

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики
Кафедра інтелектуальних програмних систем

**Кваліфікаційна робота
на здобуття освітнього рівня бакалавра**

за спеціальністю 121 Інженерія програмного забезпечення

на тему:

ПРОГРАМА РОЗПІЗНАВАННЯ ЖЕСТІВ ТА ОБ'ЄКТІВ НА ФОТО

Виконав студент 4-го курсу
Георгій ГОЛОДОВ

(підпис)

Науковий керівник:
доцент, кандидат фіз.-мат. наук
Володимир ШЕВЧЕНКО

(підпис)

Засвідчую, що в цій роботі немає
запозичень з праць інших авторів без
відповідних посилань.

Студент

(підпис)

Роботу розглянуто та допущено до захисту
на засіданні кафедри інтелектуальних
програмних систем

«___» _____ 202_ р.,

протокол №14 від 28.05.2021

Завідувач кафедри

Олександр ПРОВОТАР

(підпис)

Зміст

РОЗДІЛ 1. МАШИННИЙ ЗІР ЯК НАПРЯМОК ШТУЧНОГО ІНТЕЛЕКТУ.	2
РОЗДІЛ 2. ПОСТАНОВКА ЗАДАЧІ.	4
РОЗДІЛ 3. ОГЛЯД ІСНУЮЧИХ РІШЕНЬ.	6
РОЗДІЛ 4. МЕТОДИ.	17
РОЗДІЛ 5: АЛГОРИТМ.	27
РОЗДІЛ 6. РЕАЛІЗАЦІЯ.	32
РОЗДІЛ 7. РЕЗУЛЬТАТ РОБОТИ.	40
РОЗДІЛ 8. ВИСНОВКИ.	41
Література	43

РОЗДІЛ 1. МАШИННИЙ ЗІР ЯК НАПРЯМОК ШТУЧНОГО ІНТЕЛЕКТУ.

1. Штучний інтелект. Розпізнавання образів.

Штучний інтелект – це напрям комп'ютерних наук який вивчає методи розв'язання задач, для яких не існує способів вирішення. Системи штучного інтелекту можуть оперувати даними та самонавчатися.

Розпізнавання образів - це віднесення вихідних даних до певного класу за допомогою виділення істотних ознак, що характеризують ці дані, із загальної маси несуттєвих даних.

Живі системи, включаючи людей, постійно стикалися із завданням розпізнавання зразків з моменту їх створення. Зокрема, інформація, що надходить від органів почуттів, обробляється мозком, який, у свою чергу, сортує інформацію, забезпечує рішення, а потім за допомогою електрохімічних імпульсів передає потрібний сигнал далі, наприклад, до органів руху, які виконують необхідні дії. Потім відбувається зміна навколишнього середовища, і вищевказані явища трапляються знову. Отже, кожен етап супроводжується визнанням.

З розвитком обчислювальної техніки стало можливим вирішити ряд завдань, що виникають в процесі життєдіяльності, полегшити, прискорити, підвищити якість результату. Наприклад, робота різних систем життєзабезпечення, взаємодія людини з комп'ютером, поява роботизованих систем та інших. Тим не менш, відзначимо, що забезпечити задовільний результат в деяких завданнях (розпізнавання подібних об'єктів, що швидко рухаються, рукописного тексту) на даний час не вдається.

Під образом мається на увазі деяка упорядкована сукупність ознак. Класом образів називається сукупність об'єктів з однаковими властивостями.

Розпізнавання образів значно спрощує взаємодію людини з комп'ютером. Базовим в розпізнаванні є невизначене поняття множини. У комп'ютері

множина представляється набором неповторюваних однотипних елементів. Слово "неповторюваних" означає, що якийсь елемент в множині або є, або його там немає. Універсальна множина включає всі можливі для розв'язуваної задачі елементи, порожнє не містить жодного. У класичній постановці задачі розпізнавання універсальна множина розбивається на частини-образи. Образ будь-якого об'єкта задається набором його приватних проявів. Методика віднесення елемента до якого-небудь образу зветься вирішальним правилом. Ще одне важливе поняття - метрика, спосіб визначення відстані між елементами універсальної множини. Чим менше ця відстань, тим більш схожим є те, що ми розпізнаємо. Від вибору представлення образів і реалізації метрики залежить ефективність програми, один алгоритм розпізнавання з різними метриками буде помилятися з різною частотою.

Мета комп'ютерного зору - розробка методів, які дозволять машині «розуміти» або аналізувати зображення і відео.

У роботі було проведено дослідження другого методу введення інформації – через жести людини. Під час дослідження було розроблено та впроваджено алгоритм розпізнавання жестів, отриманих із відеопотоку. Важливою вимогою було звести до мінімуму обмеження на вхідне відео та записуючі пристрої, щоб полегшити подальше використання програми. В якості кінцевого контрольного набору жестів було вибрано підмножина міжнародної азбуки глухонімих.

Розпізнавання образів (об'єктів, сигналів, ситуацій, явищ або процесів) - завдання ідентифікації об'єкта або визначення будь-яких його властивостей за його зображенню (оптичне розпізнавання) або аудіозапису (акустичне розпізнавання) та іншим характеристикам.

Образ - класифікаційне угруповання в системі класифікації, яке об'єднує певну групу об'єктів за деякою ознакою. Образи мають характерну властивість, яка виявляється в тому, що ознайомлення з кінцевою кількістю явищ з одного і того ж набору дає можливість дізнатися якомога більший

кількості його представників. Образи характеризуються об'єктивністю властивостями в тому сенсі, що різні люди, що навчаються на різному матеріалі спостережень, здебільшого однаково і незалежно один від одного класифікують одні й ті ж об'єкти. У класичній постановці проблеми розпізнавання універсальний набір ділиться на частини-образи. Кожне відображення предмета на сприймаючих органах розпізнавальної системи, незалежно від його положення щодо цих органів, називається зображенням об'єкта, і багато таких зображень у поєднанні з деякими загальними властивостями є образами.

Причиною такого вибору стала поширеність і вільне використання цього набору жестів певною групою людей.

РОЗДІЛ 2. ПОСТАНОВКА ЗАДАЧІ.

Метою даної роботи була розробка і реалізація алгоритму, який отримує на вхід відеопотік, знятий з одної веб камери, розбиває його на окремі жести, кожен з яких класифікує в рамках обраної підмножини з абетки глухонімих. Формат вхідного відео було вирішено максимально наблизити до природних умов використання.

Вхідна інформація:

- У кадр може потрапляти вся людина, відео записується в зручному форматі, коли людина сидить перед відеокамерою і показує жести.
- Протягом запису відео можлива зміна освітлення.
- Знімати можна як при денному, так і при штучному освітленні.
- Час затримки на певному жесті може бути мінімальним, порядку пари секунд.
- На відео можливо шумове, фонове або рух, що повторюється на задньому плані.

При розробці програми основний упор робився на спрощення введення інформації в комп'ютер, тому однією з головних вимог була мінімізація обмежень на вхідне відео.

Для звукозаписуючої техніки було вирішено використовувати лише одну камеру. Формат відео повинен бути таким, ніби людина просто сіла перед камерою, вбудованою в комп'ютер або телефон, і почала показувати жести. Інтервали між жестами повинні бути мінімальними, щоб не затримувати потік введеної інформації.

Враховуючи те, що передбачалося записувати відео безпосередньо перед камерою, використання кольору шкіри як ознаки при локалізації долоні було неможливим через те, що світло, що надходить від монітора, могло суттєво спотворити колірну гамму відео.

В якості підсумкової множини жестів було вибрана підмножина латинської абетки глухонімих, що складається з 6-ти букв. Вибір був обумовлений природністю і поширеністю цих жестів.

РОЗДІЛ 3. ОГЛЯД ІСНУЮЧИХ РІШЕНЬ.

Методи аналізу даних

- **Фільтрація.**

До цієї групи належать методи, що дозволяють виділити необхідні області на зображеннях, не аналізуючи їх. Більшість із цих методів виконують єдине перетворення до всіх точок зображення. На момент фільтрації зображення не аналізується, але точки, що фільтруються, можна розглядати як ділянки зі спеціальними характеристиками.

Бінарізація за порогом.

Найпростішим і найпоширенішим перетворенням є бінарізація зображень за порогом. Для RGB та зображень у відтінках сірого граничним значенням є значення кольору.



Рисунок 1.3 – Демонстрація бінарізації за порогом.

- **Вейвлети**

Вейвлет-аналіз зображення - це пошук будь-якого патерна на зображенні за допомогою згортки з моделлю цього патерна. Існує набір класичних функцій, що використовуються при вейвлет-аналізі. Сюди входять вейвлет Хаара, вейвлет Морле, вейвлет мексиканських капелюхів тощо.

