

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Центр післядипломної освіти _____
(повна назва)

Кафедра _____ Програмної інженерії _____
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

_____ другий (магістерський) _____
(рівень вищої освіти)

Дослідження методів та алгоритмів розпізнавання образів в мобільному застосунку на пристроях Apple
(тема)

Виконав: студент 2 курсу, групи ІПЗмзд-17-1 _____
спеціальності 121- Інженерія програмного забезпечення
(код і повна назва спеціальності)

Освітньо-наукової програми
_____ Інженерія програмного забезпечення _____
(повна назва освітньої програми)

_____ Клименко А.С. _____
(прізвище, ініціали)
Керівник _____ доц. Лановий О.Ф. _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри, проф. _____

З.В.Дудар

2019 р.

Харківський національний університет радіоелектроніки

Факультет _____ Центр післядипломної освіти _____

Кафедра _____ Програмної інженерії _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 121 – Інженерія програмного забезпечення _____

(код і повна назва)

Освітньо-наукова програма _____ Інженерія програмного забезпечення _____

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри

_____ (підпис)

« _____ » _____ 2019 р.

ЗАВДАННЯ

НА АТЕСТАЦІЙНУ РОБОТУ

студентові Клименко Антону Сергійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження методів та алгоритмів розпізнавання образів в мобільному застосунку на пристроях Apple

затверджена наказом по університету від “ _____ ” _____ 20 ____ р № _____

заповнюється вручну після отримання наказу

2. Термін подання студентом роботи до екзаменаційної комісії:
“ _____ ” _____ 20 ____ р.

3. Вихідні дані до роботи теоретичні відомості про технологію розпізнавання образів, алгоритми розпізнавання за обличчям, ідентифікація за ключовими точками, приклади використання технології розпізнавання.

4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз проблемної галузі і постановка задачі, огляд методів розпізнавання образів, ідентифікація особи за обличчям, методи погодження різноракурсних зображень, пошук обличчя на зображенні та наживо, методи швидкого детектування обличчя за допомогою розроблених алгоритмів перетворення.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Мета дослідження, обґрунтування розробки, постановка задачі, об'єктна модель системи, базові моделі, методи й алгоритми, UML-діаграми, архітектура системи, демонстраційні матеріали.

6 Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта	
		підпис	дата
Спецчастина	доц. Лановий О.Ф.		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1.	Аналіз предметної галузі	11 лютого 2019р.	
2.	Огляд існуючих методів	27 квітня 2019р.	
3.	Методи розв'язання задачі	10 травня 2019р.	
4.	Підготовка пояснювальної записки	25 травня 2019р.	
5.	Спецчастина	26 травня 2019р.	
6.	Підготовка презентації та доповіді	28 травня 2019р.	
7.	Попередній захист	30 травня 2019р.	
8.	Нормоконтроль, рецензування	02 червня 2019р.	
9.	Занесення диплома в електронний архів	03 червня 2019р.	
10.	Допуск до захисту у зав. Кафедри	04 червня 2019р.	

Дата видачі завдання _____ 2019 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Лановий О.Ф.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ / ABSTRACT

Пояснювальна записка до атестаційної роботи магістра: 109 с., 49 рис., 3 додатки, 15 джерел.

SWIFT, OBJECTIVE-C, OPENCV, OPENGL, РОЗПІЗНАВАННЯ ОБРАЗІВ, ОЗНАКИ ХААРА, НОРМАЛІЗАЦІЯ

Об'єктом дослідження є процес ідентифікації особи за визначеним на зображенні обличчям людини для пристроїв Apple.

Метою роботи є вдосконалення механізмів та підвищення ефективності програмних систем для ідентифікації осіб за обличчям в режимі реального часу.

Методи дослідження базуються на методах цифрової обробки сигналів, методах теорії математичної статистики, моделюванні та дослідженні алгоритмів розпізнавання та ідентифікації особи, технологіях Swift, Objective-C, OpenGL ES, OpenCV, середовищі розробки xCode.

У результаті роботи було здійснено програмну реалізацію модифікованого алгоритму розпізнавання за обличчям в режимі реального часу на пристроях Apple.

SWIFT, OBJECTIVE-C, OPENCV, OPENGL, EXPRESSION, SIGNS OF HAAR, NORMALIZATION

The object of the research is the process of identifying a person in a person-defined image on an Apple device.

The aim of the work is to improve the mechanisms and increase the efficiency of software systems for facial identification in real time.

The research methods are based on the methods of digital signal processing, methods of mathematical theory theory, modeling and research of identification and identification algorithms, Swift, Objective-C, OpenGL ES, OpenCV, xCode development environments.

As a result of the work, the software realization of the modified face recognition algorithm on the Apple devices was implemented.

ЗМІСТ

Вступ	6
1 Характеристика предметної області та огляд існуючих рішень	8
1.1 Задачі машинного зору та сфери його застосування	8
1.2 Процес розпізнавання за обличчям	15
1.3 Аналіз існуючих програмних реалізацій	26
1.4 Постановка задачі дослідження	30
2 Моделювання системи розпізнавання за обличчям	33
2.1 Загальні проблеми розпізнавання за обличчям	33
2.2 Методи 2D та 3D розпізнавання за обличчям	36
2.3 Застосування методики розпізнавання за обличчям	38
2.4 Методі від Apple	52
3 Програмна реалізація системи	55
3.1 Опис програмних засобів та бібліотек	55
3.2 UML модель системи	59
3.3 Приклади алгоритмів	63
3.4 Опис програмних рішень	66
3.5 Аналіз отриманих результатів	71
Висновки	75
Список використаних джерел	76
Додаток А. Програмний код	78
Додаток Б. Результати тестових експериментів	92
Додаток В. Слайди презентації	100

ВСТУП

Розвиток інформаційних технологій та їх використання в усіх сферах людського життя є результатом технічного прогресу. Однією з ключових інформаційних технологій, що за останні роки починають відігравати ключову роль, стає розпізнавання образів. Розпізнавання образів – це віднесення вихідних даних до певного класу за допомогою виділення істотних ознак, що характеризують ці дані, із загальної маси несуттєвих даних. При цьому це може бути представлено будь-яким образом: від дерева до особи людини. Саме на розпізнанні певного образу з наступним процесінгом базується технологія віртуальної реальності – створений за допомогою технічних засобів світ, який передається людині через його відчуття. Віртуальна реальність – основна і найбільш важлива складова цього концепту, що поєднує властивості розпізнаного образу з віртуальним. І якщо задача розпізнавання людини за обличчям досить глибоко вивчена та широко використовується на різних програмно-апаратних платформах, то задача виконання антропометричних вимірювань, визначення окремих частин тіла людини та їх розпізнавання з використанням мобільних технологій ще детально не вивчена.

Розпізнавання обличчя – найбільш поширений спосіб ідентифікації, заснований на тому, що риси обличчя і форма черепа кожної людини індивідуальні. Комп'ютер лише автоматизує процедуру, виконуючи аналогічну процедуру, з тією різницею, що замість фото застосовуються біометричні дані, записані в еталонному образі. Це самий інтуїтивно зрозумілий метод ідентифікації, найбільш близький до того, як люди ідентифікують один одного.

Розпізнавання образів значно спрощує взаємодію людини з комп'ютером, створює передумови для застосування різних систем штучного інтелекту. Багато операцій пов'язаних з процесами автентифікації та ідентифікації можливо прискорити за допомогою використання комп'ютерних або мобільних систем розпізнавання образів. В останні роки розпізнавання образів знаходить все більше

застосування в різних напрямках: охоронні системи, криміналістика, комп'ютерна графіка. Саме тому розробка нових ефективних методів розпізнавання образів становить інтерес для вирішення широкого спектра задач, що виникають в техніці, медицині, біометрії. Так застосування елементів розпізнавання образів під час проведення ділових зустрічей дозволяє зменшити ризики прийняття економічно невірних рішень. Нарешті, останнім часом набуває все більшої популярності інтернет-торгівля, і в цій сфері розпізнавання образів має великі перспективи, оскільки дозволяє віртуально визначити, чи личить той чи інший вид товару саме цій людині – наприклад, окуляри, прикраси, капелюх тощо.

Розпізнавання за рисами обличчя має ряд переваг перед іншими біометричними технологіями:

- немає потреби в безпосередньому контакті людини, обличчя якої встановлюють, зі сканером, за винятком систем розпізнавання за обличчям у складі стандартних електронних охоронних систем, де людина при верифікації дивиться прямо в камеру;

- при відповідному обладнанні розпізнавання за рисами обличчя можливо на значній відстані, в групі людей, не привертаючи до цього уваги;

- це єдиний біометричний спосіб ідентифікації з точки зору можливості багатоцільового застосування, що не вимагає спеціальної техніки;

- при ідентифікації використовується загальнодоступні біометричні характеристики, що зазвичай не приховуються людиною.

Оскільки об'єктом дослідження є конкретні методи розпізнавання образів, в роботі проведено аналітичний огляд цих методів та їх порівняння з іншими відомими методами вирішення задачі розпізнавання. Результатами є алгоритмічне забезпечення та програмні технології для розв'язання деяких з перерахованих вище завдань розпізнавання образів. На основі запропонованих моделей, методик та алгоритмів здійснено розробку мобільного додатку iOS для використання під час проведення бізнес-нарад.

1 ХАРАКТЕРИСТИКА ПРЕДМЕТНОЇ ОБЛАСТІ ТА ОГЛЯД ІСНУЮЧИХ РІШЕНЬ

1.1 Задачі машинного зору та сфери його застосування

Машинний зір – це міждисциплінарна область, що отримала на теперішній час широкий розвиток. Методи обробки зображень стали важливими розділами комп'ютерного зору і засновані на комбінації багатьох дисциплін. Прогрес в сфері машинного зору визначається розвитком математичних методів (математичної статистики, теорії ймовірності, методів оптимізації, методів розв'язання алгебраїчних, диференціальних та інтегральних рівнянь), а також розвитком обчислювальних ресурсів. Довгий час теоретичні дослідження в області машинного зору випереджали обчислювальні можливості ЕОМ, що ускладнювало їх використання для вирішення практичних завдань.

Машинний зір є дійсно інноваційно привабливим. Інтерес до нього виник ще на початку спроб створення штучного інтелекту. На теперішній час кількість нових рішень і актуальних додатків для розв'язання задач машинного зору продовжує зростати.

Значимість проблеми машинного зору ніколи не викликала сумніву, але одночасно її складність істотно недооцінювалася. Наприклад, легендарним по своїй показовості став випадок, коли в 1966 р. один з основоположників штучного інтелекту, Марвін Мінський, навіть не сам зібрався вирішити проблему штучного зору, а доручив це зробити одному із студентів за найближче літо. При цьому на створення програми, що грає на рівні гросмейстера в шахи, відводився значно більший час. Однак зараз очевидно, що створити програму, яка буде обігравати людину в шахи, значно простіше, ніж створити адаптивну систему управління з підсистемою машинного зору, яка б змогла просто переставляти шахові фігури на реальної дошці.

Машинний зір дає важливу інформацію для створення систем штучного інтелекту. Такі системи можуть отримувати інформацію як із отриманих

зображень, так і з наборів багатовимірних даних різної природи. Підставами для цього стали досягнення в побудові математичного апарату машинного зору, головними етапами розвитку якого стали:

- до 1970-х років сформувався основний понятійний апарат в області обробки зображень, що є основою для дослідження проблем зору. Також були виділені основні завдання, специфічні для машинного зору, пов'язані з оцінкою фізичних параметрів сцени (дальності, швидкостей руху, відбивної здатності поверхонь тощо);

- до 80-х років сформувалася теорія рівнів представлення зображень в методах їх аналізу. Свого роду відміткою закінчення цього етапу служить книга Девіда Марра «Зір. Інформаційний підхід до вивчення подання та обробки зорових образів»;

- у 90-х виявляється сформованим систематичне уявлення про підходи до вирішення основних задач машинного зору, що вже стали класичними;

- з середини 90-х відбувається перехід до створення і дослідження великомасштабних систем комп'ютерного зору, призначених для роботи в різних природних умовах;

- з початку 2000-х стрімко розвиваються методи автоматичної побудови уявлень зображень в системах розпізнавання зображень і комп'ютерного зору на основі принципів машинного навчання.

У той же час прикладні застосування обмежувалися обчислювальними ресурсами. Адже для виконання навіть найпростішої обробки зображення необхідно хоча б один раз переглянути всі його пікселі (а зазвичай – і не один раз). Для цього потрібно виконувати як мінімум сотні тисяч операцій в секунду, що довгий час було неможливо і вимагало спрощень. Наприклад, для автоматичного розпізнавання деталей в промисловості могла використовуватися чорна стрічка конвеєра, що усуває необхідність відділення об'єкта від фону, або сканування об'єктів, що рухаються лінійкою фотодіодів зі спеціальною підсвіткою, що вже на рівні формування сигналу забезпечувало виділення інваріантних ознак для розпізнавання без застосування будь-яких складних

методів аналізу інформації. У оптико-електронних системах супроводження і розпізнавання цілей використовувалися фізичні трафарети, що дозволяють «апаратно» виконувати узгоджену фільтрацію. Деякі з цих рішень були геніальними з інженерної точки зору, але були застосовані тільки в задачах з низькою апріорної невизначеністю, і тому мали, зокрема, поганою переносимістю на нові завдання.

Тож не дивно, що на 1970-ті роки припав пік інтересу і до оптичних обчислень в обробці зображень. Вони дозволяли реалізувати невеликий набір методів (переважно кореляційних) з обмеженими властивостями інваріантності, але вельми ефективним чином. Поступово, завдяки зростанню продуктивності процесорів (а також розвитку цифрових відеокамер), ситуація змінилася. Подолання певного порогу продуктивності, необхідного для здійснення корисної обробки зображень за розумний час, відкрило шлях для цілої лавини додатків комп'ютерного зору. Слід, однак, відразу підкреслити, що цей перехід не був миттєвим і триває досі.

Однак дійсно масове застосування методи комп'ютерного зору отримали лише менше десяти років тому, з досягненням відповідного рівня продуктивності процесорів у персональних і мобільних комп'ютерів. Таким чином, в плані практичного застосування системи комп'ютерного зору пройшли ряд етапів:

- етап індивідуального рішення (як в частині апаратного забезпечення, так і алгоритмів) конкретних завдань;
- етап застосування в професійних областях (особливо в промисловості і оборонній сфері) з використанням спеціалізованих процесорів, спеціалізовані системи формування зображень і алгоритми, призначені для роботи в умовах низької апріорної невизначеності, однак ці рішення допускали масштабування;
- етап масового застосування.

Машинний зір дозволяє вирішувати значну кількість задач, які умовно можна розділити на чотири групи (див. рис. 1.1).



Рисунок 1.1 – Групи задачі систем машинного зору

Розпізнавання положення. Мета машинного зору в даному застосуванні – визначення просторового розташування (розташування об'єкта щодо зовнішньої системи координат) або статичного положення об'єкту (в якому стані знаходиться об'єкт щодо системи координат з початком відліку в межах самого об'єкту) і передача інформації про стан і орієнтацію об'єкта в систему управління або контролер. Прикладом такого додатка може служити вантажно-розвантажувальний робот, перед яким стоїть задача переміщення об'єктів різної форми з бункера. Інтелектуальна задача машинного зору полягає, наприклад, у визначенні оптимальної базової системи координат та її центру для локалізації центру ваги деталі. Отримана інформація дозволяє роботу захопити деталь належним чином і перемістити її в належне місце.

Вимірювання. У додатках даного типу основне завдання відеокамери полягає у вимірюванні різних фізичних параметрів об'єкту. Прикладом фізичних параметрів може служити лінійні розміри, діаметр, кривизна, площа, висота, кількість тощо. Приклад реалізації цієї задачі – вимірювання різних діаметрів горлечка скляної пляшки.

Інспекція. У додатках, пов'язаних з інспекцією, мета машинного зору – підтвердити певні властивості, наприклад, наявність або відсутність етикетки на пляшці, болтів для проведення операції зборки, шоколадних цукерок в коробці або наявність різних дефектів.

Ідентифікація. В задачах ідентифікації основне призначення відеокамери – зчитування різних кодів (штрих-кодів, 2D-кодів тощо) з метою їх розпізнавання засобами камери або системним контролером, а також визначення різних буквено-цифрових позначень. Крім того, до задач даної групи можна віднести системи, які вирішують задачі підтримки безпеки, такі як ідентифікація особи і техніки, детектори руху тощо.

Система машинного зору включає наступні основні компоненти (рис.1.2):

- підсистему формування зображень (яка сама може включати різні компоненти, наприклад об'єктив та ПЗС- або КМОП-матрицю);
- обчислювальний елемент (процесор);
- блок обробки зображень, що включає алгоритми аналізу зображень, які можуть бути реалізовані програмно на процесорах загального призначення, апаратно в структурі обчислювального елементу і навіть апаратно в рамках підсистеми формування зображень.

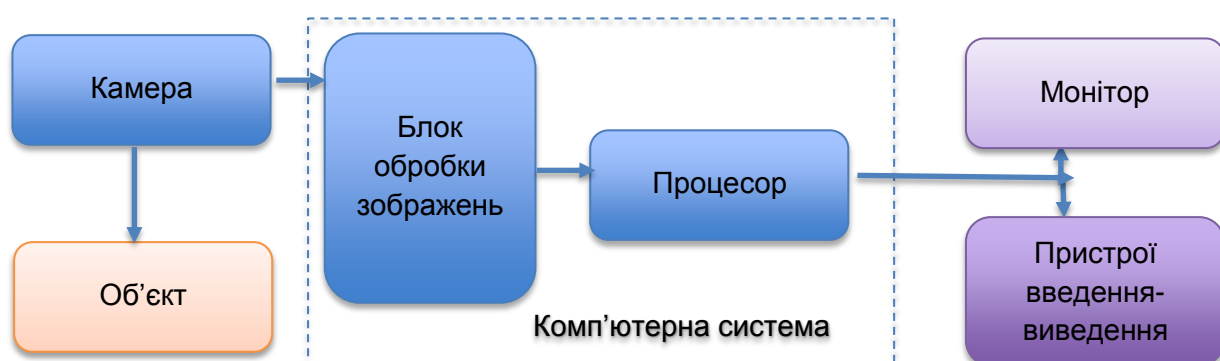


Рисунок 1.2 – Комп'ютерна система машинного зору

Масового застосування досягають системи машинного зору, які використовують стандартні камери і комп'ютери в якості перших двох компонент

(саме для таких систем більше підходить термін «комп'ютерний зір» (КЗ), хоча чіткого поділу між поняттями машинного і комп'ютерного зору немає). Однак інші системи машинного зору мають не меншу значимість. Саме вибір «нестандартних» способів формування зображень (включаючи використання інших, крім видимого, спектральних діапазонів, когерентного випромінювання, структурованої підсвічування, гіперспектральних приладів, швидкодіючих камер, телескопів і мікроскопів тощо) істотно розширює можливості систем машинного зору. У той час як за можливостями алгоритмічного забезпечення системи машинного зору істотно поступаються зору людини, за можливостями отримання інформації про об'єкти, за якими ведеться спостереження, вони істотно перевершують його.

