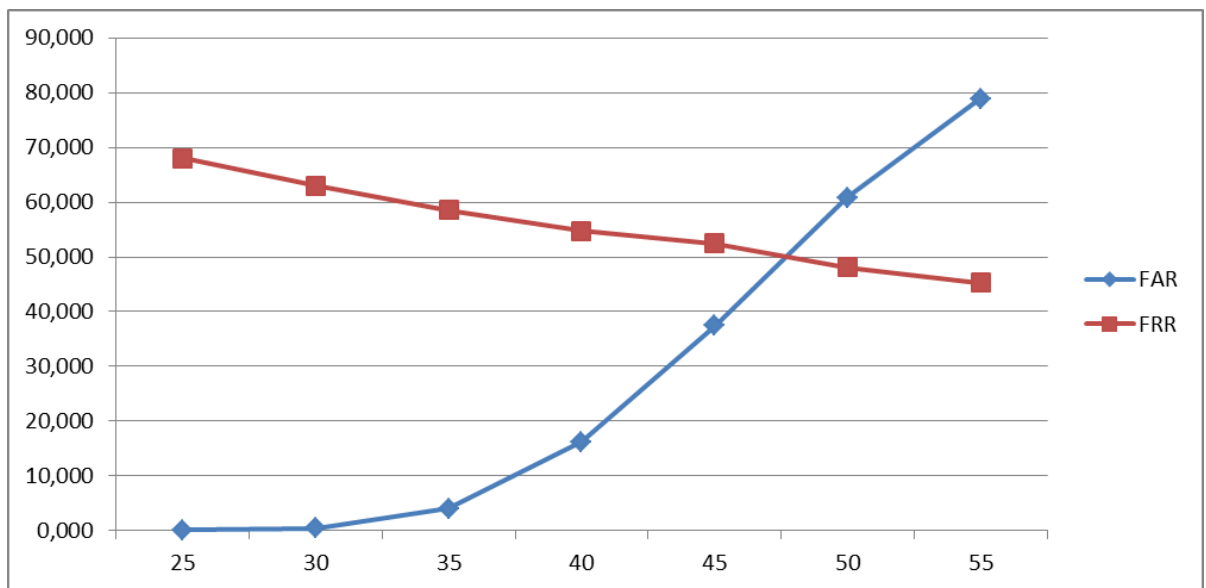


Рисунок 3.13 - Помилки FAR та FRR залежно від обраного порога розпізнавання, одне зображення обличчя на людину

Таблиця 3.13 - Результати тестування системи розпізнавання облич на базі зображень облич FERET, 2-6 зображень на людину

Поріг алгоритм-розпізнавання, %	Вірно розпізнані особи	Помилково опознані особи	Вірно розпізнані особи, %	FAR, %	FRR, %
25	93	0	51,381	0	48,619
30	96	3	53,039	0,056	46,961
35	99	7	54,696	0,131	45,304
40	106	19	58,564	0,356	41,436
45	111	97	61,326	1,818	38,674
50	117	349	64,641	6,542	35,359
55	121	850	66,851	15,933	33,149