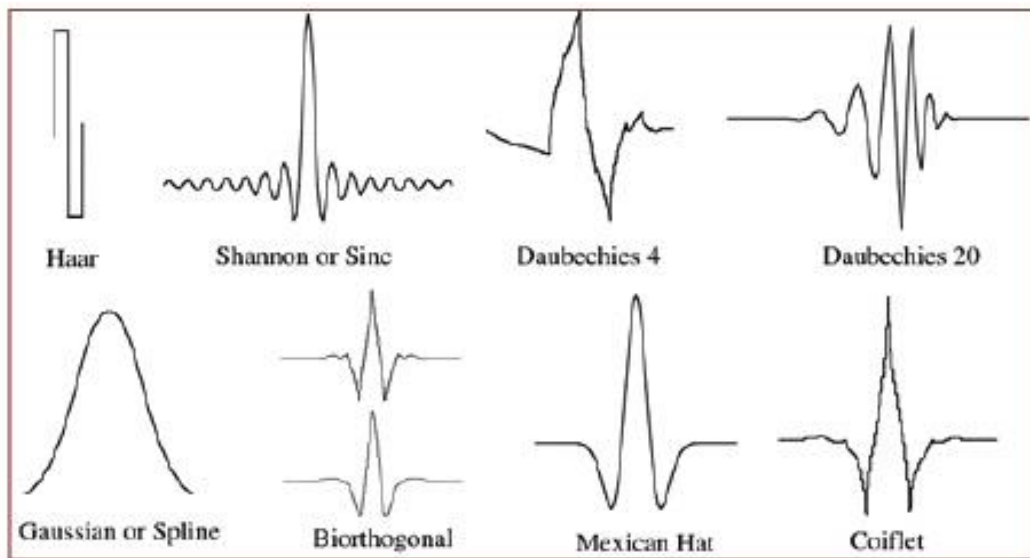


Figure 8

Examples of types of wavelets

Рисунок 1.4 – різновиди вейвлетів

- **Кореляція**

Отримання інформації про різність декількох зображень



Рисунок 1.5 – Приклад кореляції зображень

- **Контури**

Пошук обрисів об'єктів - це ефективний спосіб перейти від роботи над зображенням до роботи над графічним об'єктом. У тому випадку, якщо об'єкт має складну геометричну форму, але добре виділяється на задньому плані, тоді підбір контурів зазвичай повністю вирішує проблему фільтрації.

Існують наступні методи, які вирішують задачу фільтрації контурів:

- Оператор Кенні;
- Оператор Собля;
- Оператор Лапласа;
- Оператор Прюїтт;
- Оператор Робертса.

У більшості випадків задачу розпізнавання контурів можна вирішити саме оператором Кенні.

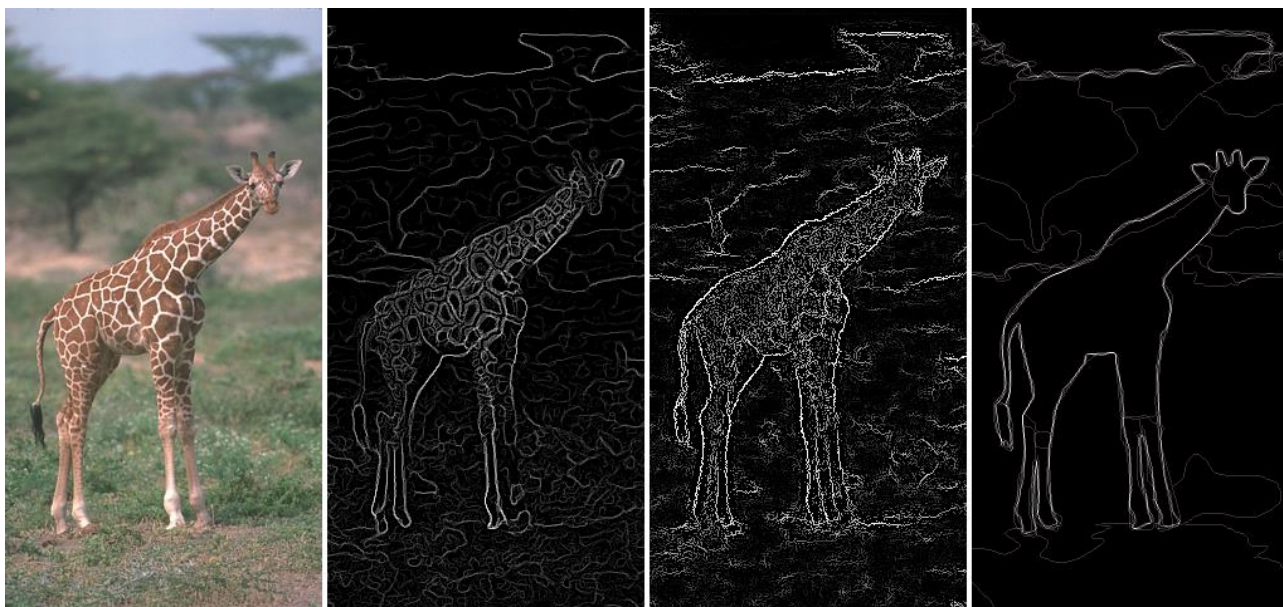


Рисунок 1.6 – приклад алгоритмів розпізнавання контурів

- **Логічна обробк результатів фільтрації**

Результатом фільтрації є набір корисних даних. Але їх часто важко використовувати без попередньої обробки.

- **Контурний аналіз**

Завданням контурного аналізу є отримання контуру зображення від межі зображення. Контури - це унікальні характеристики об'єктів, які дозволяють вирішувати класифікаційні задачі.

- **Особливі точки**

Спеціальні точки - це унікальні характеристики об'єкта, що дозволяють порівнювати об'єкт із самим собою або з об'єктами подібних категорій. Є багато способів виділити ці моменти. Деякі з них виділяють ці точки через сусідні кадри, а інші - через тривалі періоди часу та зміни освітлення. Певні алгоритми дозволяють знаходити спеціальні точки, які залишаються незмінними навіть при обертанні об'єкта. Типи особливих точок:

- **Особливі точки, які залишаються стабільними протягом секунд**

Ці точки використовуються для супроводу об'єкта у сусідніх відеокадрах або для вставки зображень із декількох камер. Ці точки включають локальні максимуми зображення, кути, точки з максимальною дисперсією, визначені градієнти тощо.

- **Особлива точка, яка залишається стабільною, коли змінюється світло і предмет трохи рухається**

Ці точки використовуються для навчання для подальшої класифікації типів об'єктів. Наприклад, класифікатор обличчя - це продукт системи, створеної в цих точках.

- **Стабільні точки**

Навіть якщо зображення обернено, ці точки залишаються стабільними. Алгоритми, що використовуються для пошуку таких точок, набагато складніші, їх існує менша кількість і вони зазвичай запатентовані.

- **Навчання**

Визначивши об'єкт дослідження та вибравши стійкі точки або контури, використовуються методи, які можуть приймати рішення. Цей процес можна описати наступним чином:

Існує тестовий зразок, який містить кілька класів об'єктів. Для кожного зображення існує набір функцій, обраних певним алгоритмом.

Алгоритм навчання повинен створити модель, яка дозволяє аналізувати нове зображення та робити висновки про те, які об'єкти знаходяться на зображенні.

Алгоритми навчання включають нейронну мережу, регресію, кластеризацію k-засобів, класифікатор AdaBoost та інші алгоритми.

Існуючі алгоритми:

- **Локалізація жестів**

Перший алгоритм вибору рук на зображенні - це фільтрування за кольором шкіри. Цей метод полягає у виділенні пікселів кольору шкіри із зображення. Дані про колір шкіри можна отримати за кольором обличчя. Результати цього дослідження не є задовільними, оскільки всі дані тестів мають шкірні тони.

Стандартний каскад Хаара також включає класифікатори долонь, кулаків, рук тощо, але їх робота набагато гірша, ніж відповідні класифікатори обличчя. Варто зауважити, що каскади Хаара дуже чутливі до кута

зображення, тому неможливо використовувати їх для складання ручних класифікаторів у всіх можливих положеннях.

- **Виділення рук на основі аналізу рухів**

Для подальших досліджень було зроблено припущення, що мова жестів за своєю природою є рухом, та саме області руху є найбільш інформативними для аналізу відеоряду.

- **Оптичний потік та лінії Фарнебеку.**

Існує багато способів аналізу руху на зображенні. З точки зору розрахунків, найбільшою інформацією і складністю є розрахунок частоти оптичних кадрів. Основна ідея цього методу полягає в розрахунку зміщення пікселів на сусідніх зображеннях.



Рисунок 1.7 – приклад знаходження оптичного потоку за допомогою алгоритму Фарнебека

Розпізнавання жестів

- **Метод гістограм кольорів**

Ідея методу кольорової гістограми для порівняння зображень полягає в наступному. Весь набір кольорів ділиться на набір різних, що повністю охоплює підмножину зображення. Сформуйте гістограму для зображення, яка відображає частку кожного підмножини кольорів у колірній гамі зображення. Для порівняння гістограм вводиться поняття відстані між ними.

Коли кольори RGB діляться на яскравість, обчислюється інтенсивність кожного кольору. Значення результату, вкладене між 0 і 255, потрапляє в один з 16 інтервалів, що ділить діапазон можливих значень. Сума різниць між відповідними елементами гістограми використовується як відстань між гістограмами; деякі вдосконалення цього методу реалізовані шляхом обчислення відстані на основі поелементного порівняння гістограм сусідніх елементів. Цей метод найбільш ефективний для чорно-білих зображень.

