

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ інфокомунікацій _____
(повна назва)

Кафедра _____ інформаційно-мережної інженерії _____
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____

Дослідження якості методів сегментації зображення людини в умовах дії
адитивних завад _____
(тема)

Виконав:
студент 2 курсу, групи ІМІм-21-1 _____
Глубока Ю.О. _____
(прізвище, ініціали)

Спеціальність 172 «Телекомунікації та
радіотехніка» _____
(код і повна назва спеціальності)
Освітня програма Інформаційно-мережна
інженерія _____
(повна назва освітньої програми)

Керівник _____ доц. Омельченко С.В. _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ Безрук В.М. _____
(підпис) (прізвище, ініціали)

2022 р.

Не містить відомостей, заборонених до відкритого публікування

Студент

(підпис)

Глубока Ю.О.

(прізвище та ініціали)

Керівник

(підпис)

Омельченко С.В.

(прізвище та ініціали)

Харківський національний університет радіоелектроніки

Факультет _____ інфокомунікацій _____
 Кафедра _____ інформаційно-мережної інженерії _____
 Рівень вищої освіти _____ другий (магістерський) _____
 Спеціальність _____ 172 «Телекомунікації та радіотехніка» _____
 (код і повна назва)
 Тип програми _____ освітньо-професійна _____
 Освітня програма _____ Інформаційно-мережна інженерія _____
 (повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 2022__ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____ Глибокій Юлії Олександрівні _____
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Дослідження якості методів сегментації зображення людини в умовах дії
адитивних завад _____

затверджена наказом по університету від « 21 » жовтня 2022 року № 1376 Ст

2. Термін подання студентом роботи до екзаменаційної комісії _____ 2022 р. _____

3. Вихідні дані до роботи _____

Проаналізувати існуючі методи сегментації зображень. Розібратись у різниці методів, Зрозуміти які методи сегментації краще підходять до яких умов використання, проаналізувати можливості роботи алгоритмів в умовах адитивних завад.

4. Перелік питань, що потрібно опрацювати в роботі _____

Вступ

1 Мета сегментації зображень

2 Класифікація та методи сегментації зображень

3 Природу адитивних завад у зображеннях

4 Сучасні методи для сегментації зображення людини

Висновки

5. Перелік графічного матеріалу із зазначенням креслень, схем, плакатів, комп'ютерних ілюстрацій (слайдів) _____

Слайди у форматі Power Point : 19 слайдів

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	<i>Ознайомлення із завданням. Уточнення ТЗ.</i>	<i>24.10.22</i>	
2	<i>Підбір літератури за темою роботи</i>	<i>26.10.22 -11.11.22</i>	
3	<i>Аналіз літератури з досліджуваної проблеми</i>	<i>11.11.22 – 13.11.22</i>	
4	<i>Виконання розділу 1</i>	<i>13.11.22 – 25.11.22</i>	
5	<i>Виконання розділу 2</i>	<i>25.11.22 – 30.11.22</i>	
6	<i>Виконання розділу 3</i>	<i>30.11.22 – 05.12.22</i>	
7	<i>Виконання розділу 4</i>	<i>05.12.22- 11.12.22</i>	
8	<i>Оформлення пояснювальної записки</i>	<i>11.12.22 – 14.12.22</i>	
9	<i>Оформлення презентаційного матеріалу,</i>	<i>14.12.22 – 18.12.22</i>	
	<i>підготовка до захисту у ЕК</i>		

Дата видачі завдання 24 жовтня 2022 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Омельченко С.В.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи: 43 с., 6 рис., 1 додаток, 13 джерел.

Об'єкт роботи – методи сегментації зображень.

Мета роботи – аналіз можливостей існуючих методів сегментації в умовах наявності адитивних завад.

Було проаналізовано популярні існуючі методи сегментації зображень. Було виявлено який саме вплив на зображення мають адитивні завади. На основі принципу роботи методів було зроблено висновки про їх ефективність в умовах наявності адитивних завад. Були проаналізовані деякі сучасні методики, створені спеціально з метою сегментації та подальшого використання сегментованих зображень, на яких зображені люди.

МЕТОДИ СЕГМЕНТАЦІЇ, СЕГМЕНТАЦІЯ ЗОБРАЖЕНЬ ЛЮДЕЙ,
АДИТИВНІ ЗАВАДИ

ABSTRACT

Explanatory slip 43 p., 6 fig., 1 app., 13 sources.

The object of work is image segmentation methods.

The purpose of the work is to analyze the possibilities of existing segmentation methods in the presence of additive interference.

Popular existing methods of image segmentation were analyzed. It was found what kind of effect additive interference has on the image. Based on the principle of operation of the methods, conclusions were made about their effectiveness in the presence of additive interference. Some modern techniques, created specifically for the purpose of segmentation and further use of segmented images depicting people, were analyzed.