Подальше підвищення продуктивності мобільних процесорів дозволяє ставити нові завдання для систем комп'ютерного зору в побутових роботах, число продажів яких по всьому світу обчислюється вже мільйонами примірників на рік. Крім завдань навігації, призначених для персонального використання, інтерес викликає рішення задач розпізнавання людей та їх емоцій по особам, розпізнавання жестів, предметів обстановки, включаючи столові прилади і посуд, одяг, домашніх тварин тощо.

Розпізнавання особи за обличчям — завдання, з яким люди справляються надзвичайно легко та швидко. Ця очевидна простота виявилася небезпечно оманливою, адже завдання автоматичного розпізнавання обличчя виявилось проблемою, що досі є далекою від остаточного розв'язання. Незважаючи на інтенсивні дослідження, велику кількість опублікованих статей та конференцій, які були присвячені цій тематиці, досі неможливо заявити, що штучні системи можуть «позмагатися» з людськими показниками.

Автоматичне розпізнавання обличчя є «заплутаним» завданням в першу чергу через складні вимоги до зображень (освітлення та зміна позиції і повороту лица під час руху людини), а також через інші різноманітні фактори, такі як старіння, вираз обличчя, оклюзії тощо. Дослідники з комп'ютерного зору, аналізу зображень та їх обробки, розпізнаванню паттернів та іншими сферам працюють

разом над вирішенням цього питання, вони змотивовані великою кількістю можливих практичних застосувань технології розпізнавання обличчя.

Розпізнавання обличчя стало одним з трьох методів ідентифікації, що використовуються у електронних паспортах, а також є вибором для багатьох програм із забезпечення безпеки. З-поміж шести біометричних атрибутів (рис. 1.3) ознаки обличчя мають найкращу сумісність у системі машинозчитувальних проїзних документів (Machine Readable Travel Documents – MRTD), і це базується на оцінних факторах — реєстрації, оновленні, вимогах до машини, і суспільному сприйнятті [2].



Рисунок 1.3 – Сценарій застосування MRTD-системи при паспортному контролі (зліва), і порівняння різних біометричних ознак щодо MRTD-сумісності (справа)

Системи розпізнавання за обличчям повинні «знаходити» обличчя на зображеннях та відео автоматично. Вони можуть функціонувати у одному з двох режимів: (1) – верифікація обличчя (або автентифікація), та (2) – ідентифікація обличчя (або розпізнавання). Верифікація обличчя включає в себе порівняння один до одного, тобто обличчя із зображення з шаблоном обличчя тієї людини, яка підлягає верифікації.

Ідентифікація за обличчям включає в себе порівняння один до багатьох — обличчя із зображення до усіх шаблонів облич в базі даних для визначення особи, яку підозрюють у присутності на зображенні. Інший сценарій розпізнавання за

обличчям включає в себе порівняння один до декількох — обличчя із зображення порівнюється з декількома шаблонами облич тих людей, що обрані користувачем (наприклад, полісмен обирає підозрюваних для визначення злочинця).

Продуктивність систем розпізнавання за обличчям значно покращилась з того часу, як було розроблено першу автоматичну систему розпізнавання. Крім того, виявлення обличчя на зображенні, виділення ознак обличчя та їх розпізнавання можуть зараз бути виконані у «realtime» для тих зображень, що були отримані з дотриманням бажаних (тобто, доволі суворих) вимог.

1.2 Процес розпізнавання за обличчям

Розпізнавання за обличчям — це одна із задач розпізнавання візуальних паттернів. Обличчя являє собою тривимірний об'єкт, що може піддаватися різним рівням освітлення, змінювати позу, вираз і таке інше, тож ідентифікація за обличчям базується на двовимірному зображенні (хоча в технічних системах досить часто використовуються тривимірні зображення, які отримані з використанням лазерів та 3D камер).

Система розпізнавання образів загалом складається з чотирьох модулів (див. рис. 1.4): виявлення, оцінка положення, виділення ознак та співставлення, де локалізація та нормалізація (виявлення обличчя та його положення) — це етапи попередньої обробки, що виконуються безпосередньо перед тим, як проводити розпізнавання за обличчям (виділення ознак обличчя та співставлення його з шаблонами бази даних).

Визначення обличчя відмежовує область обличчя від фону. У випадку використання відеоряду для виявлення обличчя можуть бути застосовані спеціальні компоненти трекінгу. Оцінка положення обличчя має за мету досягнення більш точної локалізації та нормалізації обличчя, в той час, як виявлення обличчя забезпечує лише грубу оцінку місця розташування та масштаб

кожного з виявлених обличчя. Виявляються окремі компоненти обличчя, такі як ніс, очі, рот та контур обличчя; на основі точок місцезнаходження вхідне зображення обличчя нормалізується у відношенні до геометричних властивостей, таких як розмір та поза, при цьому використовуються геометричні перетворення або морфінг.

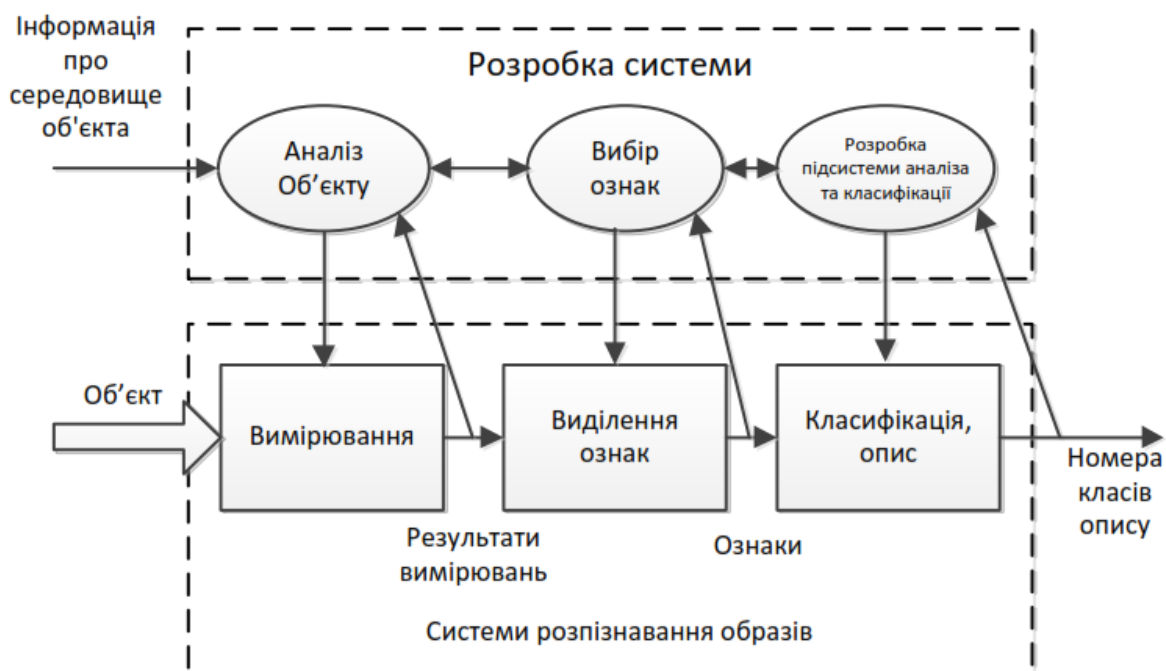


Рисунок 1.4 – Схема процесу розпізнавання

Обличчя зазвичай і далі нормалізується по відношенню до фотометричних властивостей, таких як освітлення та відтінки сірого.

Після того, як обличчя нормалізовано геометрично та фотометрично, проводиться виділення ознак обличчя для забезпечення ефективної інформації для того, щоб відрізнити обличчя різних людей одне від одного, причому ця інформація повинна бути настільки точною, щоб відрізнити обличчя можна було навіть при різному наборі геометричних та фотометричних даних.

Для співставлення за обличчям виділений вектор ознак вхідного обличчя порівнюється з векторами обличчя з бази даних; система виводить результат (розпізнавану особу) у тому випадку, якщо співпадіння виявлено при

задовільному порозі точності, або ж виводить інформацію, що обличчя належить невідомій людині.

Результати розпізнавання за обличчям залежать від виділених ознак, що представляють шаблон обличчя, та методів класифікації, які використовуються для порівняння одного обличчя з іншим, в той час, як локалізація та нормалізація обличчя є базисом для виділення ефективних ознак.

Задача розпізнавання осіб в кадрі за обличчям можна поділити на два кроки. Перший крок – це виявлення обличчя в кадрі. Другий крок – розпізнавання виявленого обличчя. Виходячи з цього розглянемо окремо існуючі методи виявлення і розпізнавання осіб за обличчям.

1.2.1 Алгоритми виявлення обличчя

Задача виявлення особи в більшості випадків є «першим кроком» для розв'язання задач більш високого рівня. Сама інформація про присутність і кількості осіб в кадрі може бути корисна для охоронних систем та інших програм. З усього розмаїття існуючих алгоритмів виявлення розглянемо найактуальніші, а також проаналізуємо їх недоліки, переваги та відмінні риси кожного.

1.2.1.1 Sparse Network of winnows (SNoW)

Алгоритм виявлення особи в кадрі, який являє собою двошарову мережу. Вхідний шар складається з вузлів, кожен такий вузол є деякою характеристикою вхідного зображення (при наявності особливості генерується одиниця, у випадках її відсутності на зображенні – нуль). Вихідний шар складається з двох вузлів, один відповідає розпізнається класу зображення «обличчя», а інший – «не

обличчя» [7]. У цьому алгоритмі використовуються ознаки Successive Mean Quantization Transform (SMQT). Це перетворення дозволяє «витягувати» з локальної області зображення складову, що не залежить від освітленості. Воно полягає в квантуванні області зображення з порогом квантування рівним середньому значенню пікселів, що входять в цю область [8].

Перевагами цього алгоритму є:

- досить висока точність виявлення обличчя;
- швидкість роботи за рахунок просіювання компонент вектору ознак;
- нечутливість до змін освітлення.

До недоліків слід віднести:

- чутливий до шумів і викидів даних.

1.2.1.2 Метод опорних векторів

Метод опорних векторів (Support Vector Machines, МОВ) використовується для вирішення задач класифікації та регресійного аналізу. Суть методу полягає в необхідності знаходження гіперплощини, яка поділяється на два класи. При виборі з множини гіперплощини необхідно вибирати саме ту, відстань до якої має найбільшу відстань для кожного класу. Така гіперплощина називається оптимальною гіперплощиною, а лінійний класифікатор, які їй відповідає, має назву оптимального класифікатора [8].

Переваги:

- висока швидкість роботи щодо нейронних мереж;
- є можливість зменшити чутливість до шуму завдяки зниженню точності, але пошук оптимального співвідношення вимагає точного налаштування;
- висока стійкість до перенавчання.

Недоліки:

- точність гірша за деякі інші методи.

1.2.1.3 Adaptive Boosting (AdaBoost)

Алгоритм, за допомогою якого проводиться посилення класифікаторів шляхом об'єднання їх в клас.

Вперше цей алгоритм було запропоновано Робертом Шапіро та Йоав Фройндом в 1999 році [8]. Він може використовуватись разом з декількома іншими алгоритмами класифікації для покращення їх ефективності.

Кожен наступний клас класифікаторів вибудовується за об'єктами, які були невірно класифіковані попередніми класами. Дуже часто цей алгоритм використовується в поєднанні з іншими алгоритмами класифікації для їх поліпшення.

Алгоритм AdaBoost побудований на основі алгоритмів навчання (наприклад, дерева рішень) і об'єднання їх. Мета полягає в тому, щоб створити один сильний класифікатор на основі декількох слабких класифікаторів. Алгоритм AdaBoost успішно використовується для пошуку осіб на статичних зображеннях, сайти постійно підтримують пошук інформації на основі зображень з використанням саме AdaBoost. При використанні цього алгоритму може застосовуватись довільна кількість класифікаторів та проводиться навчання на одному наборі зразків, по черзі застосовуючи їх на різних кроках алгоритму.

Переваги:

- простота реалізація;
- швидкість роботи;
- підлаштовується під проблемні об'єкти вибірки.

Недоліки:

- досить тривале навчання, тривалість якого залежить від кількості класифікаторів і розміру вибірки;
- чутливість до шумів і викидів даних.

1.2.1.4 Нейромережевий підхід

Штучні нейронні мережі успішно використовувалися для вирішення задач класифікації об'єктів починаючи з середини минулого століття. З точки зору машинного навчання штучні нейронні мережі є обчислювальними моделями, які здатні вирішувати завдання розпізнавання образів. Нейронну мережеву архітектуру можна розділити на дві основні групи: мережі прямого зв'язку і зворотного поширення мережі [3]. При вирішенні задачі класифікації віку і статі людини на основі аналізу зображень застосовується велика кількість нейронних мереж різних архітектур. Зокрема, використовуються імовірнісні нейронні мережі (probabilistic decision-based neural networks), багатоваршівні перцептрони та інші. Штучні нейронні мережі є одним з пріоритетних напрямків досліджень в області машинного навчання [4].

Принципи, що лежать в основі, – це послідовне перетворення сигналу паралельно працюючими функціональними елементами – нейронами. Процес навчання цих мереж зводиться до мінімізації середньоквадратичної помилки. Слід зазначити, що системи виявлення об'єктів на кадрах, що будуються на нейромережевих методах, мають чітко визначену ієрархічну структуру. На першому кроці вектор ознак обробляється досить грубою мережею із значним рівнем помилок, а потім, якщо вектор не класифікований, рішення поправляється більш точною та повільною мережею [11].

Переваги:

– досить висока точність виявлення.

Недоліки:

– сприйнятливості до шумів;

– необхідно дуже точно налаштовувати параметри нейронної мережі для отримання якісних результатів;

– повільна швидкість роботи через високу обчислювальну складність;

– схильність до перенавчання.

1.2.1.5 Метод Віола-Джонса

Цей метод було запропоновано Віолою і Джонсоном в 2001 році [6], він став безумовним проривом в області розпізнавання осіб за зображенням обличчя.

Принцип дії цього алгоритму: прямокутна область, за розміром менша, ніж зображення, пересувається з деяким кроком, і завдяки каскаду класифікаторів може визначити, чи знаходиться в зображенні обличчя. Алгоритм Віола-Джонса високоефективно використовується в задачах розпізнавання об'єктів і комп'ютерного зору.

Переваги:

- розпізнавання декількох осіб;
- один з найпопулярніших і один з найпоширених алгоритмів розпізнавання осіб;
- завдяки наявності класифікаторів – висока швидкість навчання;
- малий відсоток помилкових спрацьовувань, за рахунок чого досягається висока точність.

Недоліки:

- необхідність великої вибірки, а також багато часу для навчання;
- погано працює при неправильному положенні особи.

1.2.2 Алгоритми розпізнавання обличчя на зображеннях

Алгоритми розпізнавання осіб за обличчями можна поділити на дві підгрупи: перша – це методи, які засновані на значеннях пікселів, а друга – методи, що базуються на характеристичних точках [12].

1.2.2.1 Алгоритми, що базуються на значеннях яскравості пікселів

Виходячи з назви вищевказаної групи методів, з метою розпізнавання виявленого обличчя застосовується виключно колір або яскравість пікселів на зображеннях. Найбільш оптимальним способом прийнято вважати порівняння. В такому випадку мірою схожості виступає відстань між векторами яркостей пікселів зображень. Разом з тим, зазначений метод абсолютно непристосований до змін освітлення і положення особи, а також – до масштабування. Додатковим недоліком є значна обчислювальна складність, що робить технологію неефективною для розпізнавання в реальному часі. З цієї причини найчастіше перевагу віддають методам, що застосовують перехід векторного опису зображень в просторову форму з меншим показником розмірності. В результаті порівняння буде більш ефективним [12].

На теперішній час існує чотири основні методи розпізнавання особи, які розрізняються складністю реалізації та метою застосування [7]:

- метод автоматичної обробки зображення особи;
- «EigenFaces» (нім. «власне обличчя»);
- аналіз відмінних рис;
- аналіз на основі нейронних мереж.

Метод автоматичної обробки зображення обличчя – найбільш проста технологія, що аналізує відстані і відношення відстаней між легко визначеними точками обличчя. Особливо важливі характерні частини обличчя, а також ті, які практично не змінюються з плином часу: очі, вилиці, кінець носа, куточки рота (див.рис.1.5). Хоча даний метод не дуже потужний, він може бути досить ефективно використаний в умовах слабкої освітленості.

Технологія «EigenFaces» використовує представлення зображення обличчя в градаціях сірого у вигляді статистично обґрунтованих, стандартних блоків даних (областей обличчя). Даний метод базується на тому, що всі обличчя можуть бути отримані з репрезентативної вибірки облич з використанням сучасних

статистичних прийомів. Вони охоплюють пікселі зображення обличчя і універсально представляють форми обличчя (двомірні зображення-шаблони). Комбінуючи 100-120 різних шаблонів, можна представити велику кількість облич.