Для кольорових RGB-зображень інший метод забезпечує найкращі результати, розділивши кольори RGB на поля. Кольоровий простір RGB вважається тривимірним кубом, кожна вісь відповідає одному з трьох основних кольорів (червоному, зеленому або синьому). У цьому випадку зображення RGB будь-якого кольору може бути представлено точкою куба. Для побудови гістограми кольорів кожна сторона ділиться на 4 рівних інтервалу, відповідно RGB-куб ділиться на 64 прямокутних паралелепіпеда. Гістограма зображення відображає розподіл точок RGB-простору, відповідних кольорам пікселів зображення. Як відстань між гістограмами використовується покомпонентна сума модулів різниці між ними. Незважаючи на граничну простоту підходу, він показує досить стабільні результати.

Основним недоліком методу кольорової гистограми є втрата інформації про просторове положення об'єктів. Абсолютно різні фотографії можуть мати схожі кольорові гистограми. Наприклад, зображення осіннього листа може містити безліч дрібних плям червоного кольору. Це дасть схожу кольорову гистограму з великим червоним об'єктом.

Те, як елемент присвоюється зображенню, називається вирішальним правилом. Ще одне важливе поняття - метрика, спосіб визначення відстані між елементами універсальної сукупності. Чим менша відстань, тим більше схожі предмети (символи, звуки тощо) - те, що ми впізнаємо. Зазвичай елементи подаються як набір чисел, а метрика - як функція. Продуктивність програми залежить від вибору подання зображення та реалізації метрики, один алгоритм розпізнавання з різними метриками буде блукати з різною частотою. Навчанням зазвичай називають процес розвитку в системі реакцій на групи однакових зовнішніх сигналів шляхом неодноразового впливу на систему зовнішнього регулювання. Така зовнішня перебудова у навчанні називається "заохоченнями" та "покаранням".

Механізм генерації цього регулювання майже повністю визначає алгоритм навчання. Навчання відрізняється від навчання тим, що немає додаткової інформації про точність реакцій системи.

Адаптація - це процес зміни параметрів та структури системи та, можливо, управлінських впливів, заснований на поточній інформації для досягнення конкретного стану системи з початковою невизначеністю та зміною умов експлуатації.

Навчання - це процес, в результаті якого система поступово набуває здатності реагувати на необхідні реакції на конкретні набори зовнішніх впливів, а адаптація - це коригування параметрів і структури системи з

метою досягнення необхідної якості управління при постійних змінах зовнішніх умов.

Приклади задач розпізнавання образів:

- Розпізнавання букв;
- Розпізнавання штрих-кодів;
- Розпізнавання автомобільних номерів;
- розпізнавання осіб та інших біометричних даних;
- Розпізнавання мови

Як правило, існує три методи розпізнавання зразків:

Метод пошуку. У цьому випадку проводиться порівняння з базою даних, де для кожного типу об'єкта існують різні модифікації відображення.

Наприклад, для оптичного розпізнавання зображень можна використовувати метод сортування типу об'єкта під різними кутами, масштабами, зміщеннями, деформаціями тощо. Для букв потрібно сортувати шрифт, властивості шрифту тощо. Для звуку, малюнка розпізнавання, відповідно, здійснюється за деякими відомими зразками (наприклад, за словом, сказаним кількома людьми).

Другий підхід - більш глибокий аналіз особливостей зображення. У разі оптичного розпізнавання це може бути визначення різних геометричних ознак. У цьому випадку зразок звуку аналізується за частотою, амплітудою тощо.

Іншим методом є використання штучних нейронних мереж (НС). Цей метод вимагає або великої кількості прикладів завдання розпізнавання на навчанні,

або спеціальної нейромережевої структури, яка враховує специфіку завдання. Однак він характеризується вищою ефективністю та продуктивністю.

Дослідження розпізнавання жестів проводилося багато разів, але через численні відмінності у форматі вхідних даних та наборі впізнаваних жестів це питання не можна вважати вичерпаним. Роботи в цій галузі можна розділити на кілька груп.

Завдання розпізнавання мають наступні характерні риси.

Це інформаційні завдання, що складаються з двох етапів:

- Перетворення вхідних даних до виду, зручному для розпізнавання;
- Власне розпізнавання (вказівка приналежності об'єкта певного класу).

У цих завданнях ви можете ввести поняття аналогії або подібності об'єктів і сформулювати правила, на основі яких об'єкт зберігається в одному класі або в різних класах.

У цих завданнях може бути використаний набір прецедентів - прикладів, класифікація яких відома і які у вигляді формалізованих описів можуть бути представлені алгоритмом розпізнавання для адаптації до завдання в процесі навчання.

Для цих завдань важко будувати офіційні теорії та застосовувати класичні математичні методи (часто відсутність інформації для точної математичної моделі або вигоди від використання математичних моделей і методів непорівнянні із витратами).

Виділяють такі типи завдань розпізнавання :

- Завдання розпізнавання - віднесення пред'явленого об'єкта за його опису до одного з заданих класів (навчання з учителем) ;

- Завдання автоматичної класифікації - розбиття множини об'єктів , ситуацій , явищ за їх описами на систему непересічних класів (таксономія , кластерний аналіз , самонавчання) ;
- Завдання вибору інформативного набору ознак при розпізнаванні ;
- Завдання приведення вихідних даних до виду, зручному для розпізнавання;
- Динамічне розпізнавання і динамічна класифікація - завдання 1 і 2 для динамічних об'єктів;
- Завдання прогнозування - суть попередній тип , в якому рішення повинне ставитися до деякого моменту в майбутньому.

1. Системи розпізнавання динамічних жестів.

У ряді робіт за контрольну множину приймаються динамічні жести, що складаються з послідовних переходів руки з одного стану в інший. Атомарним жестом в даному контексті є рух руки вгору або вниз, наближення чи віддалення. Розпізнавання, в такому випадку, найчастіше засноване на прихованих Марковських ланцюгах, що і є головною причиною не застосовності досліджень до статичних жестів, оскільки базою для розпізнавання є послідовність змін, а не сам стан.

2. Використання додаткових інструментів.

В іншій групі досліджень для запису даних використовувалося інше обладнання. Це можуть бути кілька камер або пристроїв, які генерують додаткове випромінювання для отримання карти глибини. Згідно з додатковими даними, для розпізнавання будується тривимірне зображення. Отримані в цьому відношенні результати вражають, але в той же час поточні витрати програм, написаних за допомогою цих методів, зростають, а додаткові умови вхідних даних зменшують їх мобільність.

3. Орієнтування на колір шкіри.

Деякі успішні роботи висувають додаткові вимоги до вхідних даних: освітлення записаного відео не повинно спотворювати остаточну кольорову гаму. Це пов'язано з методом пошуку області зображення, де розташована долоня, яка відбирається за щільністю пікселів кольору шкіри. У повсякденних ситуаціях, наприклад при використанні вбудованої камери для запису відео, ця умова зазвичай порушується.

4. Використання маркерів.

Ще однією поширеною вимогою для введення відео є використання додаткових маркерів, які допомагають знаходити та моделювати долоню. У більшості випадків кожен палець рукавичок забарвлюється в інший колір.

РОЗДІЛ 4. МЕТОДИ.

Метод розпізнавання зразків можна розділити на дві категорії: навмисний та розширений. Відмінною особливістю навмисних методів є те, що вони використовують різні особливості ознак та їх взаємозв'язки як операційні елементи для побудови та застосування алгоритмів розпізнавання зображень. Цими елементами можуть бути одне значення або інтервал власних значень, середнє значення та дисперсія, матриця взаємозв'язку між ознаками тощо. Форма операцій, що виконуються над ними, є аналітичною або конструктивною. Отже, об'єкти в цих методах не розглядаються як необхідні інформаційні одиниці, а служать показниками для оцінки взаємодії та поведінки атрибутів. До таких методів відносяться:

- Метод оцінки щільності, заснований на розподілі власних значень (або подібності та відмінності об'єктів). Вони використовуються для вирішення проблем із відомими розподілами, як правило, звичайними розподілами, але вимагають збору та пошуку великих обсягів статистичних даних. Отже,

пошук усієї навчальної вибірки для розпізнавання може бути дуже великим, тому він буде швидко зростати в геометричній прогресії. Недоліком є висока чутливість до нерепрезентативних навчальних зразків та артефактів.

- Логічні методи можна використовувати лише для завдань з невеликим розмірним простором, що визначається високою евристичністю таких методів. При виборі правил логічного рішення (кон'юнктивних) потрібен повний пошук, що призводить до вищої обчислювальної складності.
- Лінгвістичні (структурні) методи. Вирішуються завдання відновлення (визначення) граматики за деякою множиною висловлювань (описів об'єктів). Такі завдання важко формалізувати, тому вони важко піддаються математичному представленню.