**SEGMENTATION METHODS, SEGMENTATION OF PEOPLE'S IMAGES,
ADDITIVE INTERFERENCES**

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ	8
ВСТУП	9
1 ЗАДАЧІ СЕГМЕНТАЦІЇ	10
1.1 Мета сегментації	10
1.2 Властивості зорової системи людини	11
2 АДИТИВНІ ЗАВАДИ	13
3 АЛГОРИТМИ СЕГМЕНТАЦІЇ	17
3.1 Класифікація	17
3.2 Морфологічні методи	19
3.3 Метод виділення зв'язкових компонентів	20
3.4 Порогові методи	20
3.5 Методи нарощування областей	22
3.6 Текстурні методи	25
3.7 Інші	27
4 СПЕЦІАЛЬНІ МЕТОДИКИ ДЛЯ РОЗПІЗНАВАННЯ ОБЛИЧЬ	30
4.1 Сегментація обличчя	29
4.2 Бази даних для сегментації	32
4.3 Геометричні методи	34
4.4 Цілісні методи	35
4.5 CRF	37
4.6 Гібридні методи	38
4.7 Методи глибокого навчання	38
ВИСНОВКИ	42
ПЕРЕЛІК ПОСИЛАНЬ	43
ДОДАТОК А	Ошибка! Закладка не определена.

ПЕРЕЛІК СКОРОЧЕНЬ

NOT – логічна операція «НЕ»;

AND – логічна операція «І»;

OR – логічна операція «АБО»;

XOR – логічна операція що виключає «АБО»;

2D (2-Dimensional) – двовимірний;

3D (3-Dimensional) – тривимірний;

GMM – модель суміші Гауса;

FCN (fully-convolutional networks) — повнозвертальні мережі;

SOA (state-of-the-art) – найновітніший стан методів, способів сприйняття, зображення чи створення, тощо;

ASM – активне моделювання форми;

CNN – згорткові нейронні мережі;

CRF (Conditional random fields) – умовні випадкові поля;

RBM (restricted Boltzmann machine) – обмежена машина Больцмана

ВСТУП

Сегментацією зображення називається розбиття зображення на несхожі за деякою ознакою області. Передбачається, що області відповідають реальним об'єктам або їх частинам, а межі областей відповідають межах об'єктів. Сегментація відіграє важливу роль у завданнях обробки зображень та комп'ютерного зору.

Обличчя людини є важливим джерелом інформації при спілкуванні для людей. Вираз обличчя, міміка, артикуляція під час розмови, рухи головою є зручним, природним і, що важливо, необтяжливим способом передачі інформації. Нездатність комп'ютера сприйняти і розпізнати такі природні для людини способи спілкування ускладнює передачу та сприйняття інформації.

У повсякденному житті купа людей користуються продуктом сегментації зображення людей, навіть не розуміючи цього. Здебільшого це розважальний контент, проте технологія може використовуватись і у сфері безпеки, кожен хто коли небудь користувався face id мав справу з нейромережею, яка за допомогою різних методів сегментації відрізняє одне обличчя від інших. Сегментація особи є важливим проміжним етапом для багатьох програм аналізу особи. Сегментація особи надає багату та достатню інформацію для кількох завдань середнього рівня зору. Аналіз осіб особливо складний, коли зображення збираються у диких умовах

Відповідно до цього всього завади, що запобігають правильній роботі сегментації можуть як спричинити незначні незручності так і значно погіршити роботу важливих для суспільства систем.

Актуальність дослідження полягає у зростаючій потребі в більш ефективних методах сегментації для подальшого використання отриманих даних у комп'ютерному аналізі змісту зображень, обробці або зміні частин або цілих зображень.

1 ЗАДАЧІ СЕГМЕНТАЦІЇ

1.1 Мета сегментації

У комп'ютерному зорі сегментацією називають процес поділу цифрового зображення на кілька сегментів, тобто інформативні фрагменти.

Метою сегментації є спрощення та/або зміна представлення зображення таким чином, щоб його було легше аналізувати. Сегментація зображення зазвичай використовується для виділення об'єктів і меж (ліній, кривих тощо) на зображеннях. У більш ранніх техніках використовується розщеплення та злиття регіонів, що відповідає роздільним та агломераційним алгоритмам у літературі з кластеризації. Сучасні алгоритми найчастіше оптимізують деякі глобальні критерії, такі як внутрішньо-регіональна узгодженість та міжрегіональні довжини кордонів [1].

Є велике різноманіття застосувань сегментації зображень. Деякі види сегментації можуть використовуватись у безпілотних автомобілях та роботах. Камери служать для отримання точної візуальної інформації про навколишнє середовище, але критично важлива умова правильної роботи – наскільки добре алгоритм розуміє, що зараз оточує робота.

Інша сфера застосування – медицина, де за допомогою сегментації можна проаналізувати рентгенографічні та електронно-мікроскопічні знімки – наприклад, у стоматології, пульмонології, онкології, генетиці тощо. За допомогою сегментації аналізують супутникові зображення та карти, скупчення людей та об'єкти глибокого космосу.