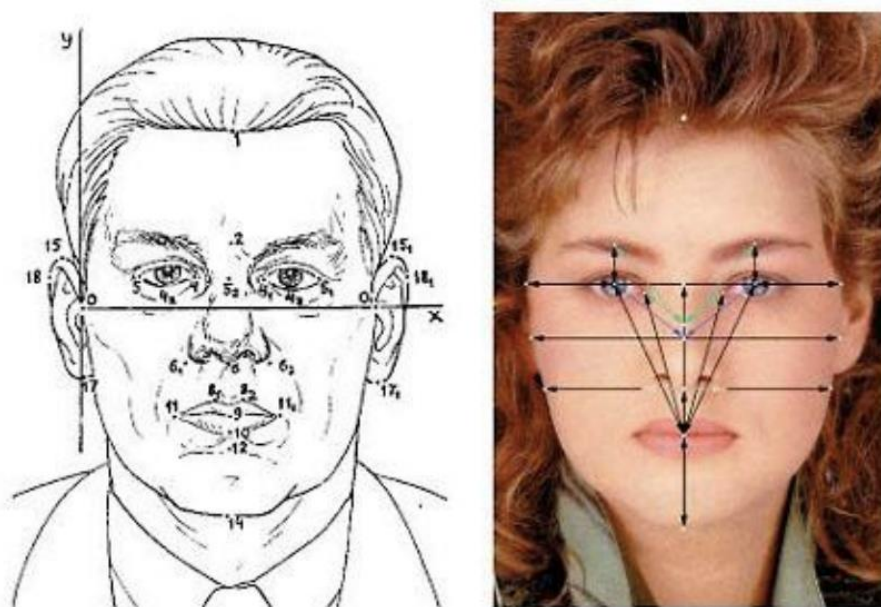


Рисунок 1.5 – Визначення ключових точок обличчя

При реєстрації вигляд кожної конкретної людини представляється рядом коефіцієнтів, що вказують найбільш відповідні шаблони. Для режиму встановлення автентичності, коли проводиться перевірка ідентичності, біометричний образ користувача обробляється і порівнюється з раніше зареєстрованим набором коефіцієнтів, з метою визначення коефіцієнта відмінності. Ступінь відмінності між шаблонами і визначає факт ідентифікації. Технологія «eigenface» оптимальна в добре освітлених приміщеннях, при можливості сканування особи в фас. Метод використовується в якості основи для інших методів розпізнавання особи.

Методика аналізу відмінних рис подібна методиці «eigenface», але в більшій мірі адаптована до зміни зовнішності або міміки людини. У технології аналізу відмінних рис використовуються не тільки характерні особливості областей обличчя, а й враховано їх відносне положення. Тобто ідентичність обличчя визначається не тільки характерними елементами, але і способом їх

геометричного об'єднання. Індивідуальна комбінація цих параметрів визначає особливості кожного конкретного обличчя.

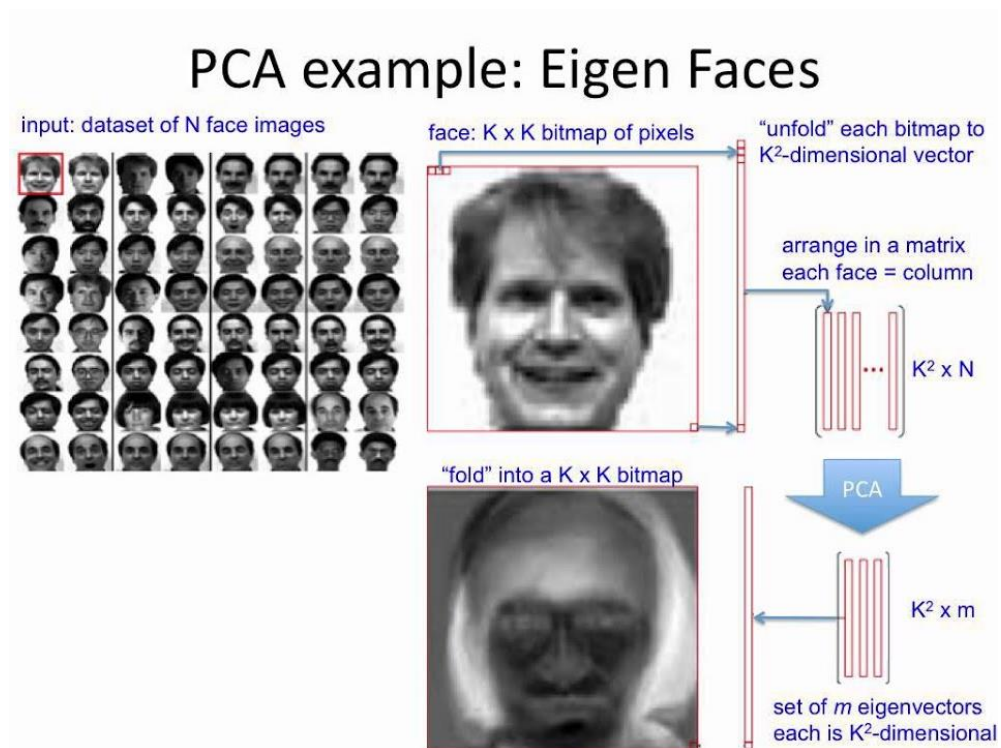


Рисунок 1.6 – Принцип дії алгоритму EigenFaces

Для порівняння з графічними зображеннями-шаблонами застосовуються два основних алгоритми: мінімальної середньої кореляційної енергії (MACE) [6] і локальні бінарні шаблони (LBP) [9].

Локальні бінарні шаблони (LBP) використовують обробку пікселя цифрового зображення. Алгоритм LBP популярний для розпізнавання графічного зображення в цілому, а останнім часом застосовується і для розпізнавання за обличчям. Непараметричне ядро LBP аналізує піксельну структуру зображень. Воно є інваріантним до монотонних сіро-масштабних перетворень, тобто менш чутливе до освітленості, що має важливе значення.

Принцип роботи MACE-фільтру базується на визначенні середнього ступеня кореляції до заздалегідь підготовлених зображень; коефіцієнт кореляції дорівнює нулю на всьому зображенні крім областей, які збігаються з шаблонами, тобто в цих областях ступінь кореляції більше. Для роботи необхідна база шаблонів для розрахунку ступеня кореляції. Для забезпечення більшої надійності

в базі потрібно мати порівняно велику кількість зображень обличчя, в різних умовах освітлення і зміни міміки.

1.2.2.2 Алгоритми, що базуються на характерних точках

Функціонування цих способів базується на характерних точках та їх розташуванню на зображенні. Такого роду елементами є: центри очей, лінія брів, рота та інше [8]. До вказаної групи також слід віднести активні моделі зовнішнього вигляду і форми.

Активні моделі зовнішнього вигляду (Active Appearance Models, AAM) являють собою зображення, модернізовані під оригінал. Їх структура містить в собі параметри форми і зовнішнього вигляду. Головним недоліком слід визначити необхідність попередньо протестувати модель на великій кількості вже розмічених вручну зображень.

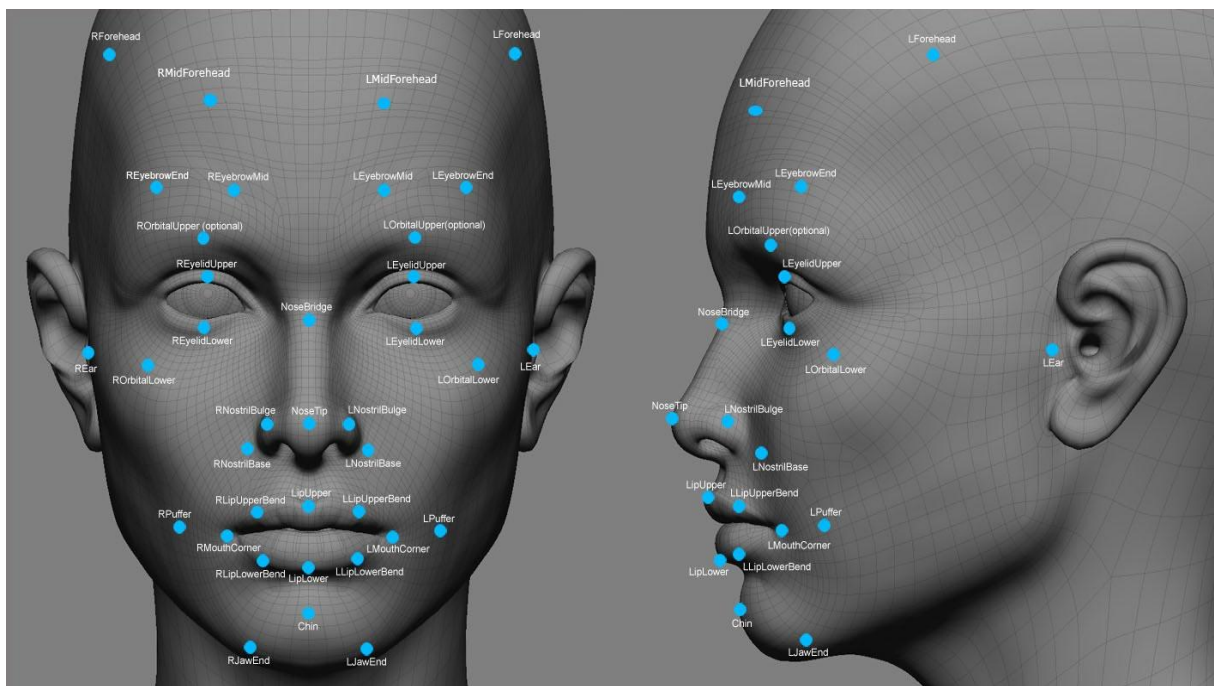


Рисунок 1.7 – Визначення ключових точок на обличчі

Активні моделі форми (Active Shape Models, ASM) функціонують з урахуванням взаємозв'язків щодо позиції антропометричних точок. На експерта покладається розмітка їх позиції на кожному із зображень. За допомогою так званого «Прокрустова аналізу» вдається звести їх координати в єдине ціле. Потім необхідно обчислити середню форму і матрицю коваріації для групи образів. Наступним кроком є обчислення власних векторів, які розподіляються по спадаючій. Локалізації ASM моделі на новому зображенні виконується в ході пошуку рішення оптимізаційної задачі [10].

Слід врахувати, що моделі вищевказаного типу спочатку розроблялися з метою виявлення детального розташування відмінних точок на зображеннях осіб, а не були створені з метою їх розпізнавання. Їх розміщення сприяє оперативному проведенню дій, спрямованих на вирівнювання осіб і можливість упорядкувати єдину систему координат для ефективного розпізнавання за допомогою інших технологій. Як правило, для реалізації такої мети достатньо мінімальної кількості точок, що сприяє збільшенню продуктивності роботи алгоритму. У свою чергу, задачі розпізнавання за обличчям вимагають значної кількості точок, наявність яких зробить класифікацію більш точною, що, у свою чергу, знизить швидкість функціонування системи [9].

Виходячи з вищевказаного варто зазначити, що найбільш оптимальним методом з максимальним показником ефективності розпізнавання осіб за обличчям на етапі обробки в режимі реального часу є методи Eigenfaces і Fisherfaces.

1.3 Аналіз існуючих програмних реалізацій

Основними і пов'язаними між собою етапами зі створення математичної моделі розпізнавання осіб за обличчям є:

- визначення обличчя на кадрі;

– розуміння програмною системою, що при наявності неприродного повороту особи або недостатньої освітленості – людина на кадрі є тією ж самою особою, що і в базі даних;

– виділення унікальних характеристик особи, які дозволяють відрізнити одне обличчя від інших: наприклад, розріз і форма очей, форма та довжина обличчя, розташування рота та носа на обличчі. Мозок людини проробляє все це автоматично і миттєво. Фактично, люди надзвичайно добре розпізнають обличчя і, в кінцевому підсумку, бачать обличчя в повсякденних предметах: комп'ютери нездатні до такого високого рівня узагальнення, тому доводиться вчити їх кожному кроку в цьому процесі окремо.

Існують різні погляди на визначення числа етапів розробки і проектування інтелектуальних систем. Це залежить від багатьох факторів, зокрема від характеру функцій майбутньої інтелектуальної системи, галузі використання, наявності розвинених інструментальних засобів і багатьох інших.

Розглянемо деякі з існуючих програмних систем автоматичного розпізнавання обличчя. Підґрунтям для цього є у першу чергу мобільні рішення для заміни обличчя образом-моделлю.

Рішення заміни обличчя образом-моделлю вже були застосовані на початку 2000-х, але не набули широкого поширення через низьку якість апаратно-програмних рішень, недостатню точність і малу обчислювальну потужність пристроїв того покоління. Наприклад, компанія Logitech включала подібний програмний додаток до складу драйверів для своїх веб-камер. Він був орієнтований лише на топові моделі камер скоріш за все через погану якість більш дешевих аналогів.

Під час аналізу конкурентів не можна не звернути увагу на MSQRD та FaceApp. Ці додатки являють собою цілком розважальні продукти, тому оцінка буде саме з урахуванням цього аспекту.

MSQRD – додаток «Маскарад» призначений для створення анімованих маски та накладає їх на обличчя. Анімовані маски автоматично підлаштовуються під розмір особи і рухаються як живі. Сильними сторонами MSQRD є якісна

реалізація і абсолютна безкоштовність. На даний момент всередині програми можна знайти всього близько 15 масок. Але скромна кількість контенту компенсується відмінною оптимізацією: кожна з масок ідеально лягає на риси обличчя і цілком повторює рухи. Завдяки чому є можливість створювати забавні відеоповідомлення з накладенням анімованих масок. Присутня можливість запису відео та фото, рис. 1.8.



Рисунок 1.8 – Зовнішній вигляд додатку «MSQRD»

FaceApp – у січні 2017 року асортимент каталогу App Store поповнився додатком під назвою FaceApp, за допомогою якого можна змусити посміхнутися навіть самого сурового. Дуже нагадує інші програми, які також використовують нейронні мережі для зміни фото- і відео контенту. Однак в той час, як MSQRD демонструє результати в режимі реального часу, FaceApp змінює тільки окремі фрагменти зображення, а вихідні дані відправляються на віддалений сервер, де і відбувається обробка. Продукт не працює у режимі реального часу та потребує часу для змін обличчя. Результат можна експортувати до фототеки девайса, рис. 1.9. Додаток використовує нейронну мережу для того, щоб змінювати обране фото.

Головним недоліком FaceApp є неможливість працювати у режимі реального часу. Під реальним часом мається на увазі можливість виконувати достатню кількість обчислень (принаймні не менше 21 на секунду, або приблизно 16 мсек для рендерінгу одного кадру). Потужність сьгоднішніх мобільних пристроїв є цілком достатньою для того, щоб працювати із подібними задачами.

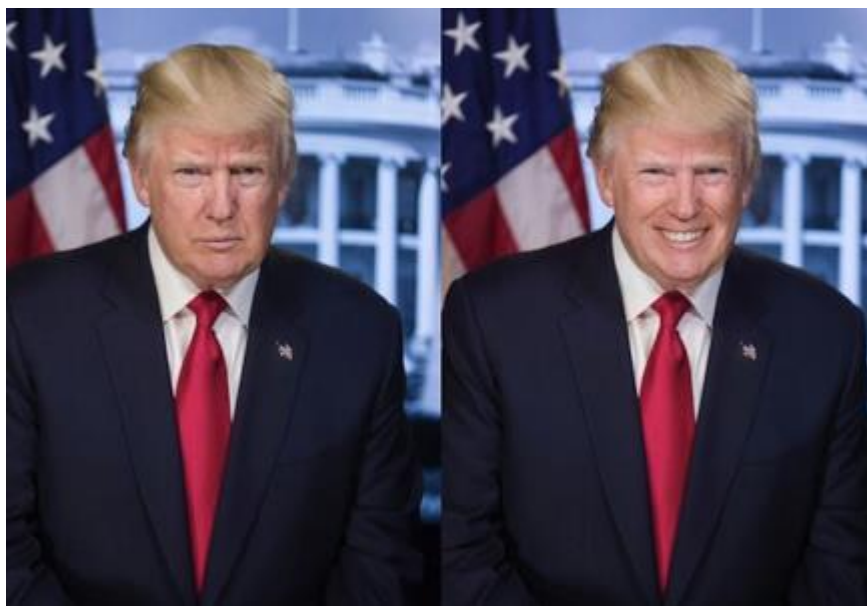


Рисунок 1.9 – Результат роботи FaceApp

Найпростіші функції розпізнавання осіб також реалізовані в цифрових фотоапаратах багатьох фірм, в тому числі Canon та Fujі. Вбудовані програми пошуку можуть автоматично знаходити в зображенні видошукача людські обличчя за характерними ознаками – очами, вухами, носом тощо. Фірма Sony розробила цифрову камеру, яка утримує затвор від спрацьовування до тих пір, поки люди не посміхнуться, досліджуючи положення куточків рота, розмикання губ, мімічні зморшки навколо очей. Розробляються програми для розпізнавання обличчя за допомогою камер мобільних пристроїв. Смартфони Apple реалізують цю функцію.

Компанії, які є лідерами в розробці технологій розпізнавання обличчя: ZN Vision Technologies (системи ZN-Face, ZN-Phantomas і ZN SmartEye), SAFLINK (біометричні add-on-модулі для Windows), Imagis Technologies (CABS –

інтегрована система обліку правопорушень і злочинців). Також на ринку представлені компанії Geometrix (3D сканери обличчя), Genex Technologies (3D сканери обличчя) в США, Cognitec Systems GmbH (SDK, спеціальні обчислювачі, 2D камери) в Німеччині, Bioscrypt (3D сканери обличчя). Однак їх використання разом з мобільними пристроями та смартфонами на теперішній час неможливе.

1.4 Постановка задачі дослідження

В рамках атестаційної роботи магістра планується дослідити алгоритми, за якими здійснюється розпізнавання образів, розробити модифікований алгоритм з можливістю роботи в режимі реального часу, розробити додаток для пристроїв Apple, в якому буде використовуватися метод послідовного перетворення зображення, отриманого з відеоряду, для його подальшого використання при розпізнаванні особи за обличчям для отримання доступу до оперативної інформації про особу.

Процес, який покладено в основу додатку, передбачає виконання обробки кадрів відеопотоку в режимі реального часу та складається з двох ключових стадій. Перший крок – виявлення обличчя методом Віюлі-Джонса, другий – виконання розпізнавання обличчя шляхом застосування методів Fisherfaces або Eigenfaces.

Слід врахувати, що на показник ефективності алгоритмів безпосередньо впливає ряд факторів, до яких входять особливості освітлення, розташування особи, наявність сторонніх предметів тощо. З цієї причини доречно скласти детальний опис умов, що сприяють продуктивній роботі алгоритму:

– виключно монотонна зміна освітлення. Під час проведення зйомки для навчальної та тестової вибірки необхідно забезпечити ідентичні умови освітлення;

- застосування фронтального або наближеного до нього положення осіб з нейтральним виразом їх обличчя;

- виключення можливість перекривання осіб іншими об'єктами.