У методі розширення групи кожен об'єкт дослідження має більш-менш незалежне діагностичне значення. По суті, ці методи близькі до клінічних методів. Метод не розглядає об'єкт як запланований ланцюжок функцій, а як ціла система кожна система є незалежною і має особливе діагностичне значення. Цей метод не виключає і не втрачає інформацію про кожен об'єкт, що виникає при використанні навмисних методів керівництва, які використовують лише об'єкт для ідентифікації та запису поведінкових зразків його атрибутів.

Основна операція розпізнавання зразків із використанням розглянутого методу полягає у визначенні подібності та відмінності об'єктів. Об'єкти цього набору методів відіграють роль діагностичного прецеденту. При цьому, залежно від умов конкретного завдання, роль конкретних прецедентів може змінюватися в найширшому діапазоні: від первинної та детермінованої до дуже опосередкованої участі в процесі ідентифікації. У свою чергу, стан проблеми може вимагати участі різної кількості діагностичних прецедентів

для успішного вирішення: від одного об'єкта в кожній розпізнаній категорії до повного обсягу вибірки та різних методів обчислення подібності та різниці об'єктів. Ці вимоги пояснюють подальший поділ методів розширення на підкатегорії.

- Метод порівняння з прототипом. Його недоліком є висока залежність результатів класифікації від міри відстані (метрики) та неможливість створення заздалегідь визначеної оптимальної метрики.
- Метод к найближчих сусідів. Застосовується на завданнях невеликої розмірності за кількістю класів і ознак. Має високу залежність результатів класифікації від міри відстані (метрики). В завданнях такого типу виникає необхідність повного перебору навчальної вибірки при розпізнаванні, що веде до значної обчислювальної трудомісткості.
- Колективи вирішальних правил (КВП). Відзначається дуже висока технічна складність методу, невирішеність низки теоретичних проблем, як при визначенні областей компетенції приватних методів, так і в самих приватних методах.

У нашій роботі увага зосереджується на задачі розпізнавання відеопотоку та відповідному алгоритму. Розпізнавання відеопотоку має багато характеристик, тому реалізація цих методів має багато проблем:

- швидкість розпізнавання образу на зображенні є недостатньою для розпізнавання 25 кадрів за секунду;
- процес розпізнавання потребує значних затрат ресурсів оперативної пам'яті;
- більшість алгоритмів потребує великої навчальної вибірки позитивних і/або негативних зображень.

Таким чином, розглянемо існуючі методи виявлення об'єкта на цифровому відеоряді.

Методи розпізнавання об'єкта на цифровому зображенні, за способом віднесення ділянки зображення до області інтересів, можна розділити на дві великі групи: **узагальнюючі і розрізняючі**.

Метод узагальнення - це абстрактне та ідеалізоване уявлення про структуру об'єкта на основі позитивного навчального образу.

Метод цього методу полягає у використанні умовної моделі, яка може навчитися розрізняти частини об'єкта. Зображення розділене на невеликі частини, і кожна частина отримує ярлик, який представляє семантичне значення місцевості.

Метод Карпури-Вінна. Наприклад, у нас є теги: "ліс", "небо", "трава", "ціль". Знання місця розташування та класифікації деталі, пошук місця розташування інших деталей обмежений. Наприклад, якщо ви знаєте місце розташування дерева, ви можете легко визначити інші дерева. Цей метод використовує далекий радіус, обмежений простором на графіку, для виявлення об'єкта.

Метод Фергюсона-Перона заснований на тому, що досліджуване зображення виділяє ділянки з особливостями, а потім використовує навчену модель для порівняння з об'єктом або фоновою областю. Розрахувавши коефіцієнт вірогідності та порівнявши його з порогом, можна визначити наявність або відсутність об'єкта.

Метод Феррарі використовує перетворення Хафа. Зображення позначає характерну точку об'єкта, яка визначає радіус-вектор, що з'єднує центр

об'єкта. При пошуку об'єкта у відеопотоці радіус-вектор вказуватиме очікуване положення центру об'єкта. Тому кожного разу, коли виявляється точка, піксель в очікуваному положенні в центрі об'єкта додає голос. Піксель, що набрав найбільшу кількість голосів, буде відповідати найбільш вірогідному центру об'єкта.

Метод дискримінації використовує спеціальну функцію класифікатора для виявлення об'єктів на цифровому зображенні. У процесі навчання вони розрізняють позитивні та негативні навчальні образи.

Метод Віоли-Джонса. Це всебічне представлення зображень на основі Хаара. На основі адаптивного алгоритму підйому будується класифікатор, який дає логічну відповідь. Цей метод використовує каскад з декількох класифікаторів, що застосовуються послідовно до зображення. Кожен класифікатор включає набір характеристик Хаара з фіксованими відносними положеннями.

Метод Папагеоргіу заснований на використанні вейвлетів Хаара. Wavelet - це математична функція, яка дозволяє аналізувати різні частотні компоненти даних. Метод Papageorgiou обчислює проміжне представлення, яке інкапсулює важливу інформацію в класі об'єктів. Це уявлення обчислюється для всіх об'єктів, отриманих як вхідні дані зображення, що призводить до набору позитивних і негативних векторних ознак. Ці векторні ознаки використовуються для навчання класифікатора розрізнення об'єктів у своєму класі.

Порівнюючи розглянуті методи виявлення об'єктів, ми можемо помітити, що жоден із методів узагальнення не дозволяє проводити аналіз відеопотоків у реальному часі. Однак метод диференціації вимагає великої вибірки. Серед цих методів метод Віоли-Джонса є найшвидшим.

Для того, щоб досягти найкращого пошуку збігу цільового об'єкта та тегу потрапляння, ми розробили відповідний алгоритм, який враховує вищезазначені фактори.

Проведений аналіз методів, використовуваних при розпізнаванні жестів, дозволив розділити алгоритм на етапи і виділити кілька підходів, які підходять для досягнення найкращих результатів.

4.1 Віднімання фону на зображенні.

Принцип, що лежить в основі алгоритмів вирахування фону, полягає в тому, що певний кадр або група кадрів вибирається за еталонний фон. Усі наступні кадри порівнюються з ним і, якщо модуль різниці яркостей для пікселя перевершує поріг, то піксель вважається переднім планом, інакше - фоном. В роботі використовується адаптивний метод заснований на суміші нормальних розподілів.

Фонове значення для кожного пікселя представлено у вигляді набору трійок: $[w, \mu, \delta^2]$ ваги пікселя, математичного очікування і дисперсії, вага - є мірою схожості між значенням пікселя і фоном. Надалі, будемо називати таку трійку "процесом" пікселя. Таке уявлення дозволяє алгоритму адаптуватися до зашумлення фону. У поточній роботі для кожного пікселя запам'ятовується 5 його процесів, значення яких він приймає з плином часу. На першому кадрі відео відбувається ініціалізація моделі: $[I, c, \delta^2_{fix}]$, де c - яскравість пікселя, δ^2_{fix} - початкове значення.

Для кожного наступного кадру, для кожного пікселя проводиться ряд кроків:

1. Вимірюється відмінність між пікселем, що поступив і його фоновими значеннями: Якщо відмінність менше порога, то процес оголошується поточним. Якщо більше, то ініціалізується новий процес і, в залежності від кількості вже створених процесів, або додається в кінець набору,

або замінюється процес з найменшою вагою. Новий процес позначається як поточний.

2. Значення поточного процесу оновлюються:

$$\mu_t = (1 - \alpha_1) * \mu_{t-1} + \alpha_1 * c.$$

$$\delta_t^2 = (1 - \alpha_2) * \delta_{t-1}^2 + \alpha_2 * (c - \mu_t)^2.$$

$$w_t = (1 - \alpha_3) * w_{t-1} + \alpha_3.$$

Для всіх інших процесів пікселя оновлюється тільки вага:

$$w_t = (1 - \alpha_3) * w_{t-1},$$

де $\alpha_1, \alpha_2, \alpha_3$ - вибрані параметри алгоритму, що впливають на швидкість адаптації.

3. Піксель класифікується як фоновий, якщо вага перевищує обраний поріг і як що належить до переднього плану в іншому випадку.

4.2 Гістограма спрямованих градієнтів.

Гістограма спрямованих градієнтів є дескриптором, широко використовуваним в розпізнаванні образів. Основна ідея полягає в тому, що об'єкт можна задати за допомогою гістограм орієнтації градієнта на зображенні.

Алгоритм складається з декількох кроків:

1. Обчислення градієнта

На першому етапі обчислення дескриптора стоїть нормалізація кольору і гамма-корекція. Для чого проводиться згортка зображення з ядром, тобто

до кожного пікселя зображення застосовується маска і його значення замінюється на суму зважених значень пікселів околиці.

2. Угрупування напрямків

На другому кроці зображення ділиться на осередки, для кожної з яких проводиться підрахунок гістограми орієнтацій градієнтів. Форма осередків може бути як прямокутної, так і круглої форми, а весь діапазон напрямків рівномірно ділиться від 0 до 360 градусів, в разі, якщо обчислюється градієнт без знака, або від 0 до 180, якщо градієнт зі знаком.

3. Блоки дескрипторів

Наступний крок полягає в групуванні маленьких осередків в більш великі блоки, таке об'єднання потрібне для нормалізації значень осередків. Найчастіше після об'єднання осередки можуть входити одночасно в кілька різних блоків, і, як наслідок, гістограма буде нормалізована кілька разів.