Завдання сегментації та розпізнавання зображень технічно комплексне. Воно ділиться на кілька великих сегментів. Основне завдання розпізнавання: зрозуміти, чи дані на зображенні до класу шуканих об'єктів. Основні технічні складності, що виникають у цьому випадку:

- Зображення пред'являються складному тлі;

- Шукані області мають складну геометрію;
- Вхідні дані мають шум або дезорієнтуючу інформацію.

Різні частини зображення мають різні характеристики (підсвічування, освітленість, завади)

Для вирішення завдання на різних етапах застосовують різні підходи та методи сегментації, нормалізації та розпізнавання.

1.2 Властивості зорової системи людини

Велика кількість інформації представлена у вигляді статичних або динамічних зображень. Оскільки ці зображення надалі досліджуються та аналізуються людиною, важливо знати механізми зорового сприйняття. Ці знання є потужним інструментом у розробці різноманітних систем обробки зображень.

Зазначимо, що психофізичні аспекти сприйняття світла до кінця не вивчені. Однак існує ряд різних методів і підходів, які пояснюють закономірності та властивості зорового сприйняття світла людиною.

Людське око є унікальним механізмом, що забезпечує адаптивну адаптацію до зовнішніх умов. Розглянемо деякі фундаментальні властивості зорової системи людини. Важливою властивістю зорової системи є чутливість, тобто здатність реагувати на зовнішні зміни. Чутливість характеризується верхнім і нижнім абсолютними порогами.

Існують різні види чутливості. Світлочутливість характеризує здатність ока реагувати на найменші потоки світла. Однак тут слід зазначити, що ймовірність розпізнати найменший світловий потік також залежить від інших факторів, наприклад кута огляду. Зоровий апарат по-різному реагує на випромінювання однакової потужності, яке, однак, випромінюється з різних спектральних діапазонів. Таку чутливість називають спектральною.

Здатність ока розпізнавати мінімальні відмінності в яскравості сусідніх ділянок зображення характеризується контрастною чутливістю. Також для

зорової системи характерна диференційована чутливість до відтінків, тобто до випромінювання з різних частин спектра. Зорова система також характеризується чутливістю до насиченості кольорів [2].

Перераховані вище види чутливості зорової системи непостійні, а залежать від багатьох факторів, насамперед від умов освітлення. Наприклад, при переході з темної кімнати в світлу потрібен час, щоб очі відновили чутливість до світла. Цей процес називається адаптацією ока до яскравості.

Сприйняття кольору характеризується трьома основними характеристиками – світлістю, відтінком і насиченістю. Кольорові простори використовуються для класифікації кольорів.

На основі властивостей і характеристик зорових систем створюються різноманітні моделі кольорового зору. Серед них виділяється модель кольорового зору Фрея. Особливістю цієї моделі є те, що зорова система представлена трьома каналами, два з яких характеризують колір, а третій – яскравість. Ця модель найбільш вдало узгоджується з багатьма властивостями кольорового зору.

2 АДТИВНІ ЗАВАДИ

Завади мають різноманітний і складний характер, що створює труднощі за умови їх класифікації. Класифікацію за ними проводять за різними ознаками, зокрема за законом розподілу, видом кореляції, характером стаціонарності, механізмом виникнення.

За законом розподілу завади поділяють на:

- гауссівські;
- негауссівські.

Такий розподіл зручний, оскільки частина найбільш характерних завад має нормальний гауссівський розподіл, а багато інших завад можуть бути зведені до гауссівських. За характером стаціонарності розрізняють завади стаціонарні та нестаціонарні. Інші завади належать до класифікації випадкових процесів. При виявленні сигналу завадах часто припускають, що завади мають характер стаціонарного випадкового процесу з нормальним законом розподілу, такі завади називають флуктуаційними. Однак деякі види завад, наприклад імпульсні, мають характер нестаціонарного процесу.

За механізмом виникнення завади поділяються на:

- природні;
- індустріальні;
- системні;
- штучні.

Завади можуть бути активними та пасивними. Активні завади всіх цих видів можуть створювати завади для різних систем. Пасивні завади зазвичай виникають за рахунок відображення сигналів від будь-яких поверхонь.

Природні активні завади поділяються на:

- завади від біооб'єкта;
- внутрішньоприймальні;
- космічні;

- атмосферні;
- ті, що можуть характеризувати тремор м'язів. та інші подібні явища нестационарності;
- ті шуми, що виникають у приймальному витраті;
- ті, що створюються довільно, змінюються у часі випромінюваннями метagalaktiki. На це загальне тло накладаються радіовипромінювання потужних дискретних джерел - планет, зірок;
- ті, що можливі лише при натуральному експерименті та обумовлені грозливими розрядами між різноіменно зарядженими масами повітря, парами води та землі.

Індустріальними завадами називають завади, що створюються електричними та електронними пристроями. Вони мають імпульсивний характер.