Визначимо необхідний функціонал для системи що розробляється:

- відеопотік буде оброблятися в режимі реального часу за допомогою вбудованої камери;

- реалізовано доступ до налаштувань параметрів, необхідних для успішного функціонування алгоритмів пошуку і розпізнавання;

- вивід інформації про розпізнаний об'єкт, що містить у собі ступінь належності особи до конкретної групи, її характеристики та графічне відображення уявлення особи, що відслідковується;

- можливість навчання та додавання груп осіб за допомогою камери через інтерфейс програми.

З метою покращення результативність роботи системи будуть задіяні проміжні етапи обробки обличчя знайдених осіб: використання фільтра Гаусса для зменшення впливу шумів та використання маски значущих областей щодо локалізованих і модернізованих зображень обличчя. Планується, що в результаті вдасться усунути негативний вплив кутових ділянок зображення під час ідентифікації при наявності заднього плану.

В результаті загальний алгоритм розпізнавання повинен складатися з наступних етапів: виявлення обличч в кадрі, обробка знайдених обличч фільтром Гаусса, використання FisherFaces або EigenFaces модифікації щодо знайдених об'єктів з подальшим застосуванням маски значущих областей.

Визначимо вимоги.

У загальному випадку під вимогами розуміють сукупність властивостей, яким повинна відповідати система. В рамках розв'язання задачі дослідження необхідно виконати програмну реалізацію системи розпізнавання осіб за обличчям. Метою є інтеграція системи розпізнавання за обличчям до систем обміну відеоповідомленнями для отримання оперативної інформації про осіб, що

приймають участь в чаті. З поставленої мети виходять мінімальні вимоги до системи.

Призначення розробки.

Програмна система з елементами розпізнавання образів призначена для отримання доступу до БД, в якій зберігається інформація про учасників відеоконференції. У разі, якщо хтось бажає надсилати замість образу свого обличчя будь-яку заміну, система повинна попереджати про це.

Система призначена для широкого кола користувачів, тобто будь-яка людина, яка активно користується телефонами та є активним користувачем соціальних мереж, є потенційним користувачем системи.

Загальні відомості про програмне забезпечення розробки. При розробці програмної системи буде використовуватись середовище розробки xCode Version 8.3.2 (8E2002) на мовах C++, Objective-C, Objective-C++, Swift 3 [8], технології OpenCV та OpenGL ES. Розробка та тестування програми буде проводитись з використанням Macbook Pro MacOS Sierra 10.12.4, iPhone 9 та iPhone X (iOS 12.3.1).

2 МОДЕЛЮВАННЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЗА ОБЛИЧЧЯМ

2.1 Загальні проблеми розпізнавання за обличчям

Будь-яка система розпізнавання за обличчям – це типова система розпізнавання образів, завдання якої зводиться до формування деякого набору ознак, так званого біометричного шаблону, згідно закладеної в систему математичної моделі. Розпізнавання обличчя в будь-якій біометричній системі виконується в кілька етапів: виявлення обличчя, оцінка якості, побудова шаблону, зіставлення і прийняття рішення [3].

На етапі виявлення обличчя система автоматично визначає в потоці відеокадрів або на фотографії обличчя людей, причому діапазон ракурсів і масштабів осіб може значно варіюватися. Виявлення обличчя є одним з ключових етапів розпізнавання, так як пропуск особи детектором автоматично означає неможливість подальшої ідентифікації. Якість роботи детектора прийнято характеризувати ймовірністю виявлення особи. Для сучасних біометричних систем, що працюють в умовах потоку людей, значення ймовірності виявлення особи становить від 95 до 99% і залежить від умов реєстрації відео (освітленість, роздільна здатність тощо).

На етапі оцінки якості здійснюється вибір з усього масиву виділених облич тільки тих зображень, які задовольняють заданим критеріям якості. На практиці системи біометричної ідентифікації змушені мати справу з не дуже сприятливими умовами роботи: відхилення обличчя від фронтального положення на кути, що перевищують 20 град.; сильне засвічування; перекриття частини обличчя; наявність тіней на обличчі; малий розмір зображення тощо. Саме стабільність роботи біометричної системи в таких складних умовах і визначає її якість. Як правило, оцінюються: ракурс обличчя (не повинен перевищувати 20-30 град.); розмір обличчя (оцінюється за відстанню між зіницями очей і повинен бути більшим за 50-80 пкс); часткове закриття обличчя (закриття не повинно бути більшим за 10-25% від загальної площі обличчя).

Побудова шаблону – це один з найбільш складних і унікальних етапів розпізнавання за обличчям, що становить ключове досягнення новітніх технологій. Суть даного етапу полягає в нетривіальному математичному перетворенні зображення обличчя в набір ознак, об'єднаних в біометричний шаблон. Принципи побудови біометричних шаблонів надзвичайно різноманітні. Найважливішою характеристикою біометричного шаблону є його розмір. Чим більший розмір шаблону, тим більше інформативних ознак він включає до себе, але тим нижча швидкість і ефективність пошуку цього шаблону.

Зіставлення і прийняття рішення – це об'єднаний етап роботи системи розпізнавання, на якому проводиться порівняння біометричного шаблону обличчя, побудованого за виділеним обличчям, з масивом шаблонів. У найпростішому випадку зіставлення здійснюється простим перебором всіх шаблонів і оцінкою міри їх схожості. На підставі отриманих оцінок та їх зіставлення із заданими порогоми приймається рішення про наявність чи відсутність ідентичної особи в базі даних. У сучасних системах зіставлення реалізується за складними оптимальними схемами порівняння, що забезпечує швидкість зіставлення від 10 000 до 200 000 порівнянь на секунду і більше. Процес порівняння може бути розпаралеленим, що дозволяє працювати системам ідентифікації практично в режимі реального часу навіть по великих масивах зображень.

Ефективність роботи алгоритму стосовно заданого біометричного параметру зазвичай оцінюють за двома критеріями [6]:

– FAR (False Acceptance Rate) – коефіцієнт помилкового доступу, процентний показник випадків, при яких перевірка особи виявилася помилково успішною.

– FRR (False Rejection Rate) – коефіцієнт помилкової відмови в доступі, процентний показник випадків, при яких перевірка особи помилково завершилася невдачею.

Теоретично система тим краща, чим менше значення мають FRR та FAR. Однак, у більшості випадків більш значимою є одна з цих величин. Зокрема, для

системи контролю логічного або фізичного доступу пріоритетом є заборона доступу не уповноважених осіб за будь-яких обставин, як більш критичної обставини.

Очевидно, що для цього необхідний дуже низький FRR. Залежно від конкретного завдання можливе настроювання на певний компроміс між припустимими значеннями FRR і FAR, або, як їх прийнято називати в теорії статистичних рішень, помилками 1-го та 2-го роду.

Ефективність розпізнавання обличчя безпосередньо залежить від таких факторів, як стійкість біометричного шаблону до різного роду перешкод, спотворень у вихідному фото- або відеозображенні [3-5].

Надійність роботи системи розпізнавання осіб залежить від декількох факторів [7]:

- якість зображення (помітно знижується ймовірність безпомилкової роботи системи, якщо людина, яку намагаємося ідентифікувати, дивиться не прямо в камеру або знята при поганому освітленні);

- актуальність фотографії, занесеної до бази даних;

- величина бази даних (обсяг баз даних при використанні мобільних гаджетів не перевищує 10000 зображень).

Якість роботи систем розпізнавання осіб прийнято характеризувати ймовірностями ідентифікації. Хоча обличчя людини і є унікальним параметром, але досить мінливим – риси обличчя змінюються в залежності від повороту голови, психоемоційного стану, мімічного виразу, наявності бороди, вусів, окулярів, косметики тощо.

Для забезпечення високої надійності розпізнавання незалежно від цих факторів, кількість, якість і різноманітність зчитувальних образів може варіюватися в залежності від алгоритмів і функцій системи, що реалізує даний метод. Таким чином, система, яку планується реалізувати в рамках атестаційної роботи, повинна бути стійкою до помилок 1-го та 2-го роду.

2.2 Методи 2D та 3D розпізнавання за обличчям

Вся множина методів розпізнавання за геометрією обличчя ділиться на два напрямки: 2D і 3D методи розпізнавання [10]. У кожного з них є переваги і недоліки, проте багато що залежить ще і від області застосування і вимог, пред'явлених до конкретного алгоритму.

2.2.1 Розпізнавання обличчя в 2D

Цей напрямок ідентифікації з'явилося давно і бере початок в криміналістиці. Розпізнавання обличчя спочатку мало низьку в порівнянні з іншими методами надійність. Високі результати досягалися лише при фіксованих зовнішніх факторах (ракурс, освітленість, дальність і т.п.). В даний час він застосовується лише в багатофакторній (перехресній) автентифікації, або в соціальних мережах (вказівка людей на фото в Facebook).

В задачах ідентифікації при використанні великих баз даних надійність і швидкодія таких біометричних систем різко знижується, змушуючи використовувати додаткові ознаки для автентифікації. На практиці також пред'являються вимоги до освітлення, відсутності зовнішніх перешкод. Обов'язкове фронтальне зображення обличчя з досить невеликими відхиленнями, багато алгоритмів не враховують можливі зміни міміки обличчя. Все це додає труднощів при ідентифікації і встановлює певні мінімальні вимоги до обчислювальної потужності апаратури. На практиці досить стандартних відеокамер з роздільною здатністю 320x240 ppi, які передають дані зі швидкістю відеопотоку, принаймні 3-5 кадрів в секунду. Інтенсивний розвиток і, як наслідок, здешевлення цифрового відео і мультимедійних цифрових технологій дозволяють

впровадити їх в широке використання. Найбільш поширеними пристроями, що дозволяють отримати двомірне зображення обличчя користувача є веб-камери.

2D система працює з відносно простим двовимірним зображенням, що помітно спрощує алгоритми і знижує інтенсивність обчислень. Переваги методу 2D розпізнавання обличчя: не потрібне дороге обладнання; при відповідному обладнанні можливість розпізнавання на значних відстанях від камери. Недоліки: низька статистична достовірність; вибагливість до освітлення; неприйнятність будь-яких зовнішніх перешкод; не враховують можливі зміни міміки обличчя, вираз повинен бути нейтральним [1].

Сучасні алгоритми здатні компенсувати наявність окулярів, вусів і бороди, а також додаткових аксесуарів на обличчі досліджуваної людини навіть на двомірному зображенні. Однак основною проблемою використання двомірних зображень є вразливість до атак з використанням муляжів. Для обману таких систем досить використання фотографії суб'єкта.

2.2.2 Розпізнавання обличчя в 3D

Реалізація являє собою досить складне математичне і технічне завдання. В даний час існує багато методів з 3D-розпізнавання обличчя. Методи неможливо порівняти один з одним, оскільки вони використовують різні сканери та бази, не для всіх з них вказані FAR і FRR, використовуються абсолютно різні підходи.

Класичним є метод проектування шаблону. Він полягає в тому, що на обличчя проектується світлова сітка. Промінь, що падає на викривлену поверхню, згинається – чим більша кривизна поверхні, тим сильніше вигинається промінь. Спочатку застосовувалося джерело видимого світла, а потім – інфрачервоне. Камера робить знімки зі швидкістю десятки кадрів на секунду, а отримані зображення обробляються спеціальною програмою.

За отриманими знімками відновлюється 3D модель обличчя, на якій виділяються і видаляються непотрібні перешкоди (зачіска, борода, вуса та окуляри). Потім проводиться аналіз моделі – виділяються антропометричні особливості, які записуються в унікальний код, що заноситься в базу даних. Крім низької чутливості до зовнішніх чинників, найважливішою перевагою методу є високий рівень надійності.

Переваги методу 3D розпізнавання обличчя: висока достовірність розпізнавання – більше інформації, ніж має звичайний знімок; стійкість розпізнавання до відхилення ракурсу особи від фронтального; стійкість розпізнавання до неоднорідності освітлення; відсутність необхідності контактувати з пристроєм; низька чутливість до зовнішніх факторів. Недоліки: має обмежену сферу застосування із-за поганих статистичних показників; дороге обладнання; зміна міміки обличчя і перешкоди на обличчі погіршують статистичну надійність методу [1].

Досить надійні системи базуються на застосуванні декількох камер, розташованих під різними кутами і забезпечують формування тривимірної моделі обличчя. У них також використовується додаткове підсвічування для зниження впливу освітлення на одержуваний результат. Подібні системи знаходять застосування на контрольно-пропускних пунктах. У той же час застосування таких систем звичайними користувачами в повсякденних умовах неможливо через їх високу вартість, складності установки і використання [6].

2.3 Застосування методики розпізнавання за обличчям

Відповідно до постановки задачі дослідження метою атестаційної роботи є практичне застосування методів та алгоритмів розпізнавання за обличчям в режимі реального часу на мобільних пристроях фірми Apple. Процес, що покладено в основу мобільного додатку, передбачає виконання обробки кадрів

відеопотоку в режимі реального часу та складається з двох ключових стадій. Перший крок – виявлення обличчя методом Віоли-Джонса, другий – виконання розпізнавання обличчя шляхом застосування методів Fisherfaces або Eigenfaces. Зі структури процесу зображеної на малюнку 2.1 видно, що для вирішення головного завдання розпізнавання осіб за обличчям необхідно вирішити завдання більш низького рівня такі як детектування обличчя на зображенні, нормалізація локалізованого обличчя, обчислення ключових ознак і порівняння їх з еталоном.

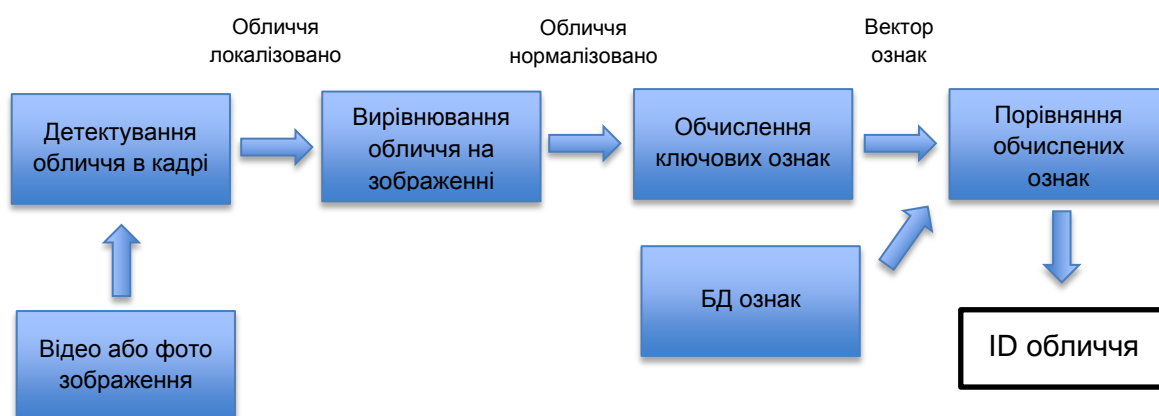


Рисунок 2.1 – Структура процесу розпізнавання особи за обличчям

Для розв’язання кожної з підзадач існує певний набір методів, від ефективності яких залежить вирішення завдання розпізнавання в цілому.

2.3.1 Метод Віоли-Джонса

Метод входить в число ключових способів пошуку об'єктів на зображенні.

Принципами його функціонування є:

- інтегральне представлення зображень;
- визначення обличчя з використанням ознак Хаара;
- каскадна класифікація із застосуванням бустінгу.

За допомогою інтегрального представлення зображень з'являється можливість оперативно визначити показник яскравості прямокутної ділянки

зображення [9]. Додатково це дає можливість оперативно розрахувати сумарну яскравість довільного прямокутника на зображенні, причому площа фігури не впливає на тривалість виконання операції.

Інтегральне представлення зображення – це матриця, параметри якої ідентичні параметрам джерела. Компоненти матриці містять загальну інтенсивність пікселів, розташованих лівіше і вище елементу. Складові матриці можна розрахувати за формулою:

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x, y), \quad (1)$$

де $I(x, y)$ – значення точки (x, y) інтегрального зображення; $i(x, y)$ – показник інтенсивності початкового зображення.

В результаті з'являється можливість обчислити показники однакового виду, але з відміними у геометричній формі, за однаковий проміжок часу.

Ознака Хаара визначається наступним чином: Ознакою f об'єкту a є відображення типу $f: A \rightarrow D_f$, де D_f – це ряд допустимих значень цієї ознаки. При наборі векторів f_1, \dots, f_n , вектор $x = (f_1(a), \dots, f_n(a))$ вважається ознакою опису об'єкту $a \in A$.

Ознаки Хаара – це прямокутні ознаки, загальний вигляд яких наведено на рис. 2.2.

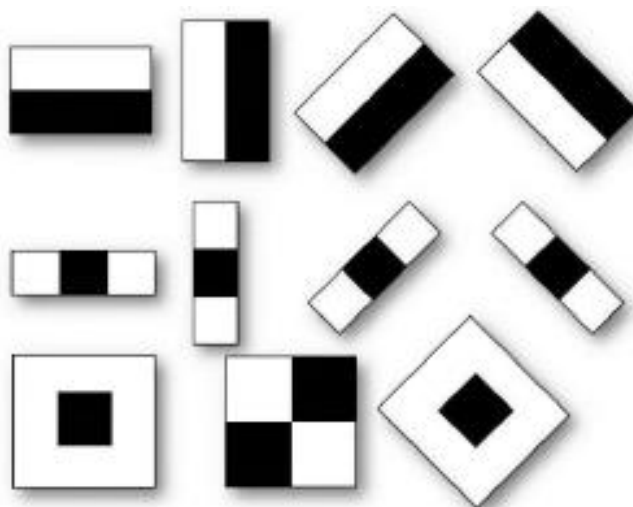


Рисунок 2.2 – Зображення базових ознак Хаара

В бібліотеці комп'ютерного зору OpenCV в методі Віоли-Джонса задіяно ряд додаткових ознак, наведених на рис. 2.3.

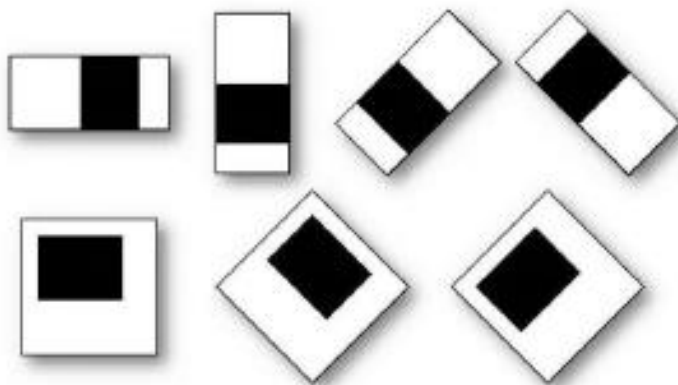


Рисунок 2.3 – Зображення додаткових ознак Хаара

Результат обчислення такої ознаки для інтегрального подання зображення можна записати наступним чином:

$$F = U - V, \quad (2)$$

де U – ознака загального значення яскравості точок, які перекриваються світлою частиною ознаки, а V – темною. За допомогою цих ознак вдається безпомилково описати коливання яскравості за осями зображення.