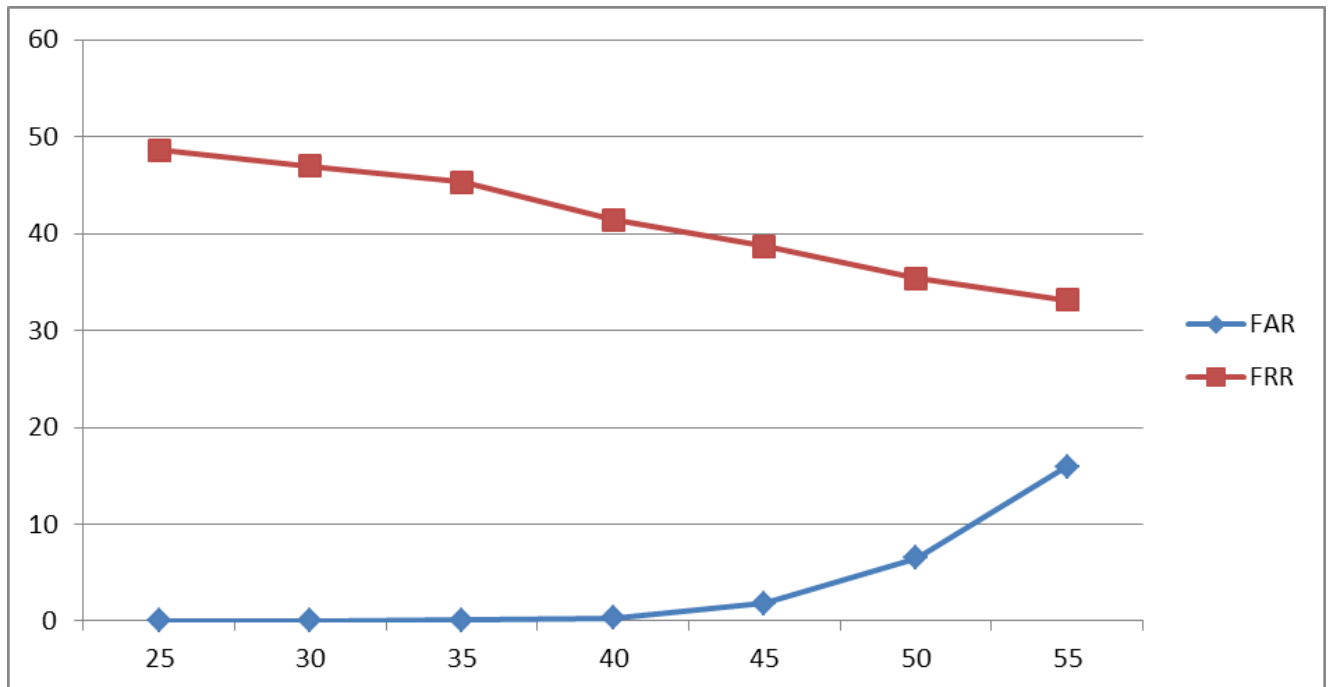


Рисунок 3.14 - Помилки FAR та FRR залежно від вибраного порога розпізнавання, 2-6 зображень на людину

Таблиця 3.14 - Результати тестування системи розпізнавання осіб на базі зображень осіб FERET, 15 згенерованих алгоритмом зображень на людині

Поріг алгоритму розпізнавання, %	Правильно розпо- знані особи	Неправдиво впізнані особи	Правильно розпо- знані особи, %	FAR, %	FRR, %
50	54	0	29,834	0	70,166
55	59	2	32,597	0,037	67,403
60	65	12	35,912	0,225	64,088
65	71	55	39,227	1,031	60,773
70	72	206	39,779	3,861	60,221
75	80	532	44,199	9,972	55,801
80	86	1023	47,514	19,175	52,486

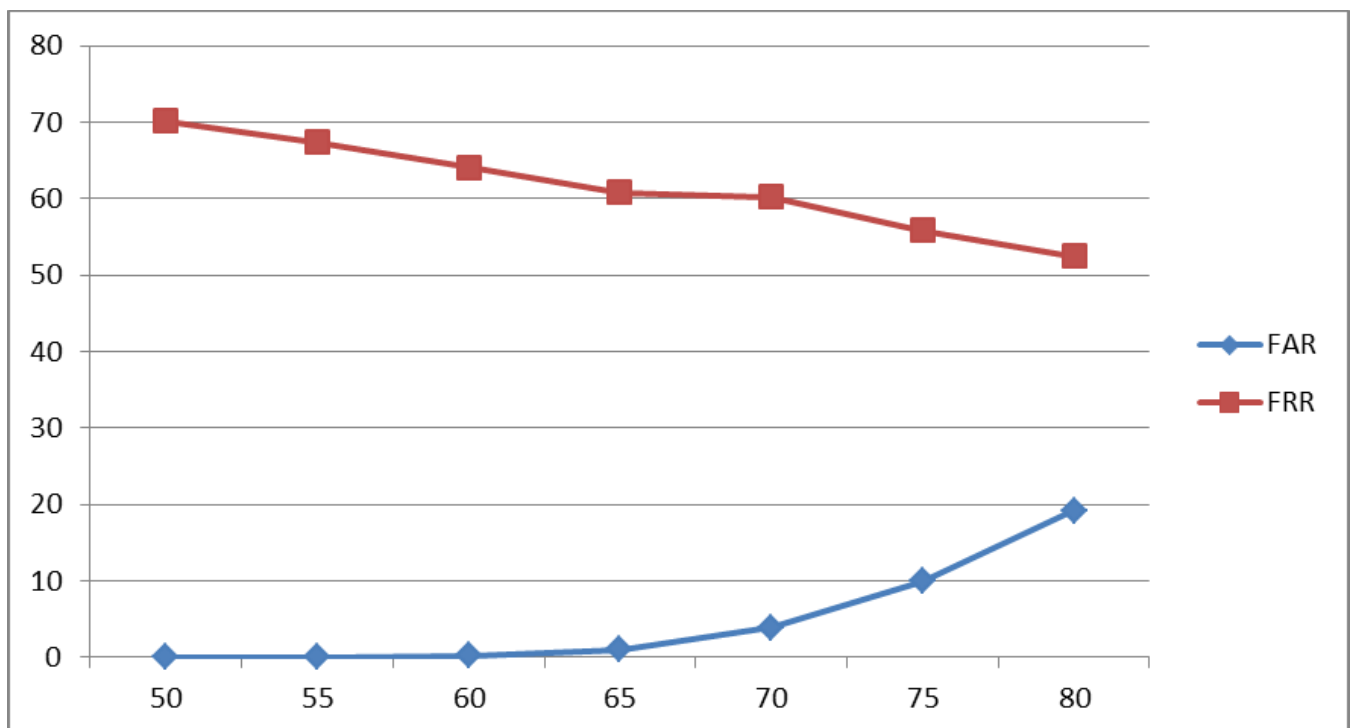


Рисунок 3.15 - Помилки FAR та FRR залежно від обраного порога розпізнавання, 15 згенерованих алгоритмом зображень на людину