4. Нормалізація блоків

На заключному етапі проводиться нормалізація блоків і об'єднання елементів гістограм блоків для отримання вектора дескриптора.

4.3 Машинне навчання.

За допомогою методів машинного навчання проводиться визначення приналежності жесту з вхідного потоку до однієї з обраних груп.

4.3.1 Постановка завдання класифікації.

Для класифікації об'єкта виділяється набір ознак, який вважається достатнім для ідентифікації класу об'єкта, набір ознак представляється у вигляді n -мірного вектора. Класифікатор може давати позитивний, в разі належності об'єкта до головної множини, або негативний результат, в іншому випадку.

Завдання класифікації є завданням з учителем, тому на вхід подається тренувальна множина з m векторів:

$$x^{(1)} = [x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}], x^{(2)}, \dots, x^{(m)}.$$

Кожному вектору ставиться у відповідність очікуваний результат класифікації: $y^{(1)}, y^{(2)}, \dots, y^{(m)}$.

4.3.2 Логістична рекурсія.

Метод логістичної регресії заснований на поданні класифікатора як параметрически заданої функції:

$$h_{\Theta}(x) = \frac{1}{1 + e^{-\Theta^T x}}, \text{ где } \Theta = [\Theta_1, \Theta_2, \dots, \Theta_n].$$

За допомогою тренувальної множини підбирається набір параметрів найбільш оптимальний для даної класифікації.

Визначається функція помилки і за допомогою методу градієнтного спуску знаходиться її мінімум:

$$J(\Theta) = \frac{1}{2m} \sum_{i=1}^n (h_{\Theta}(x^{(i)}) - y^{(i)})^2 \rightarrow \min, \text{ где } y^i \in \{0, 1\}.$$

4.3.3 Support Vector Machines

Класифікація в методі опорних векторів відбувається за допомогою поділу точок різних класів гіперплощиною. Таких гіперплощин може бути багато, тому в якості міри якості обраної являється зазор між класами. Якщо існує гіперплощина, що розділяє класи з максимальним зазором, то вона називається оптимальною розділяючою гіперплощиною, а відповідний їй лінійний класифікатор називається оптимально розділяючим класифікатором.

Будуємо розділяючу гіперплощину, яка має вигляд:

$w * x - b = 0$, де w - перпендикуляр що розділяє гіперплощини.

Окремо введемо поняття «помилка класифікації» для даного елемента - ξ_i . Результат класифікації, в такому випадку, може приймати два значення: $y_i \in \{-1, 1\}$. Згідно з цією класифікацією кожен вектор x повинен задовольняти такій умові:

$$y_i * (w * x_i - b) \geq 1 - \xi_i.$$

Для знаходження максимально можливого зазору між класами при такому лінійному розподілі необхідно мінімізувати таку функцію:

$$\frac{1}{2} * ||w||^2 + C * \sum_{i=1}^n \xi_i \rightarrow \min.$$

4.3.4 Нелінійна SVM.

Ідея класифікації за допомогою нелінійної SVM практично повністю повторює ідеї лінійної, з тією лише різницею, що кожний скалярний добуток замінюється нелінійною функцією ядра. У даній роботі в використовувалося ядро Гаусса:

$$k(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2 * \sigma^2}\right).$$

4.3.5 Мультикласифікатор.

Основою мультикласифікації був обраний алгоритм OneVsAll.

1. Необхідно провести класифікацію з m різними наслідками.розд
2. Для кожного $i \in \{1: m\}$ відбувається класифікація, де позитивним результатом вважається приналежність даного об'єкта до i -ому класу і негативним в іншому випадку.

3. В результаті виходить m -мірний вектор ймовірностей, приналежності до кожного класу, на основі якого і відбувається класифікація.

РОЗДІЛ 5: АЛГОРИТМ.

Вхідна інформація - відео потік з однієї вебкамери. Відео може містити різномірний фон і шумові рухи на задньому плані.

1. Попередня обробка зображення



Рис. 5.1: Початкове зображення



Рис. 5.2: Значення яскравості зображення



Рис. 5.3: Бінарне зображення, що розділяє вихідне на передній і задній план

Кожен кадр (Рис.5.1) переводиться в чорно-біле зображення виходячи із значення яскравості для пікселів (Рис.5.2) і за допомогою адаптивного алгоритму віднімається фон (Рис.5.3). На отриманій карті руху забирається шум і виділяються компоненти зв'язності для кожної з яких вважається вага.

2. Локалізація долоні



Рис. 5.4: Зміна активності руху на відео

Для кожної компоненти зв'язності на кадрі підраховується її вага до тих пір, поки значення не перевищить граничне. Вважаємо область найбільшої активності руху - долонею. На наступних кадрах, в залежності від того, чи є жест зафіксованим або відбувається його зміна, область долоні коригується (рис.5.4).

3. Поділ відео на жести

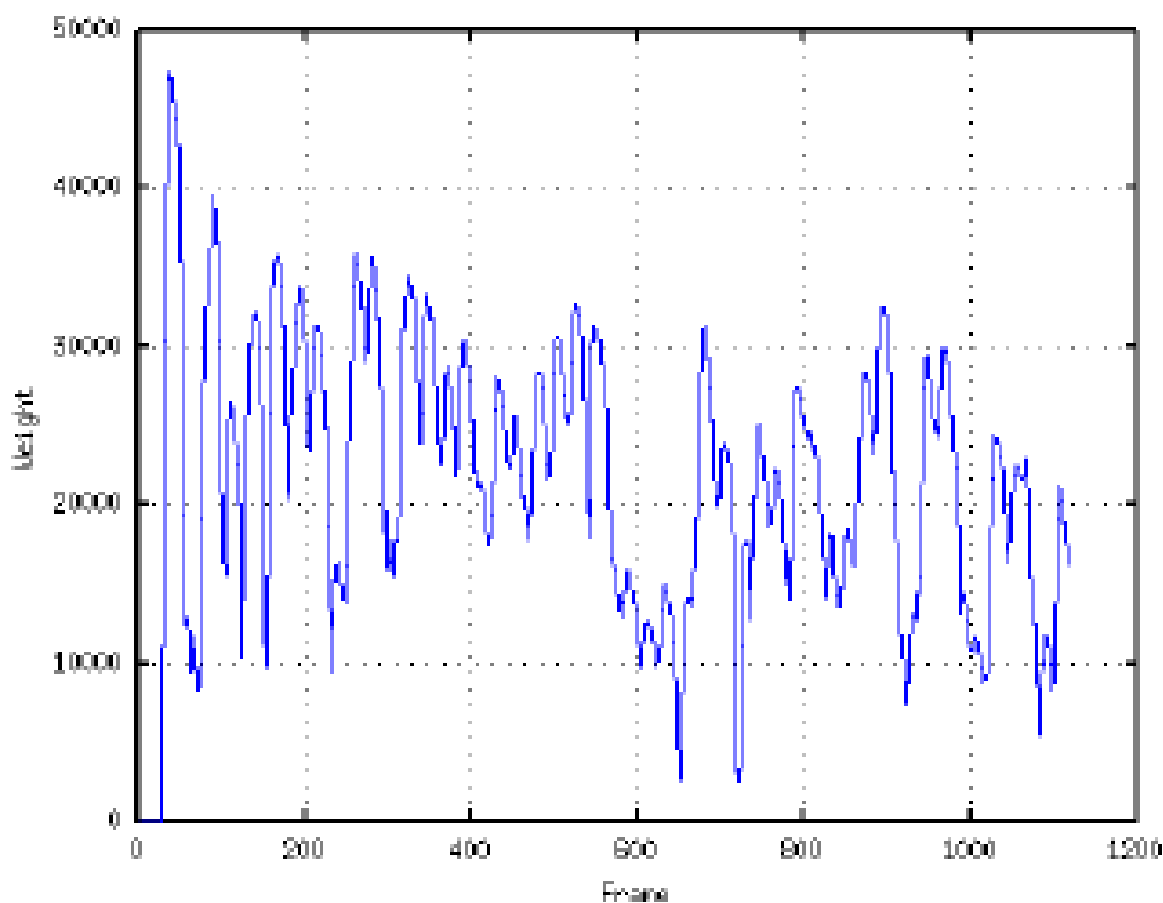


Рис. 5.5: Зміна активності руху на відео

В роботі використовується адаптивний метод вирахування фону і, як результат, різкі піки на графіку поступово згасають. Зростання активності руху є початком зміни жесту, а його загасання ознака того, що жест устоявся і його можна класифікувати (Рис. 5.5)

4. Отримання ознак жесту

В якості ознак для класифікації було використано два набори:

- *Моменти контуру і габарити долоні*

На фрагменті зображення, де була локалізована долоня за допомогою методу Канні обчислювалися контури.

Моменти - характеристика контуру обчислюється через підсумовування всіх його пікселів, обчислюється за формулою:

$$M_{ij} = \sum_{x=1}^{n_1} \sum_{y=1}^{n_2} x^i y^j I(x, y),$$
 де $I(x, y)$ - бінарне значення зображення, рівне 1, якщо контур проходить через піксель з координатами (x, y) і 0 в іншому випадку, i та j - параметри моменту.