2.3.2 Скануюче вікно

Пошук обличчя виконується за допомогою скануючого вікна, параметри якого в оригінальному алгоритмі складають 24x24 пікселя. Воно пересувається по зображенню з кроком в 1 піксель та в кожній новій точці обчислюються ознаки Хаара з урахуванням масштабу та розташуванням їх у вікні. Ці ознаки спрямовуються у класифікатор, який за їх значеннями визначає ступінь відповідності вікна та проводить аналіз, чи є виділена область зображення

обличчям або ні. Для цього це значення порівнюється з порогом, який визначається при навчанні. Якщо значення перевищує поріг, то ознака виявив об'єкт, в іншому випадку – не виявив. Ознаки реагує по-різному, тому виділяють сильні і слабкі ознаки.

Ознаки обчислюються в межах скануючого вікна, який переміщається по всьому зображенню (рис.2.4).

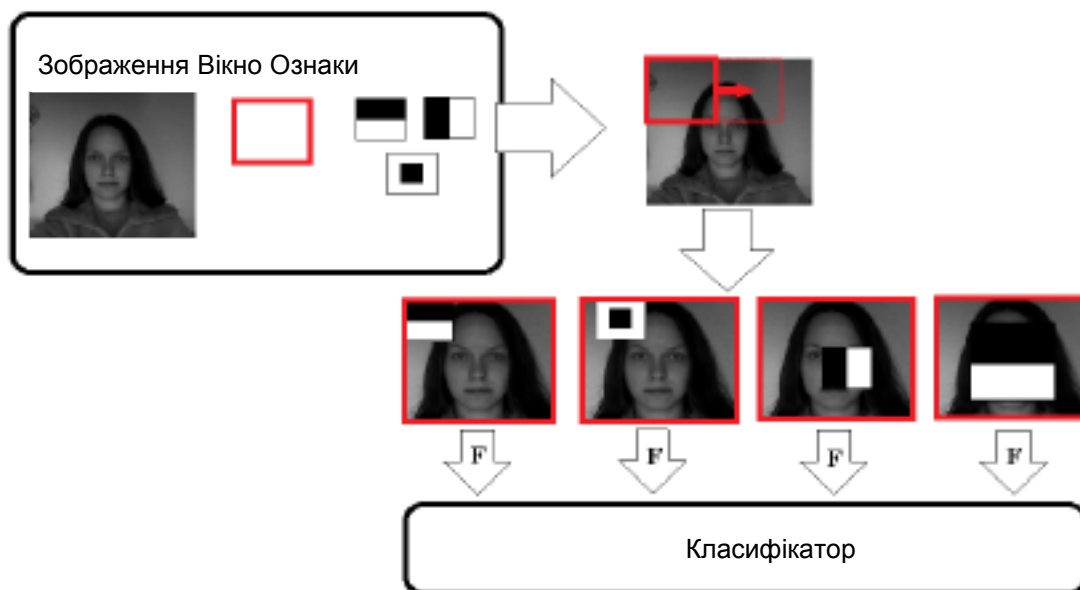


Рисунок 2.4 – Схема роботи принципу скануючого вікна

Цей процес повторюється кілька разів для інших розмірів скануючого вікна, тобто застосовується масштабування над вікном та ознаками.

Структура класифікатора вищевказаного типу сприяє збільшенню продуктивності роботи алгоритму з виявленню обличчя, приділяючи максимум уваги областям зображення, що викликають найбільший інтерес. Каскадна класифікація базується а використанні групи неактивних класифікаторів, навчених за допомогою процедури бустінга.

Бустінгом називають комплекс методів, що сприяють підвищенню точності аналітичних моделей. Бустінг – це сучасний метод класифікації та розпізнавання образів, що представляє собою процедуру послідовного побудови композицій

алгоритмів навчання, коли кожен наступний алгоритм прагне компенсувати недоліки попереднього.

Класифікатор, що допускає мало помилок називається «сильним», а класифікатор, який класифікує з достовірністю 51% – «слабким» або простим.

Отже, при мінімальних витратах на обчислення, каскадна класифікація дозволяє вже на перших стадіях розпізнавання відкинути ті зображення, на яких найімовірніше відсутній необхідний об'єкт (в конкретному випадку – обличчя). Приклад каскадної структури класифікаторів наведено на рис. 2.5.

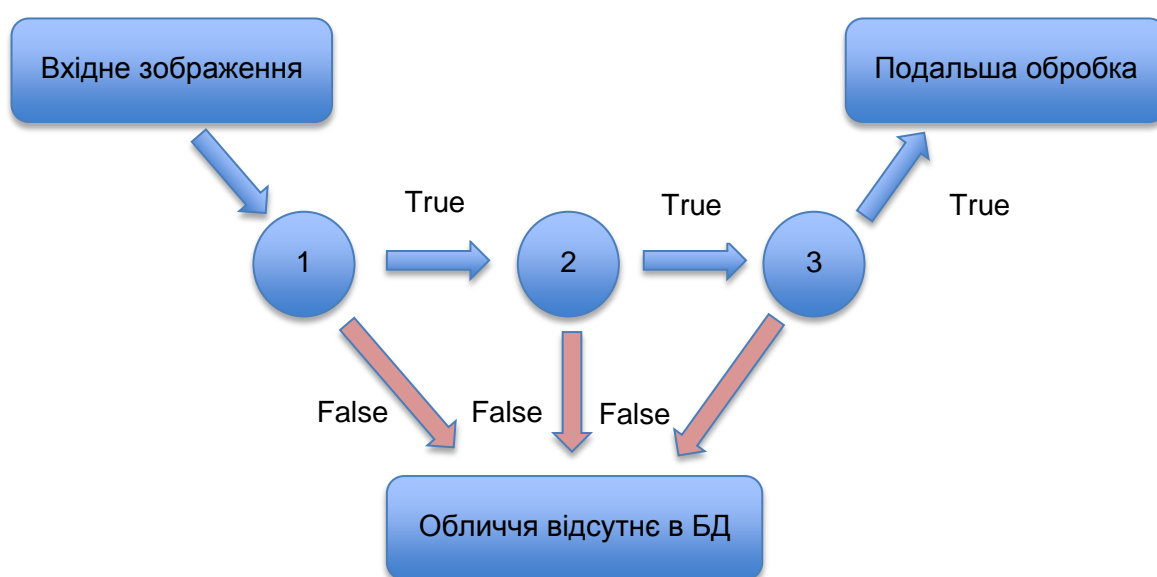


Рисунок 2.5 – Приклад каскадного класифікатора

Бустінг є жадібним алгоритмом побудови композиції алгоритмів (greedy algorithm) який на кожному кроці робить локально найкращий вибір в надії, що кінцеве рішення буде оптимальним. Бустінг вважається одним з найбільш ефективних методів з точки зору якості класифікації. У багатьох експериментах спостерігалось практично необмежену зменшення частоти помилок на незалежній тестовій вибірці по мірі нарощування композиції.

Розвитком даного підходу стала розробка більш досконалого сімейства алгоритмів бустінга – AdaBoost (AdaBoost: adaptive boosting – адаптоване поліпшення), який може використовувати довільне число класифікаторів і

виробляти навчання на одному наборі прикладів, по черзі застосовуючи їх на різних етапах.

Всі каскадні рівні навчаються за допомогою алгоритму AdaBoost. Число властивостей, застосованих у ньому, зростає до того моменту, поки не буде виявлений цільовий об'єкт та помилки першого роду. Якщо не вдалося виявити загальну помилку, з кожним етапом необхідно додавати ще 1 шар.

Негативну множину ознак можна отримати шляхом збору невірних виявлень під час застосування поточного каскаду. За підсумками класифікації вдається виявити області зображення, де знаходиться шуканий об'єкт, а потім необхідно виключити вкладені повторення при виявленні аналогічних об'єктів через масштабування скануючого вікна.

2.3.3 Опрацювання шумів

Щоб запобігти негативному впливу, які надають шуми на зображеннях обличчя, використовують фільтр Гаусса. Він являє собою фільтр розмиття зображення, застосовуваний показник стандартного поділу з метою обчислення зміни щодо кожного пікселя. Розподіл для 2-х вимірів обчислюється за формулою (3) [12]:

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)}, \quad (3)$$

де r – радіус розмиття; $r^2 = x^2 + y^2$; σ – показник нормальної розбіжності.

За допомогою цієї формули задається поверхня, яка виглядає як набір концентричних кіл зі стандартним розподілом відповідно до центральної точки. Пікселі, де розподіл відрізняється від нуля, застосовуються з метою побудови матриці згортки, яка використовується по відношенню до джерела. Початковий показник пікселя приймає максимальне значення, а сусідні з ним елементи

зменшуються пропорційно відстані до них. Слід зазначити, що розмиття по Гаусу є ефективним способом усунення небажаних шумів на зображеннях. Результат використання фільтра Гауса до цілого зображення наведений на рис. 2.6.



Рисунок 2.6 – Результат застосування фільтра Гауса до зображення з шумом

Серйозною проблемою, що стоїть перед системами розпізнавання, є велика мінливість візуальних образів, пов'язана зі змінами освітленості, забарвлення, масштабів, ракурсів спостереження (рис.2.7). Колір і яскравість окремих пікселів на зображенні також залежить від великої кількості важко прогнозованих факторів. У число цих факторів входять:

- число і розташування джерел світла;
- колір і інтенсивність випромінювання;
- тіні або віддзеркалення від навколишніх об'єктів.



Рисунок 2.7 – Проблема освітлення

Завдання виявлення об'єктів на зображенні ускладнюється також через значний обсяг даних, що містяться в зображенні. Зображення може містити тисячі пікселів, кожен з яких може мати важливе значення. Повне використання інформації, що міститься в зображенні, вимагає аналізу кожного пікселя на приналежність його об'єкту або фону з урахуванням можливої мінливості об'єктів. Такий аналіз може вимагати високих витрат необхідної пам'яті і продуктивності комп'ютера, що вкрай небажано при використанні мобільних пристроїв.

Вирішення цієї проблеми лежить в правильному виборі опису об'єктів, для виявлення і розпізнавання яких створюється система. Опис об'єкта має враховувати найбільш характерні особливості опису і бути досить повним, щоб відрізнити даний об'єкт від інших елементів навколишньої сцени. Щоб уникнути суб'єктивності при виборі потрібного опису, можна використовувати методи автоматичного вибору відповідних характеристик об'єкта.

Незважаючи на різноманітність алгоритмів та підходів, типовий метод розпізнавання обличчя складається з трьох основних компонентів:

- перетворення вихідного зображення в стандартний вигляд;
- виділення ключових характеристик;
- вибір механізму класифікації (моделювання): кластерна модель, метрика, нейронна мережа та ін.

Побудова алгоритму розпізнавання спирається на апіорну інформацію про предметну область (в даному випадку – характеристики обличчя людини) та коригується експериментальною інформацією, що з'являється по ходу розробки методу. З попереднього аналізу першоджерел нам відомо, що найбільшого поширення набули такі алгоритми розпізнавання обличчя: Eigenfaces, Fisherfaces, Local Binary Patterns Histograms. Спираючись на той факт, що програмна реалізація передбачає використання мобільних пристроїв Apple, звернемо основну увагу саме на алгоритм Eigenfaces.

В основі цього алгоритму лежить метод головних компонент (Principal Component Analysis – PCA), ідея якого полягає в тому, що масивні набори даних

описуються корельованими змінними, та для великої кількості інформації надається лише кілька значущих розмірів. PCA знаходить напрями з найбільшою дисперсією в даних, що називаються головними компонентами. Математично вказаний алгоритм можна представити наступним чином:

Якщо позначити $X = \{x_1, x_2, \dots, x_n\}$ – випадковий вектор з набором спостережень $x_i \in R^d$. Далі потрібно визначити середнє значення μ та обчислити матрицю коваріації S :

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i, \quad (4)$$

$$S = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T. \quad (5)$$

Після цього знаходять власні значення λ_i та власні вектори v_i із S :

$$Sv_i = \lambda_i v_i, i = 1, 2, \dots, n. \quad (6)$$

Зменшуючи власне значення, визначають власні вектори. До основних компонент спостережуваного вектору x додаються:

$$y = W^T(x - \mu), \quad (7)$$

де $W = (v_1, v_2, \dots, v_n)$.

Вхідні вектори є відцентрованими та приведеними до єдиного масштабу зображення обличчя. Власні вектори, обчислені для всього набору зображень обличчя, називаються власними обличчями, через це метод головних компонент в застосуванні до зображень обличчя також називають методом власних облич (Eigenfaces) (рис. 2.8).



Рисунок 2.8 – Приклад зображення власних векторів (власні обличчя)

З рис. 2.8 видно, що власних векторів недостатньо для точної реконструкції зображення, проте, досліди показують, що 50 власних векторів може бути достатньо для кодування ключових рис обличчя, що є цілком достатнім для розпізнавання обличчя на мобільній платформі.

Навчання класифікаторів йде дуже повільно, але результати пошуку особи дуже швидкі, саме тому був обраний даний метод розпізнавання осіб на зображенні. Метод Віоли-Джонса є одним з кращих по співвідношенню показників ефективності розпізнавання/швидкості роботи. Також цей детектор має вкрай низьку ймовірність помилкового виявлення особи. Алгоритм навіть добре працює і розпізнає риси обличчя під невеликим кутом, приблизно до 30 градусів. При куті нахилу понад 30 градусів відсоток виявлення різко падає. І це не дозволяє в стандартній реалізації виявляти повернене обличчя людини під довільним кутом.

Даний метод в загальному вигляді шукає особи і риси обличчя за загальним принципом скануючого вікна. У загальному вигляді, завдання виявлення особи та рис обличчя людини на цифровому зображенні виглядає саме так: є зображення, на якому є шукані об'єкти. Воно представлено двовимірною матрицею пікселів розміром $w * h$, в якій кожен піксель має значення:

- від 0 до 255, якщо це чорно-біле зображення;
- від 0 до 2553, якщо це кольорове зображення (компоненти R, G, B).

В результаті своєї роботи алгоритм повинен визначити особу та її риси і позначити їх — пошук здійснюється в активній області зображення прямокутними ознаками, за допомогою яких і описується знайдена особа та її риси:

$$rectangle_i = \{x, y, w, h, a\}, \quad (8)$$

де x та y – координати центру i -того прямокутника, w – ширина, h – висота, a – кут нахилу прямокутника до вертикальної осі зображення.

Алгоритм сканування вікна з ознаками виглядає так:

- є досліджуване зображення, обрано вікно сканування, обрані використовувані ознаки;
- далі вікно сканування починає послідовно рухатися по зображенню з кроком в 1 клітинку вікна (припустимо, розмір самого вікна є $24 * 24$ комірки);
- при скануванні зображення в кожному вікні обчислюється приблизно 200 тисяч варіантів розташування ознак, за рахунок зміни масштабу ознак та їх положення у вікні сканування;
- сканування проводиться послідовно для різних масштабів;
- масштабується не саме зображення, а скануюче вікно (змінюється розмір комірки);
- всі знайдені ознаки потрапляють до класифікатора, який «виносить вердикт».

Іншими словами, стосовно до малюнків і фотографій використовується підхід на основі скануючого вікна (scanning window): сканується зображення вікном пошуку (так зване, вікно сканування), а потім застосовується класифікатор до кожного положення [9]. Система навчання і вибору найбільш значущих ознак повністю автоматизована і не вимагає втручання людини, тому даний підхід працює швидко.

Завдання пошуку і знаходження осіб на зображенні за допомогою даного принципу часто буває черговим кроком на шляху до розпізнавання характерних рис, наприклад, верифікації людини по розпізнаваній особі або розпізнавання міміки обличчя.

2.3.4 Навчальна вибірка

Результати навчання каскадів у вигляді класифікаторів Хаара знаходяться в XML-форматі.

Для кожного класифікатора, вираженого як сутність з підлеглими атрибутами треба створити окреме XML-сховище DOM (Document Object Model) — об'єктна модель документа такого класифікатора в загальному вигляді представлена нижче (рис.2.9).