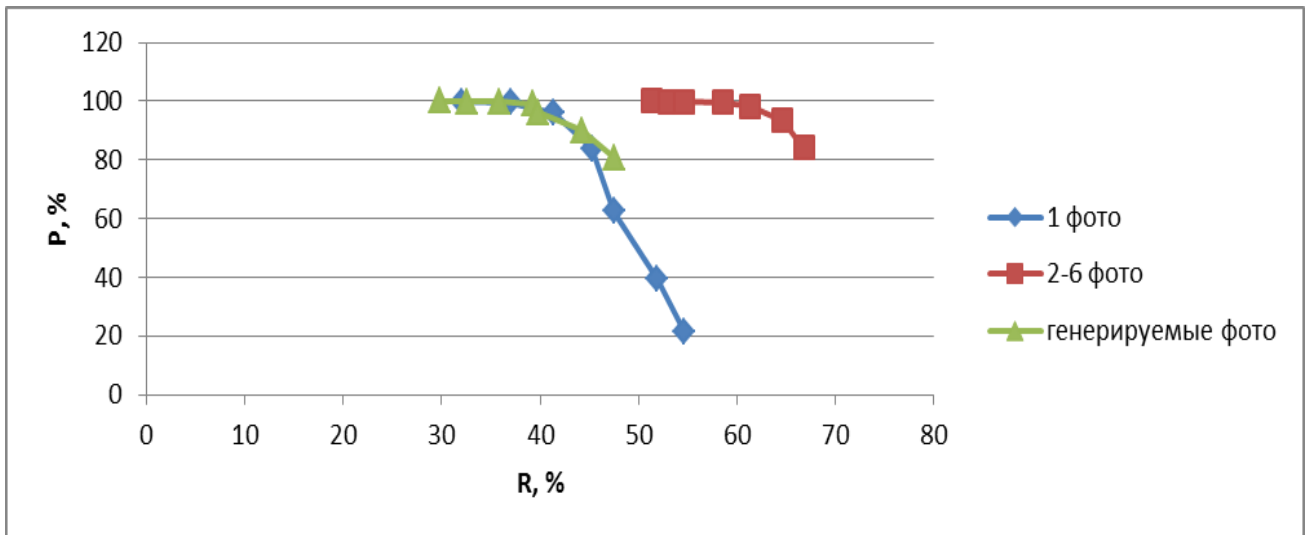


Рисунок 3.16 - PR-крива для тестування алгоритму на базі FERET

Таблиця 3.15 - F1-міра тестування на базі FERET

Поріг розпізнавання	Одне зображення особи на людину	15 зображень на людину	15 згенерованих зображень
50	48,522	<b>67,883</b>	45,957
55	53,961	<b>69,301</b>	49,162
60	57,878	<b>70,681</b>	52,814
65	<b>58,809</b>	<b>73,77</b>	56,184
70	53,998	<b>75,496</b>	56,274
75	44,685	<b>76,423</b>	59,29
80	30,517	<b>74,477</b>	<b>59,846</b>

Таблиця 3.16 - Результати тестування системи розпізнавання осіб з урахуванням зображень осіб FERET, FRR при FAR=1%

База зображень	RR, % (при FAR=1%)
Одне зображення обличчя на людину	61,878
2-6 зображень на людину	41,436
15 згенерованих алгоритмом зображень	61,326

Як можна бачити з таблиць 3.15- 3.16, використання згенерованих зображень виявляється на 20% менш ефективним для бази, що складається з великої кількості фотографій, ніж використання кількох зображень для кожної людини, але все ж таки на 0,55% більш ефективним, ніж використання одного зображення. Отримані вектори ознак порівнювали за допомогою метрики.

З таблиць 3.5, 3.11 і 3.16 можна дійти невтішного висновку, що результати тестування по вибраній метриці підтверджують ефективність запропонованого підходу: використання алгоритму розширення набору біометричних векторів на власній базі зображень підвищує точність розпізнавання порівняно з підходом, коли для формування біометричних векторів береться набір зображень із самої бази. На базах Caltech Faces та FERET тестування показало, що набір синтезованих зображень не дозволяє досягти точності розпізнавання з набором природних зображень, але підвищує точність розпізнавання порівняно з одиночними зображеннями.