Моменти які обчислюють за цією формулою залежать від вибору системи координат, тому в якості ознак було вирішено використовувати нормалізовану інваріантну модифікацію, яка обчислюється в кілька етапів.

На першому етапі проводиться централізація щодо центру мас:

$$\mu_{ij} = \sum_{x=1}^{n_1} \sum_{y=1}^{n_2} (x - x_c)^i (y - y_c)^j I(x, y),$$

где $x_c = \frac{M_{10}}{M_{00}}$, $y_c = \frac{M_{01}}{M_{00}}$ - центри масс

Отримані результати нормалізуються:

$$\nu_{ij} = \frac{\mu_{ij}}{M_{00}^{\left(\frac{i+j}{2} + 1\right)}}.$$

Моменти описують контур інваріантної обертання, відображення і масштабу:

$$\phi_0 = \nu_{20} + \nu_{02},$$

$$\phi_1 = (\nu_{20} - \nu_{02})^2 + 4\nu_{11}^2,$$

$$\phi_2 = (\nu_{30} - 3\nu_{12})^2 + (3 * \nu_{21} - \nu_{03})^2,$$

$$\phi_3 = (\nu_{30} + \nu_{12})^2 + (\nu_{21} + \nu_{03})^2,$$

$$\phi_4 = (\nu_{30} - 3 * \nu_{12}) * (\nu_{30} + \nu_{12}) [(\nu_{30} + \nu_{12})^2 - 3 * (\nu_{21} + \nu_{03})^2] + (3 * \nu_{21} - \nu_{03})(\nu_{21} + \nu_{03}) [3 * (\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2],$$

$$\phi_5 = (\nu_{20} - \nu_{02}) [(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2] + 4 * \nu_{11} * (\nu_{30} + \nu_{12})(\nu_{21} + \nu_{03}),$$

$$\phi_6 = (3 * \nu_{21} - \nu_{03})(\nu_{21} + \nu_{03}) [3 * (\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2] - (\nu_{30} - 3 * \nu_{12})(\nu_{21} + \nu_{03}) [3 * (\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2].$$

- *Гістограма спрямованих градієнтів*

Гістограма спрямованих градієнтів є дескриптором особливих точок, що задає об'єкт через інтенсивність зображення і напрямки меж об'єктів

5. Класифікація жесту

Кожен кадр в часовому проміжку, коли жест вважається усталеним, класифікується за допомогою методу опорних векторів, з усіх отриманих результатів класифікації вибирається той, який зустрівся більш, ніж в половині випадків, інакше жест вважається таким, що не належить контрольній множині.

РОЗДІЛ 6. РЕАЛІЗАЦІЯ.

6.1 Проектування

Якщо людина сидить безпосередньо перед нею, вхідні дані записуються за допомогою звичайної веб-камери. Важливою вимогою під час запису інформації є зручність використання цього методу, і використання вбудованої в комп'ютер камери повністю відповідає цій вимозі.

Однак запис відео поблизу можливого джерела світла ставить певні вимоги до подальшого розвитку: нам довелося відмовитись від пошуку однієї з найбільш очевидних ознак кольору долоні та шкіри, оскільки світло від монітора комп'ютера може спричинити серйозні спотворення кольору на відео.

В якості контрольного набору даних було вибрано підмножина мови глухонімих: A, C, V, G, H, Y

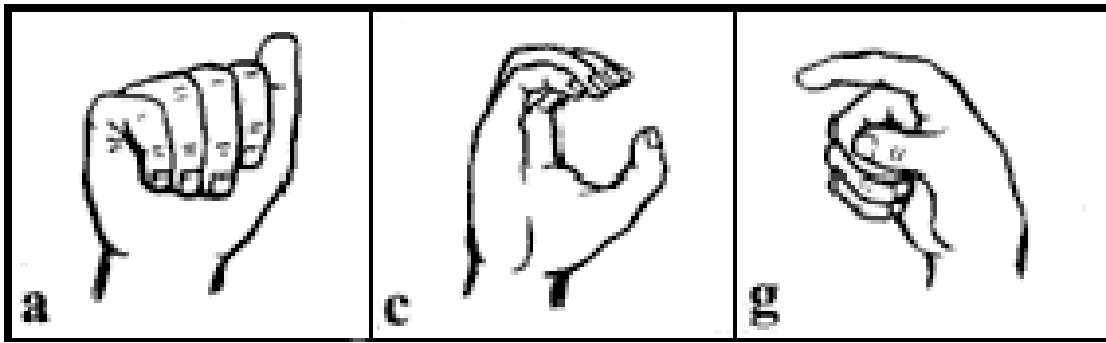


Рис. 6.1: Букви: А, С, G

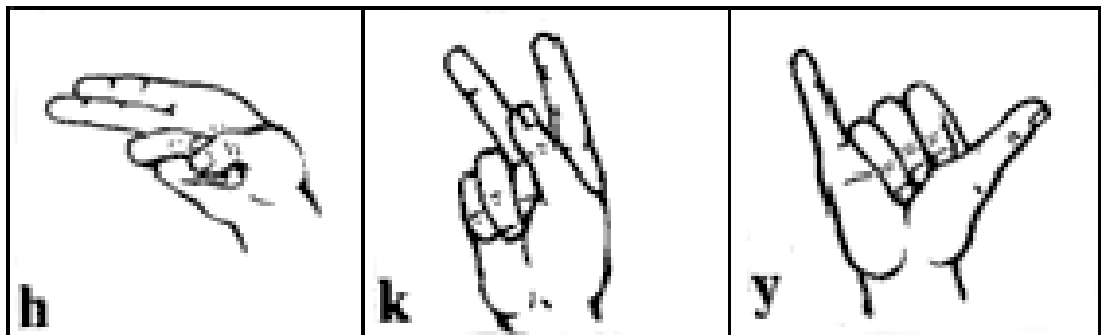


Рис. 6.2: Букви: H, K, Y

Вибір такої множини символів був обумовлений природністю і поширеністю жестів серед певної групи людей і можливістю розширення контрольного набору жестів тільки за рахунок збільшення тренувальних даних.

Для розпізнавання жестів було знято шаблони розпізнавання:

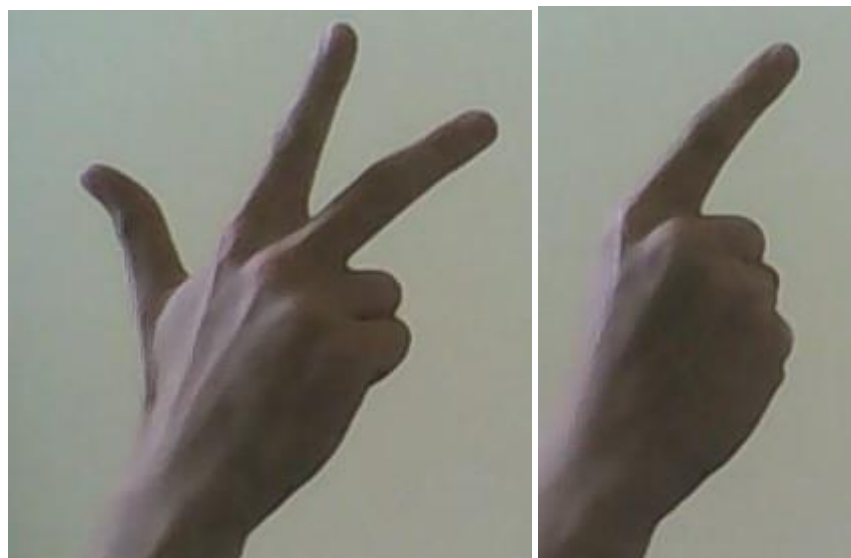


Рис. 6.3: приклад шаблонів для розпізнавання



Рис. 6.4: контрольні точки шаблонних зображень

Задля порівняння шаблонного та фактичного зображень було використано метод «співвідношення випробувань». Даний метод залучається у знаходженні двох найкращих збігів контрольних точок, та їх фільтрацію, якщо співвідношення «якості» збігу менше 0.75:

$$M_1 \geq 0.75 M_2$$

де M_1 – найкращий збіг, M_2 - другий найкращий збіг.

Дане випробування відкидає ті точки, які мають більш ніж 1 вірогідний збіг.