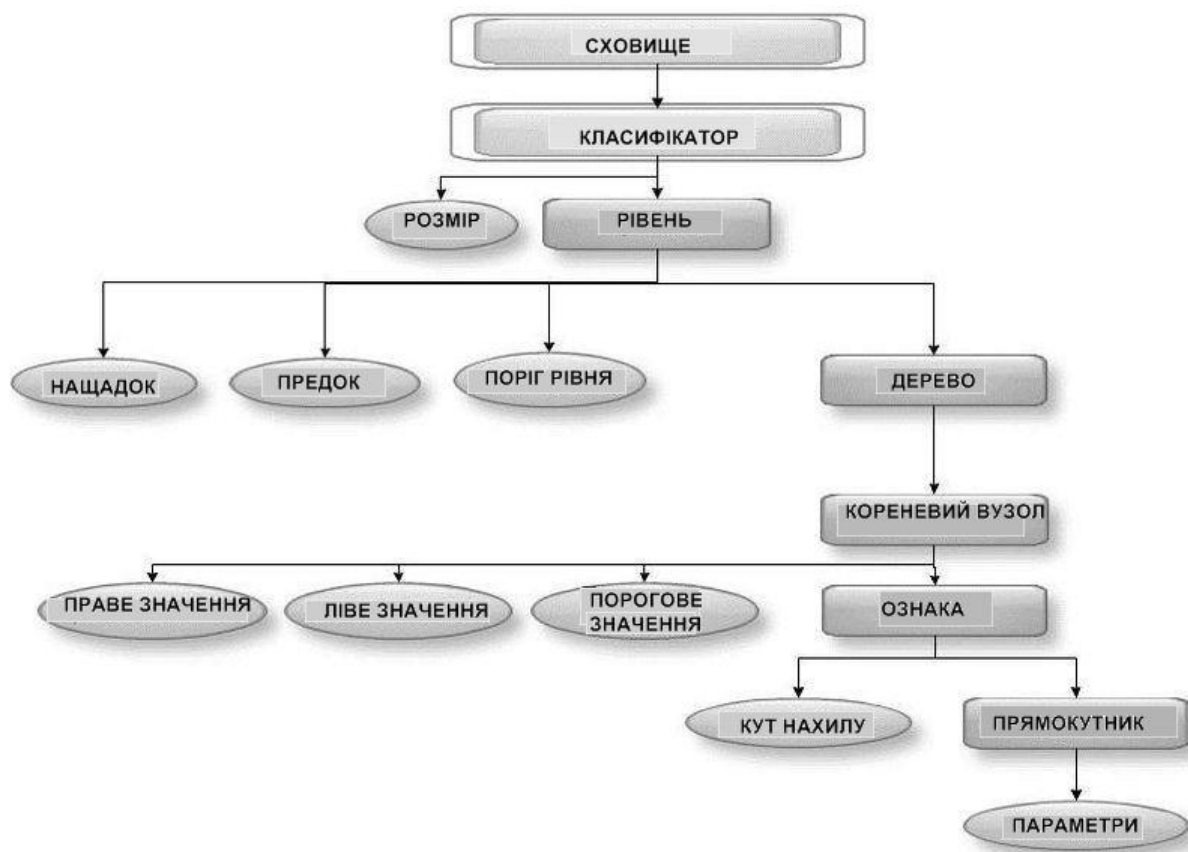


Рисунок 2.9 – Об'єктна модель документа класифікатора

Нижче наведено класифікатор, що використовується в системі розпізнавання:

```

<opencv_storage>
<haarcascade_eye_tree type_id="opencv-haar-classifier">
  <size>
    20 20</size>
  <stages>
    <_>
      <!-- stage 0 -->
  
```

```

    <trees>
  <_>
  <!-- tree 0 -->
  <_>
    <!-- root node -->
    <feature>
      <rects>
        <_>
          8 7 12 1 -1.</_>
        <_>
          8 7 6 1 2.</_></rects>
      <tilted>1</tilted></feature>
    <threshold>-0.0269871093332767</threshold>
    <left_node>2</left_node>
    <right_node>1</right_node></_>
  <_>
    <!-- node 1 -->
    <feature>
      <rects>
        <_>
          4 7 8 6 -1.</_>
        <_>
          6 7 4 6 2.</_></rects>
      <tilted>0</tilted></feature>
    <threshold>0.0506705306470394</threshold>
    <left_val>-0.8039547204971314</left_val>
    <right_val>0.6049140095710754</right_val></_>
  <_>

```

Тут `haarcascade_eye_tree type_id` — набір слабких класифікаторів, на основі яких виносяться рішення про те, чи знаходиться об'єкт на зображенні чи ні, `left_val` та `right_val` – це параметри конкретного слабкого класифікатора.

На даному прикладі видно, яка інформація зберігається у використовуваних XML. Це інформація про класифікатор (`haarcascade`) і його розміри (`size`), про

використані рівні (stage), попередника, або предка даного рівня (parent), наступного рівня, або нащадка (next), вибудовується дерево (tree) і його кореневий вузол (root node) з параметрами (threshold, left, right). Далі вибудовується інформація про самі ознаки (features) у вузлах цього дерева, які задаються прямокутниками з параметрами (rects) з певним кутом нахилу (tilted).

2.4 Меможі від Apple

Face ID – це нова функція розпізнавання обличчя, розроблена Apple [11].

Щоб зробити Face ID можливим на iPhone X, Apple довелося розміщувати багато нового обладнання в крихітну область у верхній частині дисплея. Ця область дублюється «системою камери TrueDepth», і вона складається з верхнього підсвічування, інфрачервоної камери, передньої камери, точкового проектора, датчика наближення, датчика навколишнього освітлення, динаміка і мікрофона. Коли користувач заглядає в свій iPhone X, запускається система камери TrueDepth.

Щоб обробити всі дані особи, Apple розробила новий Aion Bionic neural engine. Це чіп A11 компанії Bionic з вбудованим нейронним двигуном. Animoji використовує систему камер TrueDepth, що застосовується для Face ID, а також Aion Bionic chip, щоб захоплювати і аналізувати більше 50 різних м'язових рухів на обличчі. Потім він відображає вирази обличчя в різних емодзі для створення Animoji. Програма може не тільки автоматично додати до тексту свиню, kota, тигра або панду замість смайлика, а й емодзі з обличчям власника пристрою. Якщо людина захоче створити свій емодзі, то зможе вибрати форму обличчя, колір шкіри, очі і їх колір, рот, ніс, вуха, аксесуари, головні убори, зачіску, міміку, поворот голови і багато іншого (рис.2.10).

В iPhone X використовується розробка Microsoft. Kinect використовує кілька технологій для створення 3D зображення хорошої якості. Одна з них називається структуроване світло – пристрій проектує заздалегідь відомий візерунок і використовує машинне навчання створює 3D-сцену. В новому iPhone

Х використовується схожа схема – проектор точок створює візерунок, а інфрачервона камера його читає.

У Kinect також встановлений модуль звичайної кольорової камери, який служить для створення карти глибин. Щоб визначити наскільки далеко розташовані об'єкти, аналізується зображення. Об'єкти, які занадто далеко – зовсім розмиті.

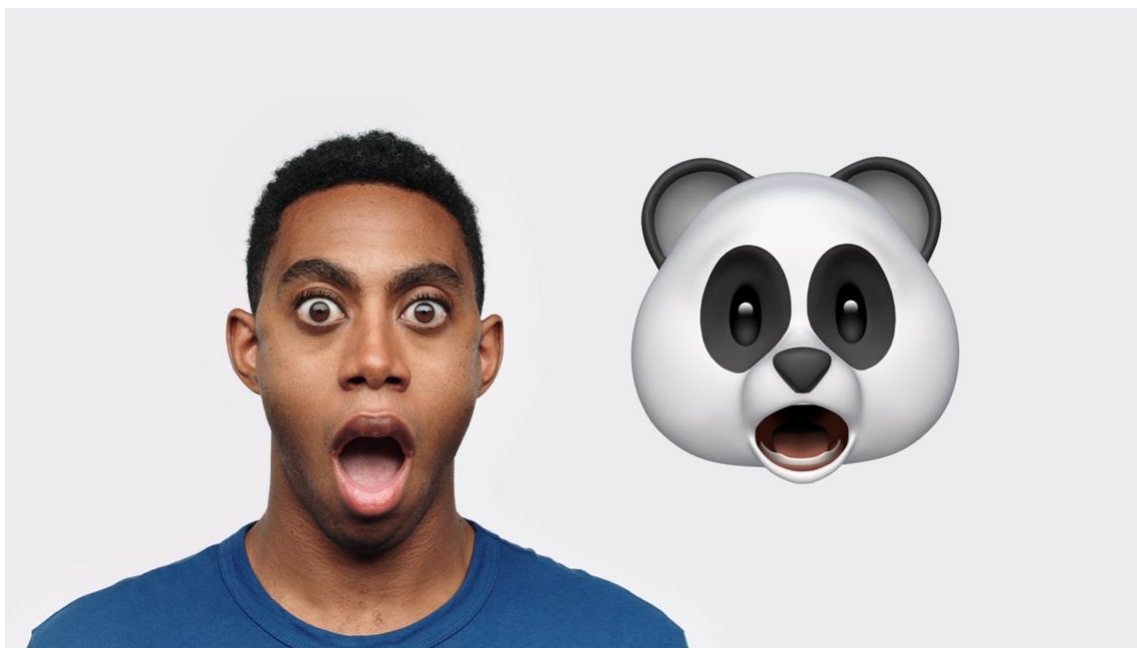


Рисунок 2.10 – Приклад відтворення мімики людини мімикою емодзі панди

В iPhone використовується ІК випромінювач, який дозволяє фронтальній камері розрізнити зображення навіть у повній темноті (інфрачервона камера використовується для читання точок проектора).



Рисунок 2.11 – Передня панель смартфона iPhone X

У будь-якому випадку обробка всіх даних зав'язується машинним навчанням та нейронними мережами. Для навчання нейронних мереж було використано тисячі прикладів положення тіла у випадку з Kinect і тисячами людей – у випадку з Apple iPhone X.

Функція Animoji використовує складну систему TrueDepth на iPhone X для створення об'ємної карти обличчя. Проте система занадто складна, але завдяки цьому вона максимально точна.

У попередніх версіях iOS стандартна програма FaceTime на пристроях Apple надавала можливість спілкуватися тільки з одним користувачем одночасно. Це не підходило для великих компаній друзів та колег. Apple перевершила очікування і відтепер одночасно можна спілкуватися з 32 друзями через FaceTime.

Таким чином на підставі проведеного аналізу методів розпізнавання за обличчям, технологій, що вбудовані в апаратне забезпечення смартфонів Apple, з'являється можливість реалізувати програмну систему, що використовує для доступу до БД саме ідентифікатор обличчя. Перейдемо до практичної реалізації системи.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ

3.1 Опис програмних засобів та бібліотек

Найбільш популярна на сьогодні реалізації великої кількості алгоритмів розпізнавання образів – бібліотека OpenCV. OpenCV (Open Source Computer Vision Library) – це бібліотека комп'ютерного зору, яка поставляється з відкритим вихідним програмним кодом. У ній зібрана велика кількість алгоритмів для використання технологій комп'ютерного зору, що надає широкі можливості. Після підключення даної бібліотеки до свого проекту користувач отримує доступ до більш ніж 500 функцій, призначених для вирішення різноманітних завдань. Окрім вирішення завдань комп'ютерного зору функціонал бібліотеки містить різноманітні чисельні алгоритми для обробки зображень. Бібліотеку реалізовано на мовах програмування C / C ++. Однак поставляється також і для інших мов, таких як Python, Java, Ruby, Matlab, Lua тощо. Бібліотека може використовуватися на різних операційних системах, до числа яких відносяться Linux, Mac OS X, iOS.

Основне завдання, яке вирішує бібліотека, це підвищення ефективності обчислень в додатках реального часу. Бібліотека OpenCV здатна використовувати багатоядерні процесори. При потребі автоматизації на різних апаратних платформах Intel можливо додатково придбати бібліотеку IPP(Integrate Performance Primitives). В склад даної бібліотеки входять процедури із низькорівневою оптимізацією, що можуть застосовуватися для різноманітних алгоритмічних областей.

Головною метою бібліотеки OpenCV є надання легкого у використанні інтерфейсу, який допоможе полегшити використання технологій комп'ютерного зору в досить складних додатках. Функції, які підтримує бібліотека, охоплюють різноманітні сфери комп'ютерного зору, від медицини, безпеки і до робототехніки.

Крім бібліотеки OpenCV, існують також інші бібліотеки комп'ютерного зору. наприклад, бібліотеки Matrox Imaging Library, Camellia Library, Open eVision, HALCON, VXL, libCVD, IVT, LTI, AForge.NET і деякі інші.

Як бібліотека для створення мобільних та веб додатків з використанням технології розпізнавання зображень на сьогоднішній день найбільш популярна бібліотека OpenCV. Такий вибір пов'язаний з тим, що в порівнянні з деякими іншими бібліотеками комп'ютерного зору, дана бібліотека має переваги. Першим і найбільш значною перевагою бібліотеки є її продуктивність. Продуктивність OpenCV перевищує продуктивність таких бібліотек комп'ютерного зору як VXL і LTI. Наявність значної кількості функцій (понад 500), орієнтованих на розв'язання різноманітні завдань від обробки зображень, комп'ютерного зору і до навчання машин, також зіграло роль при виборі OpenCV. Також бібліотека має відкритий вихідний код і ліцензію, яка дозволяє навіть з використанням функціоналу бібліотеки випускати комерційні продукти. Крім того, існує велика кількість навчальних та наукових матеріалів, в яких описуються функції і методи роботи з OpenCV, а в мережі Інтернет можна знайти велику кількість інформацій, в якій розробники і користувачі бібліотеки діляться досвідом і відповідають на питання, пов'язані з бібліотекою [10].

Останні версії бібліотеки були реорганізовані замість великих універсальних бібліотека OpenCV розділена на невеликі модулі по функціональному використанню:

- opencv_core – ядро: базові структури, обчислення (математичні функції, генерація псевдовипадкових чисел, DFT, DCT, введення / висновки в XML тощо);
- opencv_imgproc – обробка зображень (фільтри, перетворення і т. д.);
- opencv_highgui – простий UI, завантаження, збереження зображень і відео;
- opencv_ml – методи і моделі машинного навчання (SVM, дерева прийняття рішень тощо);
- opencv_features2d – різні дескриптори (SURF);
- opencv_video – аналіз руху і відстеження об'єктів (оптичний потік, шаблони руху, усунення фону);

- `opencv_objdetect` – детектування об'єктів на зображенні (вейвлети Хаара, HOG тощо);
- `opencv_calib3d` – калібрування камери та елементи обробки тривимірних даних;
- `opencv_flann` – бібліотека швидкого пошуку найближчих сусідів (FLANN).
- `opencv_legacy` – застарілий код, збережений заради зворотної сумісності.
- `opencv_gpu` – прискорення деяких функцій OpenCV за рахунок CUDA.

Основним класом для зберігання зображень в OpenCV є клас `Mat`. Використовуючи цей клас не потрібно вручну виділяти пам'ять і звільняти її, більшість функцій OpenCV буде автоматично розподіляти свої вихідні дані.

`Mat` – клас з двома частинами даних: матриця заголовка (містить таку інформацію, як розмір матриці, метод, який використовується для зберігання, в якому адреси зберігаються в матриці, і так далі) і вказівник на матриці, що містить значення пікселів (з будь-якої розмірністю в залежності від обраного методу для зберігання). Розмір заголовку матриці постійний, проте сам розмір матриці може варіюватися від зображення до зображення, і, як правило, більше на кілька порядків величини.

Для вирішення обчислювальної задачі за допомогою OpenCV більшу частину часу необхідно використовувати декілька основних функцій бібліотеки. Через це, передаючи зображення в функції є звичайною практикою. Робота з алгоритмами обробки зображень є, як правило, досить обчислювально важкі.

ROI (область інтересу) – один з фундаментів OpenCV. ROI дозволяє користувачеві задати певну прямокутну область зображення (ROI тобто завжди повинна знаходитися всередині вихідного зображення). Майже всі функції повинні підтримувати роботу з ROI, тобто роботу з виділеною областю зображення, що корисно для прискорення роботи алгоритмів. Якщо вас цікавить тільки певна область зображення – можна її виділити і працювати тільки з нею, не чіпаючи все зображення цілком. На рис. 3.1 наведено приклад виділеної області на зображенні.

Виникає логічне запитання, якщо сама матриця може належати до декількох об'єктів, який об'єкт Mat бере на себе відповідальність за її очищення, коли він не більше не потрібен. Коротка відповідь: останній об'єкт, який використовував його. Дана операція виконується за допомогою механізму підрахунку посилань. Всякий раз, коли хто-небудь копіює заголовок об'єкта Mat, лічильник збільшується для матриці. Всякий раз, коли заголовок очищається цей лічильник зменшується. Коли лічильник досягне нуля матриця теж звільняється.

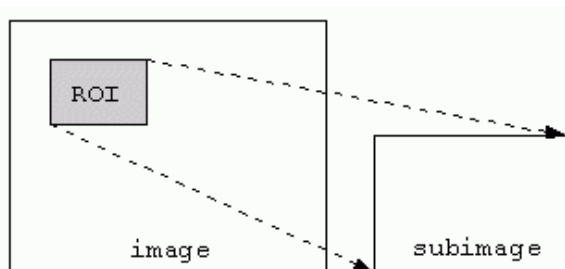


Рисунок 3.1 – Область яка цікавить на зображенні

Іноді необхідно скопіювати саму матрицю, OpenCV забезпечує функції клон (`clone()`) і `CopyTo()`.

Перед зберіганням значення пікселів необхідно вибрати колірний простір і тип використовуваних даних. Кольорова палітра ставиться до того, як ми комбінуємо колірні компоненти для того, щоб кодувати заданий колір. Найпростіший з них є шкала сірого, де кольору в нашому розпорядженні чорно-білий. Поєднання них дозволяє створювати безліч відтінків сірого.

RGB є найбільш поширеним, як наші очі використовувати щось подібне. OpenCV використовує систему кольорів BGR. HSV і HLS розкладаються кольору в їх колірний тон, насиченість і значення компонентів / яскравості, яке є більш природним для нас, щоб описати колір.

Перед безпосереднім детектуванням необхідно перетворити зображення в градації сірого. Для виконання даної операції в OpenCV передбачено метод `cvtColor()`.

Даний метод отримує три аргументи:

- вихідне зображення;
- зображення після перетворення (`gray_image`);
- додатковий параметр – вказує на те, який вид трансформації буде виконуватись. Для перетворення зображення в градації сірого необхідно використати – `CV_BGR2GRAY`.

Щоб виконати безпосереднє детектування необхідно викликати метод `detectMultiScale`, в який необхідно передати такі аргументи:

- `image`: вихідне зображення;
- `objects`: призначено для зберігання визначених об'єктів;
- `scaleFactor`: задається масштаб зображення. Цей масштабний коефіцієнт використовується для створення шкали піраміди. Припустимо, масштабний коефіцієнт 1,03, це означає, що ми використовуємо невеликий крок для зміни розміру, тобто зменшити розмір на 3%;

- `minNeighbors`: параметр, що визначає, як багато сусідів кожне виявлене обличчя повинно мати, щоб зберегти його. Цей параметр буде впливати на якість виявлених осіб: більш високі значення зменшується кількість виявлень, але з більш високою якістю;

- `minSize`: мінімально можливий розмір об'єкта. Об'єкти менші за нього ігноруються;

- `maxSize`: Максимально можливий розмір об'єкта. Об'єкти, що більші за нього, ігноруються [10].

3.2 UML модель системи

Розглянемо модель програмної системи що розробляється в термінах діаграм UML. UML може бути застосовано на всіх етапах життєвого циклу аналізу бізнес-систем і розробки прикладних програм. Різні види діаграм які підтримуються UML, і найбагатший набір можливостей представлення певних

аспектів системи робить UML універсальним засобом опису як програмних, так і ділових систем.

Діаграми дають можливість представити систему у такому вигляді, щоб її можна було легко перевести в програмний код. Основною причиною використання мови UML є спілкування розробників між собою.

Діаграма прецедентів має одного актора – Користувача. Користувач має змогу користуватися продуктом як доповненням до Messages App, або як звичайною програмою.