Це дозволяє стверджувати, що використання синтетичних зображень у різних ракурсах, отриманих за допомогою DFM, згідно з вибраними критеріями оцінки ефективності, підвищує точність розпізнавання особистості зображення обличчя на 0,5-4% у випадку для одного зображення обличчя.

### 3.3 Порівняння результатів тестування запропонованої технології з існуючими системами розпізнавання

Окрім порівняння ефективності запропонованої технології та авторських алгоритмів компенсації ракурсу за допомогою методу кластеризації проведено порівняння результатів, отриманих за допомогою запропонованої технології підготовки зображень, зі світовими аналогами систем розпізнавання на базі зображень FERET (інформація із сайту [Anonix.com](http://Anonix.com)) [12]. На (рис. 3.17) наведено графік характеристик FAR та FRR алгоритму розпізнавання обличчя із програмного комплексу Anonix SDK, а на (рис. 3.18) – аналогічні результати розпізнавання для методу кластеризації kNN з використанням запропонованої технології

попередньої підготовки зображень облич. У таблиці 3.17 показано порівняння точності та швидкості 4 алгоритмів розпізнавання, характеристики яких наведені на сайті [Ayonix.com](http://Ayonix.com).

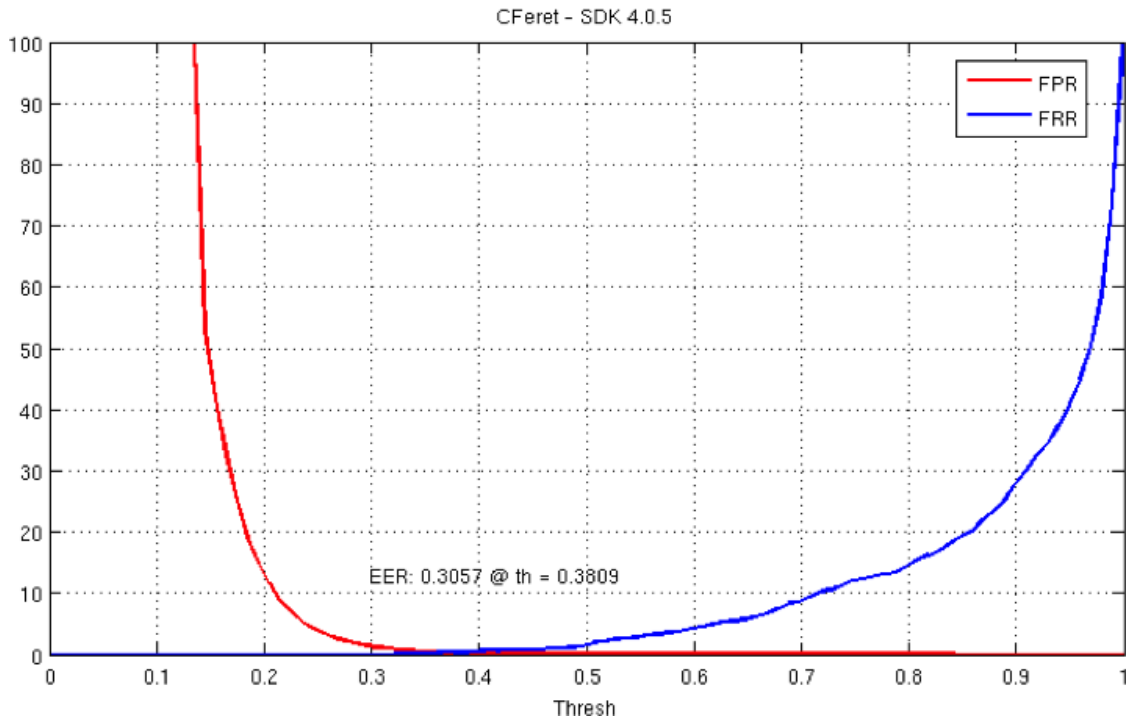


Рисунок 3.17 - Графік FAR та FRR Ayonix SDK

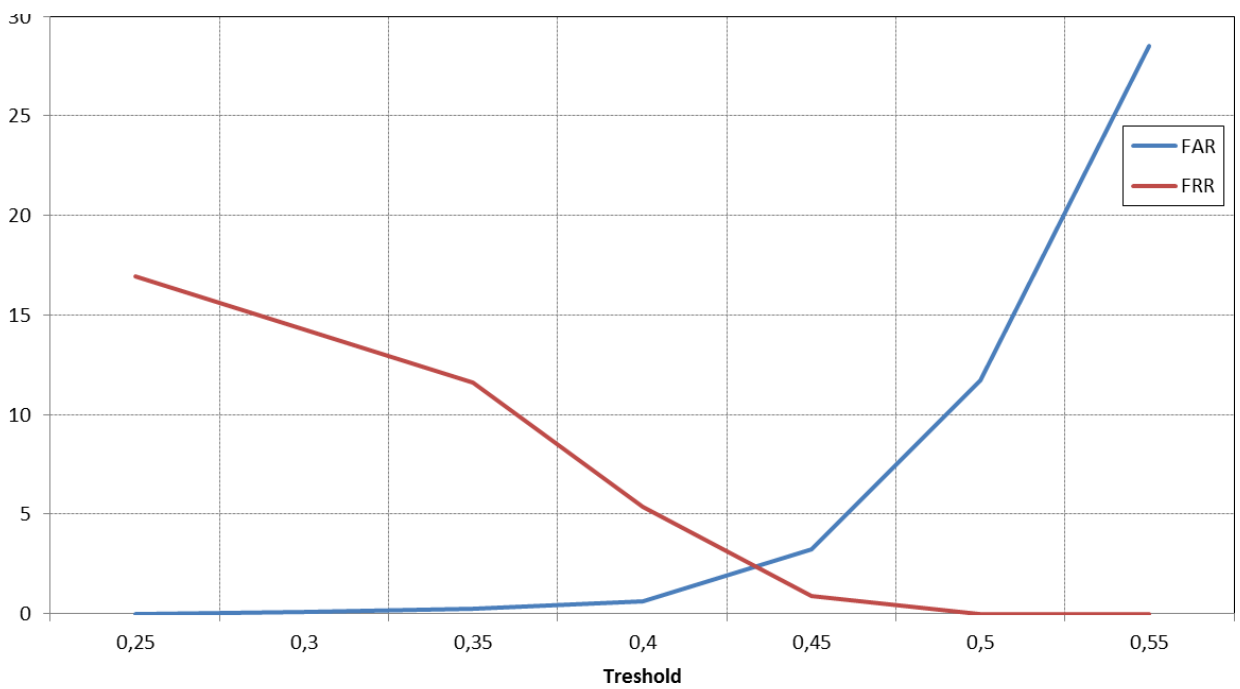


Рисунок 3.18 - Графік FAR та FRR власної реалізації

Таблиця 3.17 - Порівняння результатів розпізнавання існуючих систем та методу класифікації kNN з використанням запропонованої технології

Критерій	Німеччина	Japan Company	AyoniX v4.0 3D Release	kNN із запропонованою технологією
EER, %	97,9	97,9	99,6	97,4
Швидкість пошуку, мс	870	1200	468	152

З таблиці 3.17 очевидно, що метод класифікації kNN із застосуванням запропонованої технології попередньої підготовки зображень поступається точно провідним світовим алгоритмам розпізнавання осіб, але помітно швидше здійснює пошук за базою зображень, що підтверджує ефективність застосування технології для завдань розпізнавання осіб у режимі реального часу.