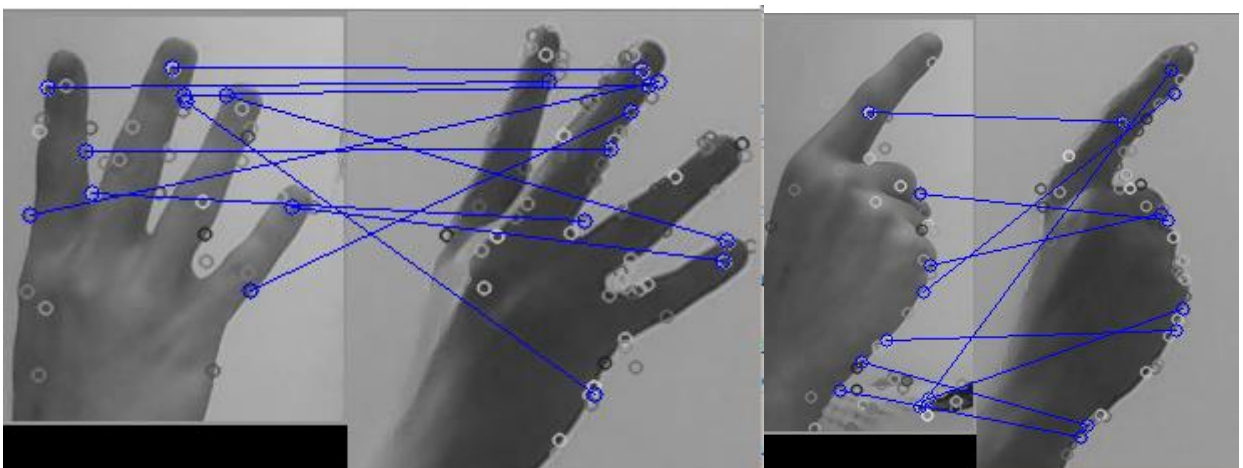


Рис. 6.5: порівняння зображень за допомогою контрольних точок

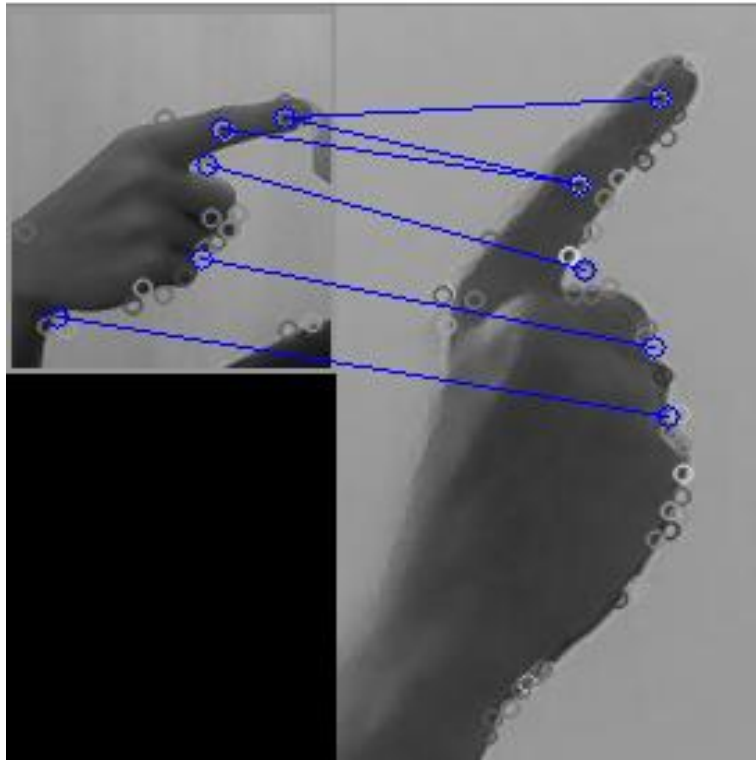


Рис. 6.6: стійкість алгоритму до афінних перетворень

6.2 Оптимізація методу порівняння

Розрахувавши попарну схожість шаблонних зображень між собою було отримано наступну таблицю:

Попарне порівняння особливих точок					
	1	2	3	4	5
1	1,00	0,43	0,40	0,34	0,29
2	0,35	1,00	0,44	0,37	0,45
3	0,26	0,25	1,00	0,27	0,28
4	0,17	0,29	0,28	1,00	0,54
5	0,19	0,23	0,32	0,35	1,00

Аналізуючи таблицю можна зробити висновок про надмірну схожість сусідніх шаблонів. Задля зменшення ступеня схожості зображень було виконано ітераційне видалення *загальних точок* (тобто таких, що мають однакові характеристики для більшості зображень, тобто не несуть інформативності для задачі класифікації).

6.3 Контурний аналіз



Рис. 6.7: Виділений контури руки з зображення

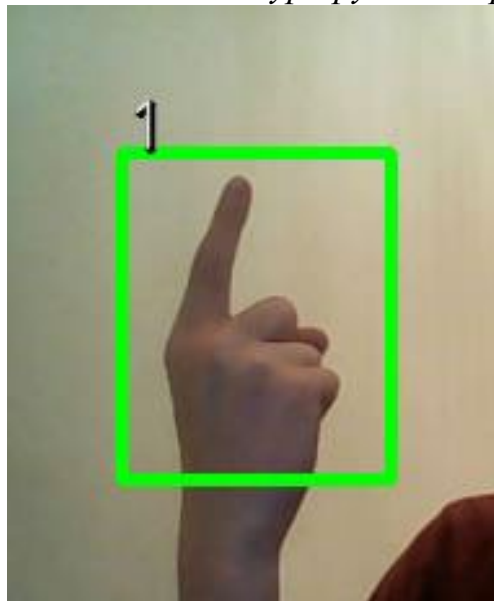


Рис.6.8: результуючий інтерфейс програми
6.4 Особливості реалізації.

Робота з великим обсягом даних.

Ми використали відео з роздільною здатністю 480x640 і розділили його на зображення зі швидкістю 24 кадри в секунду. В результаті відеоролик, який тривав лише одну хвилину, був представлений у вигляді понад тисячі картинок, кожна з яких являла собою сукупність $480 * 640 = 307200$ пікселів.

Багато алгоритмів комп'ютерного зору, наприклад, алгоритм віднімання фону, що використовується в цій роботі, потребують обробки кожного пікселя зображення. Як результат, за відсутності апаратної реалізації, час роботи алгоритму різко збільшується, що має великий вплив на процес розробки.

Зовнішні умови.

Більшість алгоритмів комп'ютерного зору дуже чутливі до зовнішніх умов. Невеликі зміни освітленості можуть суттєво спотворити їх ефективність, роблячи метод неефективним або зовсім невідповідним. Ось чому завдання встановлення обмежень для вхідного відео є настільки важливим, тому необхідно знайти баланс між простотою використання та ефективністю роботи програми.

Евристичні методи.

Як алгоритми штучного інтелекту, так і алгоритми комп'ютерного зору в навчанні покладаються на багато евристичних параметрів.

Наприклад, адаптивна швидкість алгоритму при виявленні руху на відео, або поріг класифікатора. Налаштування цих параметрів - тривалий процес, і вам

потрібно врахувати всі можливі вихідні дані, щоб у середньому отримати найкращі результати.

Збір даних і навчання системи.

Процес навчання системи можна розділити на кілька послідовних етапів: збір даних, їх розмітка, навчання параметрів класифікатора.

- 1) На першому етапі необхідно отримати велику кількість даних при різних варіаціях зовнішніх умов. Це повинно бути відео з різними людьми, з різним фоном і при різному освітленні.
- 2) Другий етап простий у виконанні, але досить трудомісткий, тому що для кожного зображення, що класифікується, необхідно вказати бажаний результат класифікації.
- 3) Останній етап являє собою саме навчання класифікатора, тобто мінімізації функції похибки для розмічених даних.

6.5 Тестування

Розроблені програми були протестовані, і ці відеозаписи були записані при різному освітленні та різному фоні. Відео змінює швидкість змін жестів, щоб перевірити ефективність поділу жестів. У всіх відео є одна людина, яка навчала систему класифікатора. Всього є близько двохсот жестів.

Для характеристики розроблюваної програми можна використати такі параметри:

- X1 – час на обробку кадра зображення;
- X2 – час на розпізнавання жеста;
- X3 – точність розпізнавання жеста.

Основні параметри програми

Параметр	Позначення параметра	Гранично допустиме значення	Значення параметра	
			Середнє отримане	Якого треба досягти
Відносна затримка на обробку одного кадру зображення, мс	X1	40	20	0
Час на розпізнавання жеста, мс	X2	1000	500	200
Точність розпізнавання жеста, %	X3	60%	40%	80%
Об'єм використаної пам'яті, МБ	X4	100	60	20