Логіка роботи програми дещо різна в залежності від того де користувач знаходиться у системі (рис. 3.1).

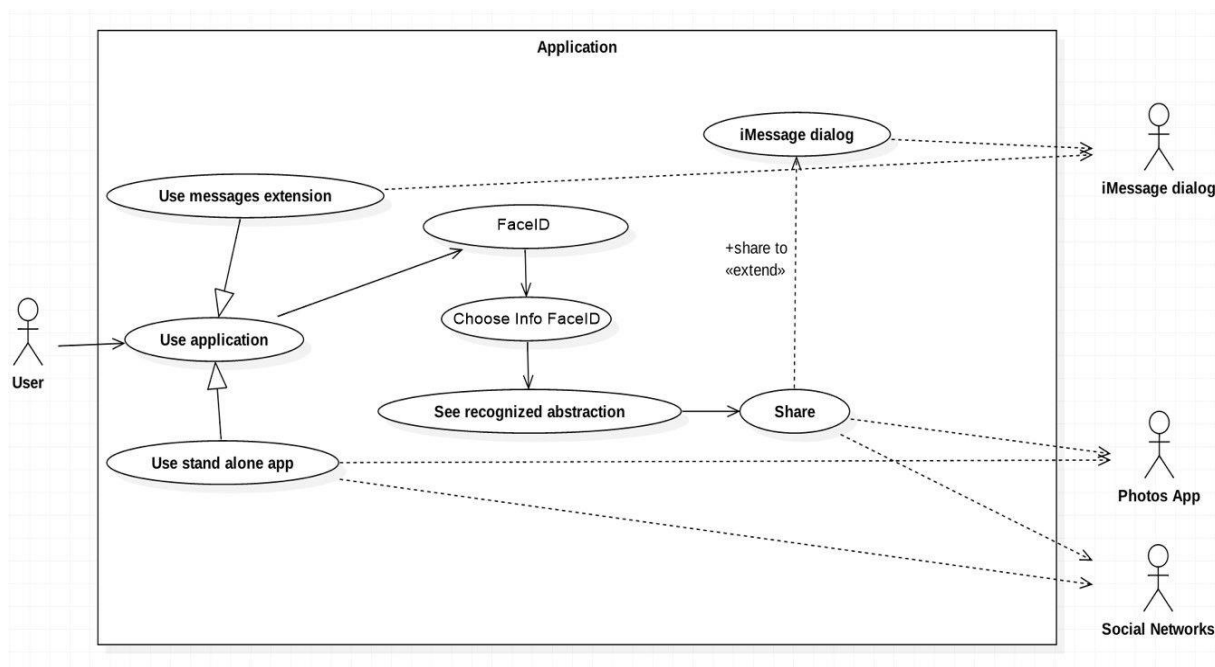


Рисунок 3.1 – Діаграма прецедентів

На діаграмі послідовностей наведена взаємодія всіх компонентів системи, які приймають участь у відображенні образу на екрані користувача. Діаграма послідовностей відображає взаємодію об'єктів у часі. На ній присутні тільки ті об'єкти які приймають участь у процесі взаємодії. Ключовий момент – це динаміка взаємодії об'єктів у часі.

В UML діаграма послідовності має виміри. Перший зліва направо у вигляді вертикальних ліній, кожна з яких зображує лінію життя окремого об'єкта, який

бере участь у взаємодії. Крайнім зліва на діаграмі зображується об'єкт, який є ініціатором взаємодії.

Праворуч зображується інший об'єкт, який безпосередньо взаємодіє з першим. Таким чином, всі об'єкти на діаграмі послідовності утворюють деякий порядок, який визначається черговістю або ступенем активності об'єктів при взаємодії один з одним.

Графічно кожен об'єкт зображується прямокутником і розташовується у верхній частині своєї лінії життя. У середині прямокутника записуються ім'я об'єкта і ім'я класу розділені двокрапкою. При цьому весь запис підкреслюється, що є ознакою об'єкта.

Другим виміром діаграми послідовності є вертикальна тимчасова вісь, спрямована зверху вниз. Початкового моменту часу відповідає сама верхня частина діаграми. Взаємодії об'єктів реалізуються за допомогою повідомлень, які надсилаються одними об'єктами іншим. Повідомлення зображуються у вигляді горизонтальних стрілок з іменами повідомлення, а їх порядок визначається часом виникнення. Тобто, повідомлення, що розташовані на діаграмі послідовності вище, ініціюються раніше тих, які розташовані нижче. Масштаб на осі часу не вказується, оскільки діаграма послідовності моделює лише тимчасову упорядкованість взаємодій типу «раніше-пізніше» (рис. 3.2).

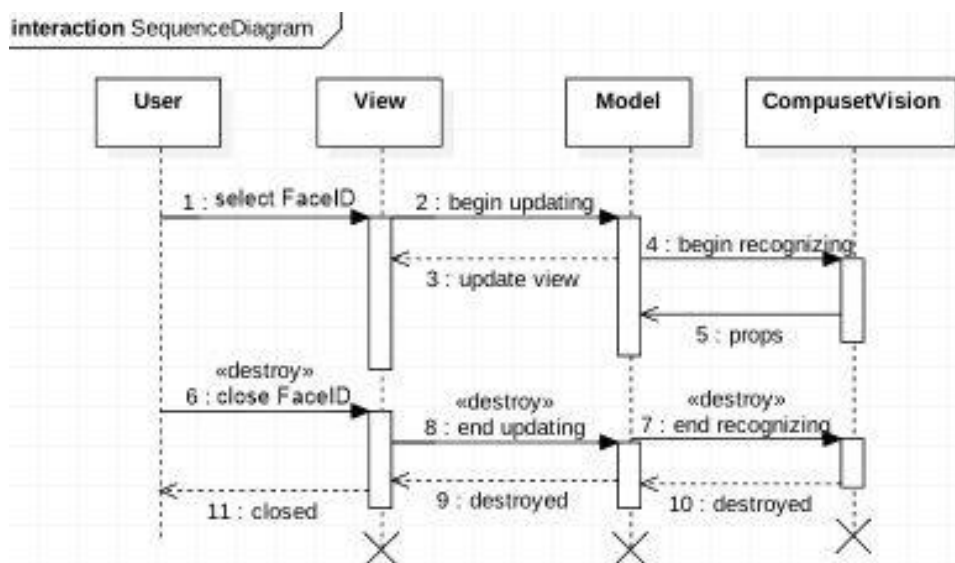


Рисунок 3.2 – Діаграма послідовностей

На діаграмі станів зображений процес визначення обличчя (рис. 3.3).

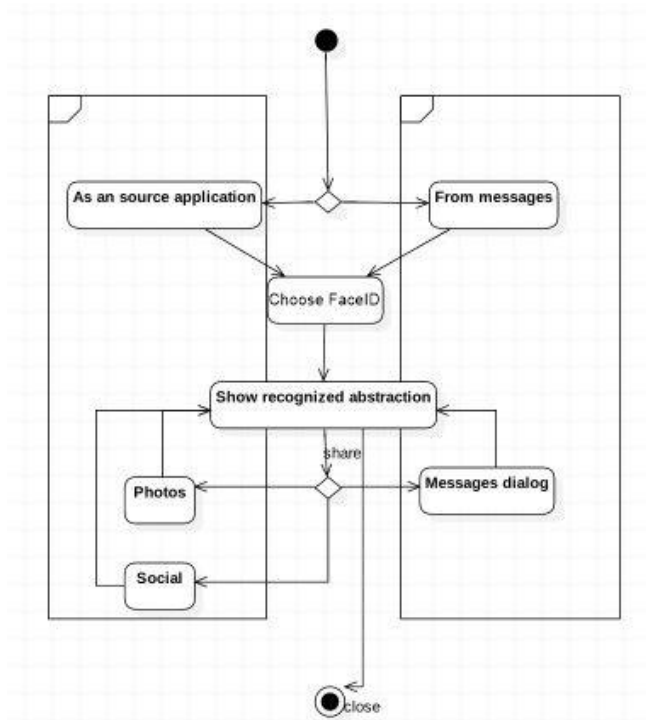


Рисунок 3.3 – Діаграма станів

Перейдемо до розробки структури програмної системи, базуючись на функціональних вимогах до системи.

Плануються, що даний продукт буде працювати на платформах Apple, тому питання архітектури посідає далеко не останнє місце. Ідея полягає у абстрагуванні бізнес логіки розпізнавання від платформи в окремий модуль, якій працюватиме як щось, що може приймати послідовність кадрів та віддавати цікаві нам параметри: `Func(video stream) ! FaceParams`.

Тому програму умовно можна розділити на декілька модулів:

- модуль отримання відеопотоку;
- модуль аналізу відеопотоку;
- модуль виводу зображення;
- модуль графічного інтерфейсу користувача.

Гнучкість рішення полягає у тому, що найважливіші модулі будуть платформи-незалежними, що дає змогу повторно використовувати їх.

3.4 Приклади алгоритмів

3.4.1 Реалізація захвату точок інтересу за допомогою OpenCV:

```

-(void)performOpticalFlowCalculationWithGray:(cv::Mat &)gray
        prevGray:(cv::Mat &)prevGray
        points:(vector<Point2f>&)points
        prevPoints:(vector<Point2f>&)prevPoints
        status:(vector<uchar>&)status
        inRect:(cv::Rect)rect
        scale:(float)scale {

TermCriteria termcrit(CV_TERMCRIT_ITER | CV_TERMCRIT_EPS, 20,0.3);

if (prevPoints.empty()) {
    //Feature detection is performed here...
    goodFeaturesToTrack(gray, points, kMaxPointsCount,
                        0.01, 45, Mat(), 3, 0, 0.04);
    cornerSubPix(gray, points, kSubPixWinSize,
                 cv::Size(-1, -1), termcrit);
    points = [self points:points byFilteringInRect:rect];
} else {
    Mat err;
    calcOpticalFlowPyrLK(prevGray, gray,
                         prevPoints, points,
                         status, err);
    cornerSubPix(gray, points, kSubPixWinSize,
                 cv::Size(-1, -1), termcrit);
}

std::swap(points, prevPoints);
cv::swap(prevGray, gray);

```

```
}

```

3.4.2 Реалізація знаходження положення голови у просторі:

```
- (void)matrixWithModelPoints:(vector<Point3f>)modelPoints
    imagePoints:(vector<Point2f>)imagePoints
    inImage:(cv::Mat &)image {

    float focal_length = MAX(image.cols, image.rows); //
Approximate focal length.
    Point2f center = cv::Point2f(image.cols / 2, image.rows /
2);

    cv::Mat cameraMatrix = (cv::Mat_<double>(3,3) <<
focal_length, 0, center.x,
                                0, focal_length, center.y,
                                0, 0, 1);

    cv::Mat
distCoeffs(cv::Mat::zeros(4,1,cv::DataType<double>::type));
    cv::Mat
rvec(cv::Mat::zeros(3,1,cv::DataType<double>::type));
    cv::Mat
tvec(cv::Mat::zeros(3,1,cv::DataType<double>::type));

    cv::solvePnP(modelPoints, imagePoints, cameraMatrix,
distCoeffs, rvec, tvec, false, CV_ITERATIVE);

#ifdef DEBUG
    vector<Point2f> projection;
    projectPoints(modelPoints, rvec, tvec, cameraMatrix,
distCoeffs, projection);
    for (int i = 0; i < projection.size(); i++) {
        cv::line(image, projection[i], imagePoints[i],

```

```

cv::Scalar(255,0,0), 2);
    }
#endif

    [self calcModelviewMat:_modelView
        fromRotVec:rvec
        transVec:tvec];

    [self calcProjectionMat:self.projection
        fromCameraMat:cameraMatrix
        withScreenSize:CGSizeMake(image.cols, image.rows)
        znear:50
        zfar:5000];
}

```

3.4.3 Реалізація початкового стану голови, проекція точок інтересу:

```

- (vector<cv::Point3f>)projectionWithRect:(cv::Rect) faceRect
    points:
(vector<cv::Point2f>)points {
    std::vector<Point2f> origin (points.size());
    std::transform(points.begin(), points.end(),
        origin.begin(), [&faceRect](Point2f point) {
        Point2f topLeft = faceRect.tl();
        return point - topLeft;
    });
    return [SSModelBuilder projectPoints:origin
        withSize:faceRect.size()
        targetSize:cv::Size(1000, 1000)];
}

```

3.4 Опис програмних рішень

З метою оцінки запропонованої методики використання системи розпізнавання за обличчям було створено прототип програмної системи, результати тестування якого дозволяють зробити висновки щодо можливості використання методу.

Програмну систему для платформи iOS ver.11.0 було реалізовано з використанням системи контролю залежностей. CocoaPods [10] є менеджером залежностей для Swift та Objective-C проектів. Вона налічує понад 32 тисяч бібліотек і використовується в більш ніж 2,1 мільйона додатків. Також CocoaPods надає допомогу в масштабуванні проектів.

Файл виглядає наступним чином:

```
platform :ios, '11.0'

target 'faceIT' do
  use_frameworks!

  # Pods for faceIT
  pod 'RxSwift', '~> 4.0'
  pod 'RxCocoa', '~> 4.0'
  pod 'AsyncSwift', '~> 2.0'
  pod 'PKHUD'

end
```

Щоб скористатися, достатньо перейти до каталогу з проектом та виконати команду `pod install`. Після цього автоматично буде згенеровано workspace з двох проектів: основного проекту та проекту зі сторонніми залежностями.

OpenGL та SceneKit. В якості фреймворку для забезпечення рендерингу було обрано SceneKit, який є дефолтним. Це достатньо потужний фреймворк для рендерингу, що може взаємодіяти з OpenGL контекстом, до якого забезпечується доступ на переважній більшості платформ, зокрема в iOS та MacOS. Це дозволяє дуже гнучко інтегрувати рішення до екосистеми Apple.

Комп'ютерний зір OpenCV. OpenCV (Open Source Computer Vision Library), що поширюється за ліцензією BSD і є безкоштовним як для наукових так і комерційних цілей. Він забезпечує інтерфейси C++, C, Python і Java, а також підтримує Windows, Linux, Mac OS, IOS та Android. Бібліотеку OpenCV було розроблено з метою забезпечення обчислювальної ефективності та з акцентом на додатках реального часу. Написана з використанням оптимізованих мов програмування C та C++, бібліотека може використовувати багатоядерну архітектуру процесора. Включно з OpenCL, вона може використовувати апаратне прискорення базової гетерогенної обчислювальної платформи.

Інструментальні засоби розробки. В якості інструментальних засобів було обрано нативні інструменти iOS: Augmented Reality (ARKit), Object Detection (Vision) та Machine Learning (CoreML).

ARKit – інструмент, який вміє обробляти велику кількість даних, отриманих від пристрою. Завдяки камері і датчикам руху фреймворк відстежує рух, знаходить поверхні і визначає освітленість. Після аналізу даних отримуємо конкретне уявлення про навколишній світ у вигляді точок перетину, координат поверхонь і положенні камери в просторі.

Основним завданням ARKit є спостереження за навколишнім світом (World Tracking) для створення віртуальної моделі реального світу. Фреймворк розпізнає особливості відеокадрів, відстежує зміни їхнього економічного становища і порівнює цю інформацію з даними від датчиків руху. Результатом є віртуальна модель реального світу. Стеження за навколишнім світом вимагає аналізу картинки, одержуваної від камери.

Основою ARKit є ARSCNView і ARSKView. Вони служать для відображення live-відео та рендерингу 3D і 2D зображень.

Програма складається з чотирьох модулів, ієрархію яких наведено на рис. 3.4. Кожен з модулів містить певний набір функцій для вирішення поставлених завдань. Перелік функцій та їх призначення наведені в додатку Б. В якості мови програмування для реалізації алгоритму було обрано Objective-C. Визначення координат області обличчя особи і центру очей на зображенні реалізовано із застосуванням бібліотеки OpenCV. Бібліотека містить реалізацію методу Віолі-Джонса з навченими класифікаторами на визначення обличчя та очей на зображенні.

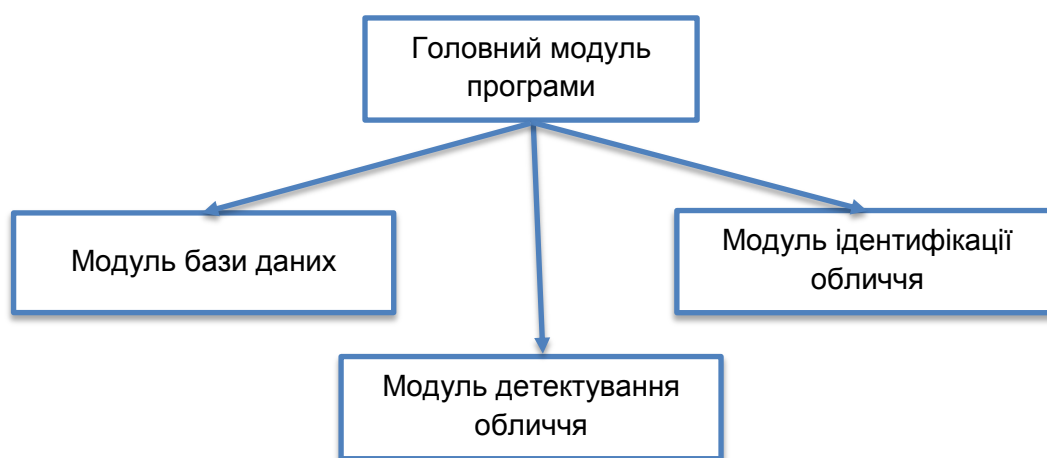


Рисунок 3.4 – Ієрархія модулів програми

Для першого запуску програми необхідно сформувати навчальну вибірку. З цією метою у якості базової моделі було обрано відеоряд, фрагмент якого наведено на рис. 3.5.

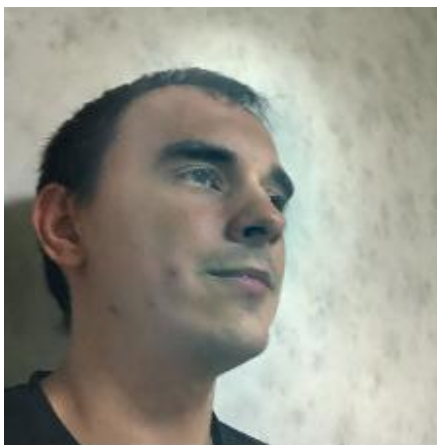


Рисунок 3.5 – Обличчя для навчання

В роботі використано алгоритм Eigenface, в основі якого лежить метод головних компонент (МГК), та який є найбільш ефективним методом на сьогоднішній день.