Імпульсивні завади є окремими імпульсами або їх послідовністю, що мають довільну форму, амплітуду і значення інтервалів між імпульсами, причому тривалість інтервалів значно перевищує тривалість самих імпульсів.

Системні, створювані технічними засобами, званими також зосередженими завадами, оскільки їхня енергія полягає у вузькому частотному діапазоні.

Штучні завади можуть бути як активні, так і пасивні. Активні штучні завади можуть бути шумовими або синусоїдальними із застосуванням різних видів модуляції. Штучні пасивні завади створюють завадове тло, що утруднює виділення корисного сигналу, і навіть імітує хибні мети. Дані види завад використовують для розвідки чи шпигунської діяльності. Канал оптичного запису часто має несиметричну амплітудну характеристику та вносить лінійні та нелінійні спотворення. Помилки фокусування та стеження за доріжкою можуть призводити до мультиплікативних завад. Неоднорідні властивості носія призводять до адитивної завади, до якої додаються шуми фотоприймача. Дефекти носія викликають випадання сигналу. Слід зазначити, що для оптичного запису, як правило, використовують носій з досить товстим захисним

прозорим покриттям робочого шару. Апертурний кут при фокусуванні променя досить великий, і діаметр променя, що фокусується, на поверхні захисного шару приблизно дорівнює 1 мм. Тому щодо випадень більшою мірою небезпечні дефекти робочого шару носія, набуті у його виробництва. При експлуатації носія пил та дрібні подряпини на захисному шарі несуттєво впливають на випадіння сигналу.

Серед адитивних завад різного походження особливе місце займає флуктуаціона завада (флуктуаційний шум), що являє собою випадковий процес із нормальним розподілом (гаусівський процес). Така завада найбільш вивчена й становить інтерес як у теоретичному, так і в практичному відношенні. Цей вид завад має місце у всіх реальних каналах.

Основні джерела шуму гауса в цифрових зображеннях виникають під час отримання. Датчик має власний шум через рівень освітленості та власну температуру, а електронні схеми, підключені до датчика, вносять свою частку шуму електронних схем.



Рисунок 2.1 – Еталонне зображення та зображення, спотворене адитивним гаусовим шумом

Типова модель шуму зображення - гауссова, адитивна, незалежна в кожному пікселі і незалежно від інтенсивності сигналу, викликаного переважно шумом Джонсона – Найквіста (тепловим шумом), у тому числі шумом скидання конденсаторів. Шум підсилювача - це основна частина шуму зчитування датчика зображення, тобто постійного рівня шуму в темних областях

зображення. У кольорових камерах, де в каналі синього кольору використовується більше посилення, ніж у зеленому або червоному каналі, в синьому каналі може бути більше шуму.

3 АЛГОРИТМИ СЕГМЕНТАЦІЇ

3.1 Класифікація

Повної та всеосяжної класифікації алгоритмів сегментації зображень не існує, кожна з них так чи інакше упускає деякі методи, у тому числі авторські задуми, не надто популярні у широкому застосуванні. Тому безліч схем класифікації алгоритмів сегментації зображень можна умовно поділити на наступні категорії:

- морфологічні методи;
- порогові методи;
- методи нарощування кордонів;
- текстурні методи;
- інші.

Метод виділення кордонів слід використовувати, якщо межі досить чіткі і стабільні. Порогові методи застосовуються при існуванні стабільних відмінностей в яскравості окремих областей. Методи нарощування кордонів ефективні при наявності стійкої зв'язності всередині окремих сегментів. Для опису та сегментації властивостей зображень (однорідності, регулярності) використовують текстурні методи, які умовно поділяються на дві категорії: статистичні та структурні.

Морфологічні методи застосовуються переважно для роботи з двійковими (чорно-білими) зображеннями. Ці методи дозволяють витягувати компоненти зображення, які можуть використовуватися для уявлення форми об'єкта. Дана група методів під час роботи використовує такі найпростіші логічні операції, як:

- NOT (логічне «НЕ»);
- AND (логічне «І»);
- OR (логічне «АБО»);

– XOR (що виключає «АБО»).

Порогові методи мають інтуїтивно зрозумілі властивості та є простими у реалізації. Існують кілька основних видів порогової сегментації, але базовими є лише два: метод з оптимальним порогом та метод з адаптивним порогом. Решта методів цього класу є похідними від двох згаданих алгоритмів.

Методи нарощування областей являють собою алгоритми, що рекурсивно виконують процедуру угруповання пікселів у підобласті за задалегідь заданими критеріями.

Текстурні методи спираються під час аналізу на дифузні властивості поверхні аналізованого об'єкта такі як колір, відбивна здатність, тощо. Представлені у цій категорії методи є наборами складних операторів, які здатні звести процес розпізнавання поверхонь до простого завдання розрізнення рівнів яскравості.

Також сама собою сегментація, а не її методи поділяється на три види:

Семантична сегментація – визначає, чи належать набори пікселів у зображенні до певних класів об'єктів (наприклад, котів, собак, людей, квітів, автомобілів тощо).