Висновки за розділом: отримано результати тестування запропонованих алгоритмів та вибраних методів у рамках технології підготовки зображень осіб до розпізнавання.

Запропонований алгоритм трекінгу успішно поєднує послідовності зображень облич у відеопотоці, збільшуючи швидкість роботи детектора облич Віоли-Джонса більш ніж на 20%.

Алгоритми оцінки контрастності та різкості не продемонстрували помітного впливу на алгоритм розпізнавання, що частково пов'язано з якістю тестових зображень осіб. Для остаточного висновку за цими методами потрібне тестування у складних умовах зйомки.

Результати експериментів з алгоритмами вирівнювання освітлення підтвердили ефективність методів DoG та Log-Gabor у підвищенні точності розпізнавання, а також високу швидкість обробки зображень за їх допомогою.

Алгоритми побудови деформованої тривимірної моделі обличчя та формування на їх основі зображень осіб у різних ракурсах показали ефективність на тестових зображеннях, а також не мали негативного ефекту в критичних ситуаціях, при порівнянні зображень осіб з кутами повороту 45 градусів та більше.

Порівняння отриманих результатів з аналогічними системами розпізнавання, чийі характеристики представлені у відкритому доступі, показало, що запропонований алгоритм попередньої підготовки зображень осіб до розпізнавання в комбінації з методом класифікації kNN демонструє дещо меншу точність розпізнавання при значному збільшенні швидкості обробки зображень облич.

Загалом можна зробити висновок, що всі розглянуті в рамках технології обробки зображень осіб алгоритми та методи ефективні на тестових даних, отже, ця технологія може бути застосована на реальних відеозаписах та при обробці відеопотоку в режимі реального часу для подальшого тестування та аналізу результатів.

## ВИСНОВКИ

У ході виконання кваліфікаційної роботи було отримано такі основні наукові та практичні результати.