Ідея методу полягає в тому, що матрицю зображення можна представити у вигляді одновимірного вектору, помістивши другий стовпець під першим, третій – під другим і т.д. Отримані вектори лежать в просторі, що має дуже велику розмірність, а необхідні дані належать меншій розмірності, тому необхідно знайти оптимальне простір, в якому можна виявити і описати індивідуальні особливості кожної особи. Для вирішення цього завдання використовуються головні компоненти розподілу осіб, які представляють собою власні вектори коваріаційної матриці набору зображень.

Опрацювання зображень відбувається у декількох етапів:

Етап 1. На вхід системи надходить тестове зображення обличчя людини в форматі RGB, яке проходить етап попередньої обробки, що включає в себе: переведення зображення в градації сірого, використання комбінованого алгоритму шумозаглушення для усунення діючих перешкод і спотворень, еквалізація гістограми;

Етап 2. Визначення локальної області зображення обличчя людини по складової яскравості. Для цього використовується алгоритм Віоли – Джонса, який для пошуку особи на складному зображенні застосовує примітиви Хаара, алгоритм AdaBoost і каскади класифікаторів;

Етап 3. Виділення і кадрування особи із загальної сцени для того, що б прискорити процес на етапі розпізнавання;

Етап 4. Формування вектору ознак особи на основі алгоритму «власних осіб» (Eigenface). Ознаки формуються як для тестового зображення, так і для навчальної вибірки;

Етап 5. Застосування класифікатора, який в області векторів ознак здійснює ідентифікацію особи на зображенні.

Вектори власних ознак упорядковуються у відповідності до величини власного значення. Кожен з цих векторів описує особливості обличчя людини і їх можна розглядати як набір характерних ознак. Так само кожна ділянка зображення вносить вклад в кожен з векторів, що дозволяє представляти вектори як деякі наближені до зображення особи, тому їх називають власними особами (Eigenface). Кожне зображення особи подається у вигляді каскаду комбінацій ознак Хаара власних осіб (рис.3.6).

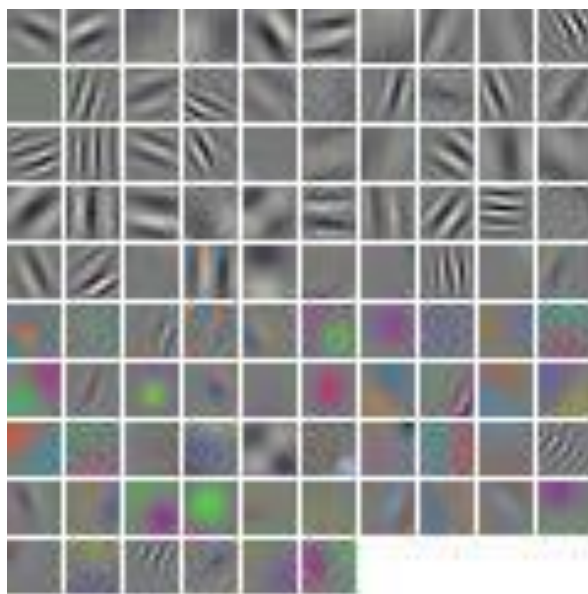


Рисунок 3.6 – Формування каскадного класифікатора за ознаками Хаара

З рис.3.6 видно, що формування класифікатора базується на методі Віоли-Джонса. На етапі виявлення в методі Віоли-Джонса вікно встановленого розміру рухається по зображенню, і для кожної області зображення, над якою проходить вікно, розраховується ознака Хаара. Наявність або відсутність предмета у вікні визначається різницею між значенням ознаки і навченим порогом. Оскільки ознаки Хаара мало підходять для навчання або класифікації (якість трохи вища, ніж у випадковій нормально розподіленій величині), для опису об'єкта з достатньою точністю необхідно більше число ознак. Тому в методі Віоли-Джонса ознаки Хаара організовані в каскадний класифікатор. Сформована на підставі класифікатора вибірка власних ознак (рис.3.7) надає

можливість сформувати набір для подальшого розпізнавання за обличчям (результати проведення експериментів наведено у додатку Б).

3.5 Аналіз отриманих результатів

Для доведення досягнення мети дослідження було проведено тестові випробування системи (додаток Б). Пошук та виділення обличчя на зображенні здійснювалось в режимі реального часу у відеопотоку. Для підвищення швидкодії алгоритму процедури визначення обличчя та сегментації зображення проводились на хмарному сервісі Amazon. Результат формування матриці власних зображень наведено на рис.3.7.

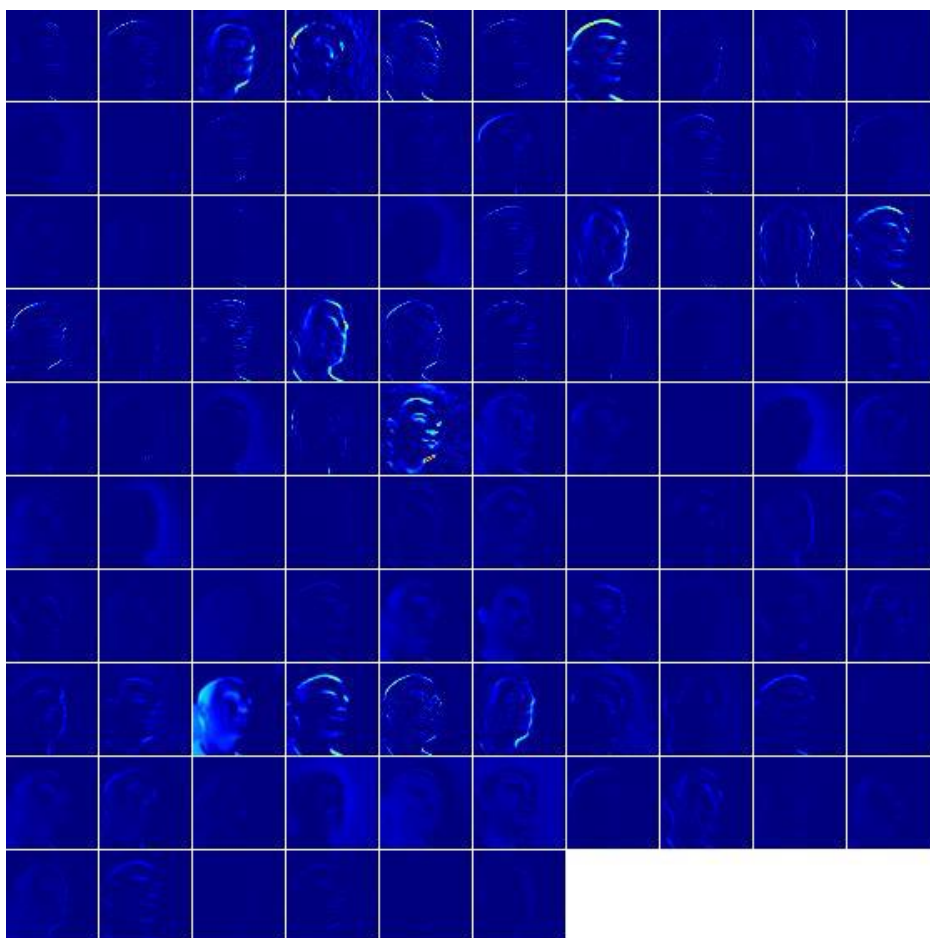


Рисунок 3.7 – Формування матриці власних зображень

Для підвищення якості зображення та покращення результату розпізнавання наступним етапом була нормалізація зображень, результат якої представлено на рис. 3.8.

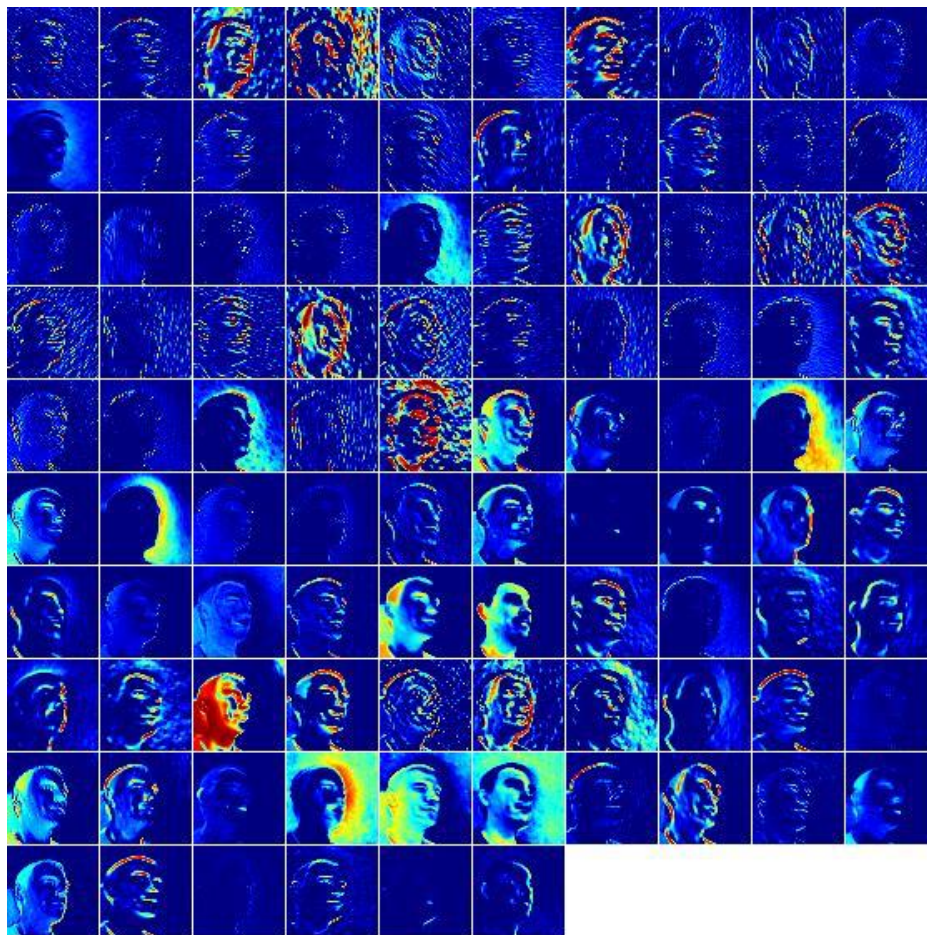


Рисунок 3.8 – Результат нормалізації зображень

Для аналізу отриманих нормалізованих зображень використовувалися характеристики, які оцінювалися шляхом визначення помилок першого та другого роду.

При помилці першого роду людське обличчя на зображенні не розпізнається (пропускається) системою, тобто виникає помилковий пропуск.

Помилка другого роду відбувається, коли зображення особи приймається за об'єкт іншого класу – виникає хибне виявлення.

В системі, яку було реалізовано в рамках виконання атестаційної роботи магістра, під час проведення тестових експериментів кількість помилок 1-го та 2-

го роду дорівнювала 0%, хоча це у своїй більшості свідчить про недостатню кількість проведених експериментів за різних умов.

Для забезпечення працездатності алгоритму було використано наступні інструментальні засоби:

- EC2 Instance Amazon (g2.2xlarge). Сервіс Amazon EC2 надає широкий вибір типів інстансів, оптимізованих для різних прикладів використання. Типи інстансів включають різні комбінації таких компонентів, як CPU, пам'ять, сховище і мережеві можливості, що дозволяє вибрати відповідний набір ресурсів для додатків. Кожен тип інстансів включає один або кілька розмірів інстансів, що дозволяє масштабувати ресурси відповідно до вимог цільової робочого навантаження;

- 1000 зображень (500 персональних фото та 500 фото для категорії «інше») для навчання алгоритму та перевірки на спрацювання;

- MacOS аплікація для знаходження та вирізки обличчя на зображеннях;

- натренована модель розпізнавання обличчя AlexNet;

- модель покращення розпізнавання FaceDetection_CNN;

- Python Script для трансформації сформованої та оптимізованої моделі до iOS CoreML формату;

- інтеграція сформованої моделі до iOS додатку.

У разі успішного розпізнавання особи, інформація про яку присутня в БД, вона виводиться у плаваючому вікні поряд з розпізнаною особою. Результат роботи додатку, що реалізує описаний алгоритм розпізнавання осіб та їх ідентифікації в середовищі iOS, наведено на рис.3.9: фрагмент початкового зображення (рис.3.9, а) з відеоряду та результат розпізнавання та спрацювання FaceID з підвантаженням інформації з БД (рис.3.9, б).

Базуючись на отриманих результатах можна зробити наступні висновки.

В атестаційній роботі магістра запропоновано та описаний комбінований алгоритм на основі спільного застосування методу Віюли-Джонса з попередньою обробкою зображення та її класифікацією у відповідності до каскадного класифікатора за ознаками Хаара, методу Eigenfaces для розпізнавання осіб на

цифрових зображення низької роздільної здатності за умови наявності шумів та в режимі реального часу.



а)

б)

Рисунок 3.9 – Результат роботи алгоритму з ідентифікацією особи:

а) – початкове зображення; б) – результат розпізнавання

З метою доведення надійності роботи алгоритму проведені експериментальні дослідження, які підтверджують гіпотезу щодо можливості використання хмарного кластеру для аналізу зображень та формування вибірки власних зображень для підвищення швидкодії розглянутого алгоритму на платформі Apple. На підставі аналізу отриманих результатів можна зробити висновки про успішну роботу створеного алгоритму і програми виділення і розпізнавання осіб.

ВИСНОВКИ

Атестаційна робота магістра присвячена дослідженню можливостей розпізнавання образів за обличчям на платформі Apple. В рамках проведення дослідження було проаналізовано предметну область, проведено аналіз сучасних методів реалізації розпізнавання за обличчям, виконано постановку задачі дослідження. На підставі проведеного аналізу програмних бібліотек комп'ютерного зору для використання в програмній системі було обрано OpenCV та виконано проектування iOS-орієнтованої системи аналізу зображень в режимі реального часу з автоматичною класифікацією.

На основі проведеного аналізу існуючих методів і алгоритмів, для розпізнавання було обрано метод Віоли-Джонса, який є одним з кращих за співвідношенням показників ефективності розпізнавання та швидкості роботи.

Розроблено архітектуру програмної системи організації доступу до БД за визначеним FaceID з автоматичною класифікацією та виділено основні складові процесу розпізнавання особи: детектування обличчя, нормалізація, отримання властивостей із зображення та приведення їх до типу, який використовується для збереженої в базі даних інформації та отримання цієї інформації з бази даних.

Для детектування обличчя на зображенні було використано каскади Хаара.

Програмна система з підтримки ідентифікації об'єктів з використанням розпізнавання за обличчям в реальному часі для платформи Apple за своїми характеристиками випереджає аналогічні програмні продукти, оскільки частину опрацювання зображень було перенесено до хмарного сервісу Amazon, що підтверджує її перспективність. Він має кращі функціональні показники, а тому є конкурентоспроможним товаром на ринку. Існуючі переваги нової розробки дозволяють зробити висновки про можливе швидке поширення її на ринку програмного забезпечення.

Розроблена комп'ютерна система може використовуватись як самостійний продукт, так і служити основою для подальших більш складних розробок.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Горелик А. Л. Методы распознавания / Скрипкин В. А. – 4-е изд. – М.: Высшая школа, 1984, 2004. – 262 с.
2. Кручинин А. В. Гистограммы яркости в OpenCV. – Режим доступа: <http://recog.ru/library/opencv/opencvhist.pdf>.
3. Handbook of face recognition.: [Электронный ресурс]. – Режим доступа : [http://read.pudn.com/downloads142/ebook/618315/Handbook Of Face Recognition.pdf](http://read.pudn.com/downloads142/ebook/618315/Handbook%20Of%20Face%20Recognition.pdf). – Дата доступа : 15.03.2015.
4. Live Face Detection Based on the Analysis of Fourier Spectra.: [Электронный ресурс]. – Режим доступа : <http://www.nws-sa.com/biometrics/facial/SPIE2004.pdf>. – Дата доступа : 15.05.2019.
5. Multimodal Person Recognition using Unconstrained Audio and Video.: [Электронный ресурс]. – Режим доступа: <http://www.cs.columbia.edu/~jebara/papers/TR-472.pdf>. – Дата доступа : 15.05.2019.
6. Томилов С. В. Алгоритмы распознавания лица на базе библиотеки OpenCV – 2-е изд. – М.: Высшая школа, 2008. – 262 с.
7. Lienhart, R. An extended set of Haar-like features for rapid object detection / Maydt, J. ICIP02, 2002 – 250 p.
8. On the Computation of Motion from Sequences of Image.: [Электронный ресурс]. – Режим доступа : <http://www.dtic.mil/dtic/tr/fulltext/u2/a211479.pdf>. – Дата доступа : 15.05.2019.
9. Peter N. Belhumeur Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection / David J. Kriegman, 1997 – 64
10. Bradsky G. Learning OpenCV – O'Reilly / Kaehler A. – O'Reilly Media, 2008 – 580 с.
11. Face ID on the iPhone X: Everything you need to know about Apple's facial recognition.: [Электронный ресурс]. – Режим доступа:

<https://www.macworld.com/article/3225406/face-id-iphone-x-faq.html>. – Дата доступу: 25.05.2019.

12. Liveness detection using cross-modal correlations in face-voice person authentication.: [Електронний ресурс]. – Режим доступу: <http://staff.estem-uc.edu.au/html/MWagner/Papers/2005/IS052290.pdf>. – Дата доступу : 15.05.2019.

13. Форсайт, Д.А. Компьютерное зрение. Современный подход: пер. с англ. / Д.А. Форсайт, Ж. Понс. – М.: Издательский дом Вильямс, 2004. – 928 с.

14. Работа каскада Хаара в OpenCV в картинках: теория и практика.: [Електронний ресурс]. – Режим доступу: <http://habrahabr.ru/company/recognitor/blog/228195/>. – Дата доступу : 10.03.2019.

15. Методичні вказівки до виконання атестаційної роботи магістра за спеціальністю 121 – Інженерія програмного забезпечення (Освітня програма – «Програмне забезпечення систем», Освітня програма – «Інженерія програмного забезпечення») для студентів усіх форм навчання / Упор.: З.В. Дудар, В.І. Каук, Н.С. Лесна, І.А. Ревенчук, Л.Д. Самофалов, І.Ю. Шубін, – Харків: ХНУРЕ, 2019.– 55 с.