Сегментація екземплярів – на відміну від семантичної сегментації, це завдання призначає окремі сегменти кожному об'єкту в межах одного класу. Наприклад, якщо зображення містить п'ять котів, двох собак і десять рослин, семантична сегментація просто вибирає всі області з котами, собаками чи рослинами, не відокремлюючи окремі об'єкти в кожному класі (розпізнає, що зображення містить котів, собак, і рослини), тоді як сегментація екземплярів виділяє кожного kota, собаку та рослину як окремий об'єкт.

Паноптична сегментація - поєднує завдання семантичної та екземплярної сегментації. Для завдання панорамної сегментації кожному пікселю зображення має бути присвоєна рівно одна мітка.

3.2 Морфологічні методи

До морфологічних методів сегментації відноситься метод виділення кордонів, і ось його основний принцип:

Кордон безлічі A , яку будемо позначати $\beta(A)$, може бути отримана шляхом виконання операції ерозії A по B , а потім отриманням множини різниці між множиною A та результатом його ерозії.

$$\beta(A) = A \setminus (A \ominus B)$$

де B - Будь-який примітив операції.

При такому підході межі областей значно відхиляються одна від одної та від фону, що дозволяє їх розпізнавати за локальними неоднорідностями інтенсивності (рівня сірого).

Іншими словами, це процес виділення кордонів. Відзначимо, наскільки важливий цей етап для розуміння властивостей зображення, оскільки добре відомо, що ці межі складаються зі значущих властивостей і містять важливу інформацію.



Рисунок 3.1 – Приклад сегментації за допомогою методу виділення кордонів.

3.3 Метод виділення зв'язкових компонентів

Відношення зв'язків між елементами зображення є фундаментальним поняттям. Для того щоб встановити, чи є два елементи зображення зв'язковими, необхідно, щоб вони були сусідами і рівні їх яскравості задовольняли заданому критерію подібності.

Нехай S - деяке підмножина елементів зображення. Два елементи p і q називаються зв'язковими в S , якщо між ними існує шлях, що повністю складається з елементів множини S . Для будь-якого пікселя p з S безліч всіх пікселів, пов'язаних з ним у S , називається зв'язковим компонентом (або компонентної зв'язаності) S . Якщо безліч S містить лише одну компоненту зв'язності, воно називається зв'язковим безліччю. Насправді виділення компонент зв'язності на зображенні займає центральне місце у багатьох прикладних завданнях аналізу зображень.

3.4 Порогові методи

Порогові методи займають центральне місце у завданнях сегментації зображень завдяки своїй простоті та високій обчислювальній ефективності у реалізації. Порогові критерії як основні, або допоміжних використовуються практично в всіх методах сегментації за яскравістю, кольоровістю та контрастності.

Пороговий метод із глобальним порогом. Цей метод є найпростішим. Після вибору глобального порогу відбувається поелементна перевірка всього зображення. Процедура має на увазі поділ зображення на дві області: перша відноситься до об'єкта, друга – до фону. У разі успішності цілком залежить від цього, наскільки добре діаграма піддається поділу. Успішного застосування цього методу очікується за умов контрольованого висвітлення.

Пороговий метод із адаптивним порогом. Раніше було зазначено, що метод із глобальним порогом гарний доти, доки ми маємо контрольоване освітлення. Як тільки освітлення стає нерівномірним, добре розділяється гістограма може перетворитися на гістограму, що погано розділяється, і метод не спрацює. У цьому випадку вихідне зображення слід розділити на підобласті, у кожній із яких для сегментації шукається і використовується свій поріг. Основною проблемою тут є завдання розбиття зображення на підобласті та вибір для кожної з них свого порога.

Оскільки поріг залежить від параметрів підобласті зображення, такий поріг називає адаптивним.



Рисунок 3.2 – Приклад сегментації за допомогою порогового методу із адаптивним порогом

Метод сегментації гістограми. Суть методу гістограм полягає в тому, щоб створити гістограму бімодального зображення, вибрати порогове значення в мінімальній точці та сегментувати зображення за допомогою вибраного порогового значення. В залежності від умов захоплення зображення поріг можна вибрати одноразово або адаптивно для кожного малюнка окремо. До того ж гістограма може мати кілька піків для відповідної кількості об'єктів. За таких умов для сегментації об'єктів певного класу вибираються два пороги та використовується метод перетину. Для комбінування сегментованих пікселей у

зв'язні множинах як основні, в цьому випадку використовується хвильовий метод.

У той же час при нерівномірному освітленні або наявності великої кількості об'єктів гістограма яскравості не має бімодальної структури. Тому для використання методу гістограми необхідно провести локалізацію об'єкта. Для цих цілей використовують скануючі маски, або розбиття зображень..