Вибрано найбільш підходящі з точки зору ефективності та швидкодії методи аналізу та обробки зображень у рамках технології підготовки зображень осіб до розпізнавання особистості в режимі реального часу за результатами оцінки різних існуючих методів та алгоритмів. Зроблено аналіз існуючих алгоритмів розміщення спеціальних точок обличчя, оцінки кута повороту, а також псевдо синтезу 3D моделей. Обґрунтовано необхідність програмної модифікації існуючих алгоритмів з метою підвищення їхньої швидкості роботи та автоматизації.

Розширено можливості алгоритму трекінгу осіб у відеопотоці, що комбінує інформацію, отриману методом background subtraction та методом Віоли-Джонса, та обчислює трек особи у відеопотоці зі статичним фоном на 20 і більше відсотків швидше за алгоритм трекінгу, заснований на детекторі Віоли-Д. Розроблено оригінальну модифікацію методу синтезування DFM, яка використовує принципи відновлення псевдо ландшафтних поверхонь методом інтерполяції Шепарда, що дозволяє автоматично синтезувати рельєф обличчя за єдиним плоским зображенням. Розроблено алгоритм формування бази зображень осіб з одного фронтального зображення, що ґрунтується на запропонованій модифікації методу DFM та дозволяє застосовувати методи класифікації осіб в умовах обмеженого набору вихідних зображень, підвищуючи точність програмних систем розпізнавання особистості, які використовують методи класифікації. Запропоновано комплексну методику підготовки зображень осіб до розпізнавання особистості, засновану на авторських методах компенсації ракурсу та трекінгу осіб та розрахована на використання у програмних системах розпізнавання у відеопотоці в режимі реального часу. Розглянуті методи та алгоритми протестовано у складі комплексної технології обробки зображень осіб для програмної системи розпізнавання особистості на різних базах зображень у відкритому доступі.

## ПЕРЕДІК ДЖЕРЕЛ ПОСИЛАННЯ

1. AJ Goldstein, LD Harmon, i AB Lesk, "Identification of Human Faces", Proc. IEEE, May 1971, Vol. 59, No. 5, 748-760.
2. MA Turk and AP Pentland, "Face Recognition Using Eigenfaces", Proc. IEEE, 1991, 586-591.
3. Adini Y. Face recognition: Проблема compensation for changes in illumination direction / Y. Adini, Y. Moses, S. Ullman // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – 19. – P. 721–732.
4. Zhao W., Chellappa R., Phillips PJ, Rosenfeld A. Face recognition: A literature survey // ACM computing surveys (CSUR). 2003. vol. 35. no. 4. pp. 399-458.
5. Заварикін Д.М., Кадейшвілі А.А., Коробкова С.В., Соколов А.Ю., Степаненко О.В. Система некооперативної біометричної ідентифікації людей // Вісник МДТУ ім. н.е. Баумана. Сер. "Приладобудування", 2012. С. 40-45.
6. Zhou H., Mian A., Wei L., Creighton D., Hossny M., Nahavandi S. Recent advances на єдиній modal i multimodal face recognition: a survey // Proceedings of IEEE Transactions on Human-Machine Systems. 2014. vol. 44. no. 6. pp. 701-716.
7. Taigman Y., Yang M., Ranzato MA, Wolf L. Deepface. 2014. pp. 1701-1708.
8. Tomasi C., Kanade T. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, 1991.
9. DCRP Review: Canon PowerShot S5 IS. Режим доступу:[http://www.dcresource.com/reviews/canon/powershot\\_s5-review](http://www.dcresource.com/reviews/canon/powershot_s5-review)(Дата звернення: 25.03.2017).
10. Watada J., Musa Z., Jain LC, Fulcher J., Human tracking: status of art survey // Proceedings of the 14th International Conference on Knowledge-based and Intelligent Information and Engineering Systems, 2010, pp. 454-463.
11. Rizvi SA, Phillips PJ, Moon H. FERET verification testing protocol for face recognition algoritms //Automatic Face and Gesture Recognition, Proceedings on Third IEEE International Conference. - IEEE, 1998. - С. 48-53.