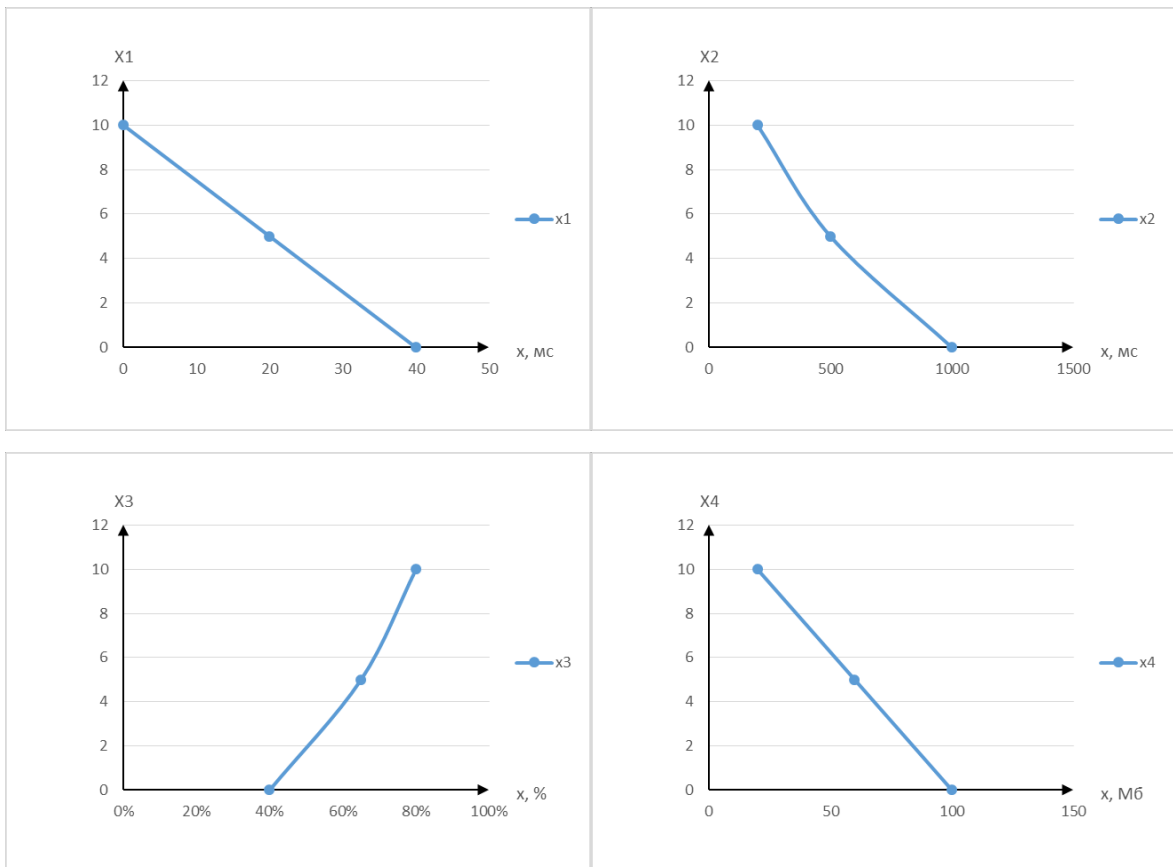


Рис. 7: оцінка основних параметрів програми. а) відносна затримка на обробку одного кадру зображення; б) час на розпізнавання жеста; в) Точність розпізнавання жеста; г) об'єм використаної пам'яті.

Вагомість параметрів визначають методом попарного їх порівняння, використовуючи результати ранжирування експертами.

РОЗДІЛ 7. РЕЗУЛЬТАТ РОБОТИ.

В ході роботи був реалізований розроблений алгоритм на мові Octave.

Навчання класифікатора проводилося на множині з 75 кадрів для кожного жесту. Налаштування класифікатора проводилася під конкретну людину, для отримання кращих результатів класифікації.

Основна вимога - зручність використання вийшло дотримати.

Інформація надходить від однієї камери, яку можна вбудувати в комп'ютер, щоб додаток можна було використовувати „з нуля” без необхідності тривалих годин налаштування та калібрування пристрою. Немає ставки на колір шкіри, що дозволяє використовувати програму в кімнаті зі специфічним освітленням, тим самим спотворюючи колір предмета.

Для оцінки класифікації використовувалися стандартні метрики:

$$precision = \frac{tp}{tp+fp}, \quad recall = \frac{tp}{tp+fn}.$$

tp (true positive) - кількість об'єктів на які було дано позитивну відповідь в тому випадку, коли він і був потрібний;

fp (false positive) - позитивну відповідь, коли був вірний негативний;

fn (false negative) - негативну відповідь, коли був потрібний позитивний.

У роботі під "false positive" розуміються ті випадки, коли класифікатору на вхід надходив жест, який не належить до контрольної групи, але який, тим не менш, був класифікований. Під "false negative", в свою чергу, розуміється кількість жестів помилково класифікованих як жести, які не відносяться до контрольної групи.

	Моменти контура	Гістограма спрямованих градієнтів
Precision	0.90643	0.96685
Recall	0.91716	0.97765
Percent	0.82703	0.93514

Таб. 7.1: Значення метрик в залежності від обраних ознак класифікатора.

РОЗДІЛ 8. ВИСНОВКИ.

В ході дослідження було визначено, що методи динамічного виділення підходять для підрахунку інтенсивності потоку. Методи семантичного

аналізу відео підходять для класифікації об'єктів на відео, однак вимагають додаткових витрат на побудову онтології. Методи з навчальними зображеннями підходять для вирішення більш складних завдань, в яких необхідно не тільки зробити підрахунок інтенсивності, але і відстежити і класифікувати об'єкт.

Запропонований алгоритм розпізнавання образів у відеопотоці дозволяє зменшити ймовірність помилкових виявлень об'єктів і підвищити швидкодію роботи системи, допомагає позбутися обмежень у вигляді шумів на зображенні і невиділення потрібного об'єкта-цілі на фоні за допомогою попередньої обробки зображення, оскільки цей метод дозволяє набагато краще працювати з напівтоновими чіткими зображеннями та позбутися проблем, пов'язаних з кутом нахилу руки шляхом тренування нових каскадів.

Розпізнавання образів (об'єктів, сигналів, ситуацій, явищ або процесів) - завдання ідентифікації об'єкта або визначення будь-яких його властивостей за його зображенням (оптичне розпізнавання) або аудіозапису (акустичне розпізнавання) та іншим характеристикам.

Подальший розвиток штучного інтелекту, зокрема проектування комп'ютерів п'ятого покоління, здатних до більш безпосередньому спілкуванню з людиною на природних для людей мовах і за допомогою мови, немислимі без розпізнавання. Тут рукою подати і до робототехніки, штучних систем управління, що містять як життєво важливих підсистем системи розпізнавання.

Саме тому до розвитку розпізнавання образів з самого початку була прикута чимало уваги з боку фахівців самого різного профілю - кібернетиків, нейрофізіологів, психологів, математиків, економістів і т.д. Багато в чому саме з цієї причини сучасне розпізнавання образів саме харчується ідеями цих дисциплін.

В ході роботи були вивчені існуючі підходи до розпізнавання образів, на основі яких було розроблено і реалізовано алгоритм по розпізнаванню латинської абетки жестів на потоці. Ряд експериментів на залежних від людини даних показав ефективність отриманого методу.

Література.

1. В.Г.Абакумов, Е.Ю.Ломакина sАвтоматическое распознавание жестов в интеллектуальных системах // «Электротехнические и компьютерные системы», 2010.
2. В.Н. Вапник, А.Я. Червоненкис «Теория распознавания образов», издательство: «Наука», 1974.
3. «Введение В Контурный Анализ», 2003 / под ред. Я.А. Фурмана.
4. А.Л.Воскресенский, С.Н.Ильин, М .Zelezny «О распознавании жестов языка глухих», 2010.
5. Л.Шапиро, Дж.Стокман «Компьютерное зрение», издательство: «Бином. Лаборатория знаний», 2006.
6. «Язык жестов» : URL:<http://jestov.net/>
7. Aran, Oya «Vision based sign language recognition: modeling and recognizing isolated signs with manual and non-manual components», 2008.
8. A. Barr, Edvard A. Feigenbaum «The Handbook of Artificial Intelligence», издательство: «Slategray Ventures», 1990.

9. Jorg Zieren, Karl-friedrich Kraiss «Robust person-independent visual sign language recognition», 2005.
10. Karl Pearson «On Lines and Planes of Closest Fit to System of Points in Space», 1901.
11. Chris Stauffer, W. Eric L. Grimson «Learning Patterns of Activity Using Real-Time Tracking», 2000.
12. Navneet Dalal, Bill Triggs «Histograms of Oriented Gradients for Human Detection», 2005.
13. Simon Lang «Sign Language Recognition with Kinect», 2011.
14. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. Метод потенциальных функций в теории обучения машин. - М.: Наука, 2004. - 384 с.
15. Горбань А., Россиев Д. Нейронные сети на персональном компьютере. //Новосибирск, Наука, 1996. – С 114 – 119.
16. Журавлев Ю.И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. М.: Наука, 2005. - Вып. 33. С. 5-68
17. Мазуров В.Д. Комитеты систем неравенств и задача распознавания // Кибернетика, 2004, № 2. С. 140-146.
18. Потапов А.С. Распознавание образов и машинное восприятие. - С-Пб.: Политехника, 2007. - 548 с
19. Минский М., Пейперт С. Перцептроны. - М.: Мир, 2007. - 261 с.

20. Растригин Л.А., Эренштейн Р.Х. Метод коллективного распознавания. 79 с. ил. 20 см., М. Энергоиздат, 2006. – 80 с.
21. Рудаков К.В. Об алгебраической теории универсальных и локальных ограничений для задач классификации // Распознавание, классификация, прогноз. Математические методы и их применение. Вып. 1. - М.: Наука, 2007. - С. 176-200.
22. Фу К. Структурные методы в распознавании образов. - М.: Мир, 2005. - 144 с.