Найважливіша перевага порогових методів є їх обчислювальна ефективність. Тому бажано використовувати пороги для швидкої, хоча грубої, обробки зображень на початковій фазі сегментації. Водночас найголовнішою перевагою гістограмних методів є їх пристосованість до розподілу яскравості об'єкта і фону, що виражаються в можливості відповідного вибору порогу для сегментації і в змінних умовах отримання вихідних даних.

Однак у більшості практично значущих випадків така адаптивність є локальною і можливий лише в певному середовищі об'єкта після його локалізації.

Тандем порогового та гістограмного методів дає нам розуміння того, що для досягнення стабільності доцільно спочатку підлаштуватися під спостережуваний розподіл яскравості зображення, а лише потім виконувати сегментацію з оптимальними параметрами.

3.5 Методи нарощування областей

Метод вододілів - один з основних алгоритмів нарощування областей, які рекурсивно виконують процедуру угруповання пікселів у підобласті за задалегідь заданими критеріями. Даний метод включає наступні три базові концепції:

- виявлення та усунення розривів;
- порогова обробка;
- обробка областей.

Завдяки даним концепціям метод вододілів дозволяє отримувати стабільніші результати сегментації (у тому числі безперервні межі областей). Цей підхід також дозволяє включати в процес сегментації додаткові обмеження.

Поняття вододілу засноване на поданні зображення як тривимірної поверхні, де як висота використовується рівень яскравості пікселя. У цьому випадку на поверхні можна виявити три типи точок:

- точки локального мінімуму;
- точки, що знаходяться на схилі, з яких вода зливається до центру водойми;
- точки, що знаходяться на гребені височини.

Лінії, утворені точками-гребнями, є лінії вододілів, тому основним завданням даного методу є саме пошук ліній вододілів.

Алгоритм:

- У місцях локального мінімуму утворюємо «дірки», якими вода почне заповнювати тривимірну поверхню.
- Якщо вода з двох боків гребеня готова поєднатися в один басейн, встановлюємо перегородку.
- Коли над водою залишаться тільки перегородки, зупиняємо алгоритм.

Отримані таким чином перегородки є необхідні лінії вододілів.

На Рисунку 3.3 наведено покрокове виконання алгоритму. На лівому верхньому малюнку показаний оригінальний «рельєф» деякого зображення, на нижньому правому – результат роботи алгоритму.

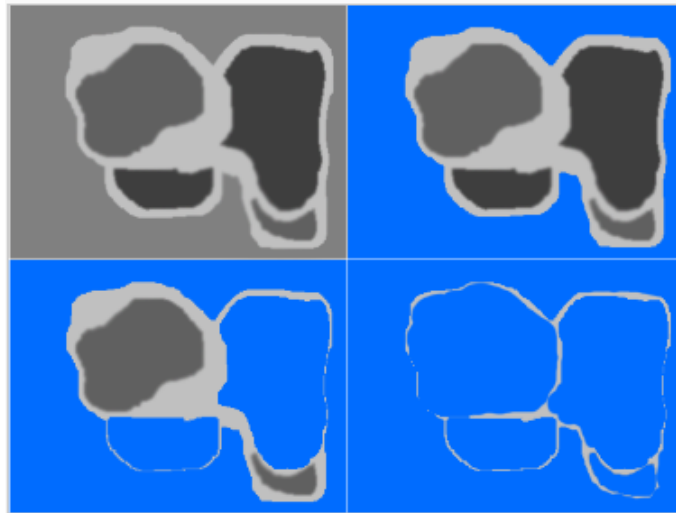


Рисунок 3.3 – Приклад роботи методу вододілів

Метод Вирощування областей. Як випливає з назви, вирощування областей є методом, який групує пікселі зображення в області за заздалегідь заданими критеріями. Суть вирощування областей полягає в тому, щоб задати піксель, що грає роль "центру кристалізації" області, а потім, починаючи з нього як зі стартового пікселя, проводити нарощування області з використанням хвильового методу шляхом приєднання сусідніх пікселів, які за своїми властивостями близькі до центру кристалізації, наприклад, за яскравістю або контрастністю. При цьому на кожній ітерації (попередньо вирощування) область, що вирощується разом з пікселями кандидатами, перевіряється на однорідність з використанням заданого критерію. Часто для цієї мети використовуються порогові та статистичні критерії, засновані на аналізі дисперсії, яка повинна не перевищувати задане порогове значення, а також критерії, що ґрунтуються на перевірці гіпотези про нормальність розподілу яскравості області. Нарощування областей - один з найбільш простих для розуміння підходів до сегментації за яскравістю: сусідні елементи з однаковими яскравостями групуються разом і утворюють область

Однією з ключових проблем при організації вирощування областей є вибір стартових пікселів. Так, якщо попередньо вирощування сегментовані пікселі деякого зображення, або його межі, то проблем із вибором стартового пікселя немає. В іншому випадку необхідно спиратися на модельні характеристики об'єкта з метою пошуку стартового пікселя. Крім того,

представляє проблему побудова критерію зупинки процесу вирощування, оскільки локальної інформації про суміжні пікселі околиці в нетривіальних завданнях не вистачає для цього. Тому найважливішим недоліком методів вирощування областей є їх низька стійкість умовах низької контрастності зображень. Описана щойно процедура вирощує області з множини початкових пікселів, що відіграють роль "центрів кристалізації". Альтернативний підхід полягає в тому, щоб провести первинне вирощування або розбиття зображення на безліч областей, що не перетинаються, і потім ітераційно здійснювати злиття (розщеплення) цих початкових областей.