12. Wu Z., Song S., Khosla A., Yu F., Zhang L., Tang X., Xiao J. 3D Shap- eNets: A глибоке representation для volumetric shapes // Відповіді на IEEE Conference on Computer Vision and Pattern Recognition. - 2015. - С. 1912-1920.
13. Comaniciu D., Ramesh V. Robust Detection and Tracking of Human Faces with an Active Camera. // Proceedings on IEEE Visual Surveillance, 2011.
14. Bradski G., Kaehler A. Learning OpenCV: Computer vision with the OpenCV library // O'Reilly Media, Inc., 2008.
15. Zagoruiko NG, Kutnenko OA. Recognition Методи, що базуються на AdDel Algorithm // Pattern Recognition and Image Analysis. 2004. Vol. 14. №2. pp.198-204.
16. Chernyshov VG, Mestetskii LM Mobile Computer Vision System for Hand-Based Identification // Pattern Recognition and Image Analysis. 2015. Vol. 25, №2, pp. 209-214.
17. Vizilter Ю.В., Желтов С.Ю. Подібні міри і порівняння metrics для зображення shapes // Journal of Computer And Systems Sciences International. 2014. Vol. 53. pp 542-555.
18. Krylov AS, Nasonov AV 3d image sharpening by grid warping // Lecture Notes in Computer Science. 2015. Vol. 9242. pp. 441-450.
19. Bui TTT, Phan NH, Spitsyn VG, Bolotova YA, Savitsky YV Розвиток algorithms для face and character recognition based on wavelet transformations, PCA and neural networks // Proceedings of IEEE International Siberian Conference on Control and Communications (5).2015. pp. 1-6.
20. Korikov AM Neural network technology for image classification [Електронні ресурси] / AM Korikov, AV Tungusova // Proceedings of SPIE. - 2015. - Vol. 9680 : Atmospheric and Ocean Optics: Atmospheric Phys-ics.
21. Bronstein A., Bronstein M., Kimmel R. Expression-invariant 3D face recognition // Proceedings of Audio and Video-based Biometric Person Autentication (AVBPA). Література Notes in Computer Science 2688, Springer. 2003. pp. 62–69.
22. Gorodnichy D., Granger E., Radtke P. Survey з комерційних технологій для face recognition in video // CBSA, Border Technology Division, Tech. Rep. 2014. Vol.22.

23. Zhang H., Berg AC, Maire M., Malik J. SVM-KNN: Discriminative near-est neighbor classification for visual category recognition // Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. - IEEE, 2006. - T. 2. - C. 2126-2136.
24. Klehm O., Rousselle F., Papas M., Bradley D., Hery C., Bickel B., Beeler, T. Recent advances in facial appearance capture // Computer Graphics Forum. 2015. vol. 34. no. 2. pp. 709-733.
25. Song H., Lv J., Wang Y. Rapid 3D Face Modeling from Video // Pacific Rim Conference on Multimedia. 2015. pp. 201-211.
26. Garrido P., Zollhöfer M., Casas D., Valgaerts L., Varanasi K., Pérez P., Theobalt C. Reconstruction of Personalized 3D Face Rigs 3 Monocular Video // ACM Transactions on Graphics (TOG). 2016. vol. 35. no. 3. p. 28.
27. Kemelmacher-Shlizerman I., Seitz SM Face reconstruction in the wild // Computer Vision (ICCV), IEEE International Conference on 2011. - C. 1746-1753.
28. Mutch J., Lowe DG Multiclass object recognition with sparse, localized features // Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. - IEEE, 2006. - T. 1. - C. 11-18.
29. Detection використовуючи Boosted Cascade of Simple Features», процедури IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
30. Abtahi S., Omidyeganeh M., Shirmohammadi S., Hariri B. YawDD: A Yawning Detection Dataset // Proceedings of ACM Multimedia Systems, Singapore, pp. 24-28, 2014.
31. Kumar S., Yadav JS Video об'єкт extration and its tracking using background subtraction in complex environments // Perspectives in Science. - 2016. - T. 8. - C. 317-322.
32. Mandellos NA, Keramitsoglou I., Kiranoudis CT A background subtraction algoritm for detecting and tracking vehicles // Expert Systems Application, 2011. № 38 (3), pp. 1619-1631.
33. Mithun NC, Howlader T., Rahman SMM Video-based tracking of vehicles

using multiple time-spatial images // Expert Systems with Applications.

– 2016. – Т. 62. – С. 17-31.

34. Abbott RG, Williams LR Multiple target tracking with lazy background subtraction and connected components analysis // Machine Vision and Applications. - 2009. - Т. 20. - №. 2. - С. 93-101.

35. Bouwmans, T. Traditional and recent approaches in background modeling for foreground detection: overview // Computer Science Review, 2014.

№11, pp. 31-66.

36. Boda R., Priyadarsini MJP Face Detection And Tracking Using KLT And Viola Jones // ARPN Journal of Engineering and Applied Sciences, Vol. 11, No. 23, December 2016, pp. 13472-13476.

37. Mikhaylov D., Samoylov A., Minin P., Egorov A. Face Detection and Tracking from Image and Statistics Gathering // IEEE Conference on Signal-Image Technology and Internet-Based Systems, 2014.

38. Charles E. Jacobs A. Finkelstein DH Sales in Fast Multiresolution Image Querying // Department of Computer Science and Engineering, University of Washington, Seattle, Washington 98195.

39. Datta R., Joshi D., Li J., Wang.

40. Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded up robust features // Computer Vision and Image Understanding. - 2008. - V. 110. - P. 346-359.

41. Leutenegger S. BRISK: Binary Robust invariant scalable keypoints / S. Leutenegger, M. Chli, RY Siegwart // Proceedings of the 2011 International Conference on Computer Vision (ICCV '11), 6 November 2011, P. 2548-2555.