Злиття-розщеплення областей. Альтернативою вирощування є розбиття зображення на безліч областей, які не перетинаються і потім виконання процедури злиття-розщеплення цих початкових областей. Головна проблема реалізації методів злиття-розщеплення є низька потужність критеріїв, що застосовуються для оцінювання однорідності областей. Разом з тим в порівнянні з методами контурної сегментації методи сегментації областей ґрунтуються вже на аналізі характеристик областей, а не окремих пікселів.

3.6 Текстурні методи

Метод текстурних дескрипторів. Наша зорова система може легко розпізнавати та розрізняти текстури. Виявляється, набагато складніше охарактеризувати та виділити більш «дифузні» властивості текстури за допомогою чітко визначених параметрів, які дозволяють комп'ютеру виконувати це завдання.

Як ми можемо описати текстуру? Будь-яка структура, яка поширюється на велику область зображення, точно не буде розпізнана як текстура. Отже, основною властивістю текстури є дрібна елементарна структура, яка періодично або квазіперіодично повторюється в просторі, як візерунок. Тому достатньо описати невелику елементарну структуру і правила повторення. Останні визначають характерний розмір текстури.

Аналіз текстури можна порівняти з аналізом структури твердих тіл, предметом, що вивчається у фізиці твердого тіла, хімії та мінералогії. Структурний аналіз ускладнюється тим фактом, що і структура, і періодичність можуть демонструвати значні випадкові зміни.

Текстури можна організовувати ієрархічно, тобто вони можуть виглядати абсолютно по-різному в різних масштабах. Розглянемо задачу, у якій необхідно виявити випадково орієнтовані об'єкти на зображенні. Нас цікавить не орієнтація об'єктів, а їх відмінність один від одного. Тому залежні від орієнтації параметри текстури нецікаві. Ми все ще можемо використовувати їх, але тільки якщо об'єкти мають форму, яка тоді дозволить нам визначити їхню орієнтацію. Ми можемо використовувати подібні аргументи для масштабно-інваріантних функцій. Якщо об'єкти, які нас цікавлять, знаходяться на різних відстанях від камери, параметр текстури, який використовується для їх виявлення, також повинен бути незмінним щодо масштабу. В іншому випадку виявлення об'єкта залежить від відстані. Однак, якщо текстура змінює властивості з масштабом, як у прикладі завіси, масштабно-інваріантні текстури можуть не існувати взагалі. Тоді використання текстур для опису об'єктів на різних відстанях стає складним завданням.

Нас цікавили самі предмети, а не їх орієнтація в просторі. Орієнтація поверхонь є ключовою характеристикою для іншого завдання обробки зображень, реконструкції 3D-сцени з 2D-зображення. Якщо ми знаємо, що поверхня об'єкта має однорідну текстуру, ми можемо проаналізувати орієнтацію та розміри текстури, щоб визначити орієнтацію поверхні в просторі. Для цього потрібні характерні розміри та орієнтації текстури.

Аналіз текстур є однією з областей обробки зображень, де досі бракує фундаментальних знань. Тому література містить багато різних емпіричних і напівемпіричних підходів до аналізу текстури. Навпаки, досить простий підхід до аналізу текстур, здається, створює складні оператори текстури з елементарних операторів.

Для аналізу текстур використовуються лише чотири основні оператори текстури:

- середнє значення;
- дисперсія;
- орієнтація;
- масштаб.

Застосовуються на різних рівнях ієрархічної структури послідовності обробки зображень. Після того, як ми обчислили локальну орієнтацію та локальну шкалу, можна знову застосувати оператор середнього значення та оператор дисперсії, але цього разу не до середнього значення та дисперсії рівнів яскравості, а до локальної орієнтації та масштабу.

Ці чотири основні оператори текстур можна розділити на два класи. Середнє значення та дисперсія не залежать від обертання та масштабу, тоді як оператори орієнтації та масштабу вказують лише орієнтацію та масштаб відповідно. Ця важлива відмінність між інваріантними або неінваріантними параметрами масштабу та обертання значно спрощує аналіз текстури. Цінність такого підходу полягає в ортогональності набору параметрів, простоті та можливості його застосування.