42. Kanjar D. Image Sharpness Measure for Blurred Images in Frequency Domain / D. Kanjar, V. Masilamani // International Conference on Design and Manufacturing. – Procedia Engineering, 2013. – P. 149 – 158.

43. Tan X. Додаткові Local Texture Feat. 19 № 6, June 2010. - P. 1635-1650.

44. Гонсалес Р. Цифрова обробка зображень / Р. Гонсалес, Р. Вудс // Москва: Техносфера, 2012. - 1104 с.

45. Zuiderveld K. Contrast обмежений adaptive histogram equalization //

Graphics gems IV, Academic Press Professional, Inc. San Diego, CA, USA, 1994. – P. 474-485.

46. Experimental estimation of direction finding to unmanned air vehicles algorithms efficiency by their acoustic emission, /Oleynikov, V., Zubkov, O., Kartashov, V., ...Sheiko, S., Babkin, S.//2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019 - Proceedings, 2019, стр. 175-178, 9061337

47 11. Features of acoustic noise of small unmanned aerial vehicles / Semenets, V.V., Kartashov, V.M., Leonidov, V.I. //Telecommunications and Radio Engineering (English translation of *Elektrosvyaz* and *Radiotekhnika*), 2020, 79(11), стр. 985-995

48 12. Geometric Analysis Of A Laser Scanner Functioning Based On Dynamic Triangulation /Sepulveda-Valdez, C., Sergiyenko, O., Tyrsa, V, Mercorelli, P., Kolendovska, M.// IEEE International Symposium on Industrial Electronics, 29th IEEE International Symposium on Industrial Electronics, ISIE 2020; Delft; Netherlands; 17 June 2020 до 19 June 2020; Volume 2020-June, June 2020, № 9152268, Pages 1398-1403

49 13. I. Y. A. Corpus, L.Lindner, O.Sergiyenko. "Transimpedance Amplifier for Laser Scanning System Range Extension," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1421-1426, doi: 10.1109/ISIE45063.2020.9152487.

50 14. Ivanov, M., Sergiyenko, O., Mercorelli, P., Hernandez, W.c, Rodriguez Quinonez, J.C.d, Katashov V., Kolendovska, M., Iryna, T. Effective informational entropy reduction in multi-robot systems based on real-time TVS. IEEE International Symposium on Industrial Electronics, 2019-June,8781209, c. 1162-1167.

51 15. Jonathan J. Sanchez-Castro ; Julio C. Rodríguez-Quiñonez ; Luis R. Ramírez-Hernández ; Guillermo Galaviz ; Daniel Hernández-Balbuena ; Gabriel Trujillo-Hernández ; Wendy Flores-Fuentes ; Paolo Mercorelli ; Wilmar Hernández-Perdomo ; Oleg Sergiyenko ; Félix Fernando González-Navarro. "A Lean Convolutional Neural Network for Vehicle Classification," 2020 IEEE 29th International Symposium on

Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1365-1369, doi: 10.1109/ISIE45063.2020.9152274.

52 16. Lindner, L., Sergiyenko, O., Rivas-López, M., (...), Gurko, A., Kartashov, V.M. Machine vision system for UAV navigation; IEEE, 2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles and International Transportation Electrification Conference, ESARS-ITEC, 2016; pp.1–6. DOI: 10.1109/ESARS-ITEC.2016.7841356.

53 17. M. Ivanov, O. Sergiyenko, V. Tyrsa, P. Mercorelli, V. Kartashov, W. Hernandez, S. Sheiko, M. Kolendovska. Individual scans fusion in virtual knowledge base for navigation of mobile robotic group with 3D TVS // Proceedings of 44th Annual Conference of IEEE Industrial Electronics Society (IECON).. -2018. – Washington DC, USA. -S. 3187-3192. . ISBN 978-1-5090-6683-4/18/.

54 18. Murrieta-Rico, F.N., Petranovskii, V., Galvan, D.H., Sergiyenko, O., Yocupicio-Gaxiola, R.I., De Dios Sanchez-Lopez, J. Phase effect in frequency measurements of a quartz crystal using the pulse coincidence principle. 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 185-190, 9152255, DOI: 10.1109/ISIE45063.2020.9152255

55 19. Oleksandr Sotnikov, Vladimir Kartashov, Oleksandr Tymochko, Oleg Sergiyenko, Vera Tyrsa, Paolo Mercorelli, Wendy Flores-Fuentes. Methods for Ensuring the Accuracy of Radiometric and Optoelectronic Navigation Systems of Flying Robots in a Developed Infrastructure. Chapter 16// Machine Vision and Navigation; Springer, Cham. pp.537–578. Editors: Sergiyenko, Oleg, Flores-Fuentes, Wendy, Mercorelli, Paolo. DOI: 10.1007/978-3-030-22587-2\_16.