3.7 Інші

Метод Grabcut дозволяє проводити сегментацію подібно магнітному ласо, для отримання сегменту достатньо виділити об'єкт з частиною фону на зображенні з допомогою миші, далі алгоритм проведе його сегментацію і ми отримаємо окремий об'єкт без фону, така сегментація на відміну від інших є повільнішою, тому, що виконується зазвичай людиною і існує ризик виникнення помилок через не уважність виконуючого. Даний метод часто використовується для обробки зображень отриманих через Lidar, на зображенні необхідно виділити набори пікселів які являють собою транспортні засоби, або

значні завади для руху. Даний метод є універсальним для сегментації будь-якого типу зображень, та об'єктів на них.

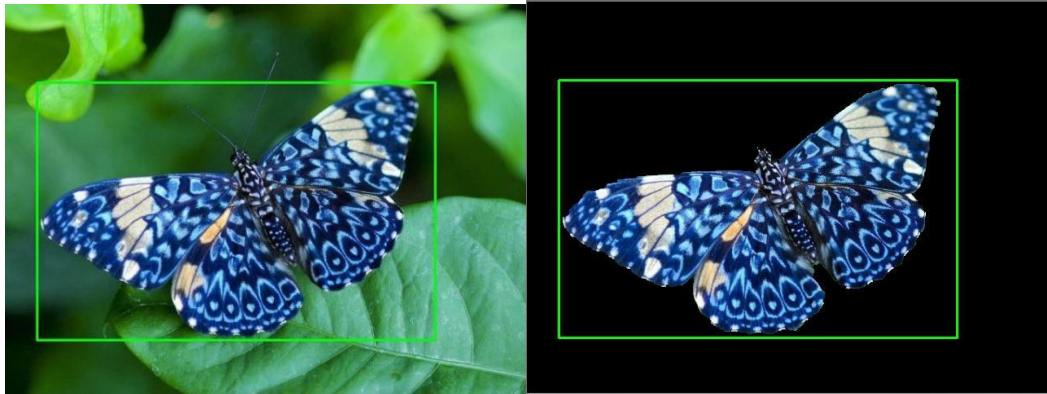


Рисунок 3.4 – Приклад сегментації за допомогою методу Grabcut

Розглянемо ідею алгоритму. На основі алгоритму інтерактивної сегментації GraphCut, який вимагає від користувача розміщувати маркери на фоні та на об'єкті. Зображення розглядається як масив. $Z(Z_1, \dots, Z_n, \dots, Z_N)$ де Z – значення інтенсивності пікселів, N – загальна кількість пікселів. Щоб відокремити об'єкт від фону, алгоритм визначає значення елементів масиву прозорості $a(a_1, \dots, a_n, \dots, a_N)$ і a_n може приймати два значення, якщо $a_n = 0$, то піксель належить фоні, якщо $a_n = 1$, то об'єкту. Внутрішній параметр θ містить гістограму розподілу інтенсивності переднього плану та гістограму фону: $\theta = \{h(z; a), a = 0, 1\}$

Завдання сегментації полягає в пошуку невідомих a_n . Розглядається енергетична функція:

$$E(a, \theta, z) = U(a, \theta, z) + V(a, z)$$

Крім того, мінімальна енергія відповідає найкращій сегментації.

$$U(a, \theta, z) = - \sum_n \log h(z_n, a_n)$$

$$V(a, z) = \sum_{(m,n) \in C} \frac{1}{dis(m,n)} [a_n \neq a_m] \exp(-\beta(z_m - z_n)^2)$$

$V(a, z)$ – термін відповідає за зв'язок між пікселями. Сума складається з усіх пар пікселів, які є сусідами, $dis(m,n)$ – евклідова відстань. $[a_n \neq a_m]$

відповідає за участь пар пікселів у сумі, якщо $a_n = a_m$ то ця пара не враховується.

$U(a, \theta, z)$ – відповідає за якість сегментації, тобто відділення об'єкта від фону.

Знайшовши глобальний мінімум енергетичної функції E , ми отримаємо масив прозорості $\hat{a} = \operatorname{argmin}_a E(a, \theta)$. Щоб мінімізувати енергетичну функцію, зображення описується як графік і шукається мінімальне перетин графіка. На відміну від GraphCut, алгоритм GrabCut розглядає пікселі в просторі RGB, тому для опису статистики кольорів використовується модель суміші Гауса (GMM).

4 СПЕЦІАЛЬНІ МЕТОДИКИ ДЛЯ РОЗПІЗНАВАННЯ ОБЛИЧЬ

4.1 Сегментація обличчя

Сегментація обличчя – основне завдання аналізу зображення обличчя. При сегментації особи комп'ютерний алгоритм сегментує зображення обличчя відповідно до різних областей обличчя. Семантична сегментація обличчя дозволяє комп'ютеру зрозуміти зміст зображення на рівні пікселів. Таким чином, для семантичної сегментації особи також використовується низка складних ознак.

Багато завдань аналізу особи можуть виграти від точної сегментації обличчя, наприклад, розпізнавання виразу обличчя, оцінка пози голови, виявлення орієнтирів обличчя, аналіз настроїв. залишається відкритою проблемою, особливо для зображень, знятих у необмежених умовах.

Сегментація області обличчя зазвичай сприймається як відправна точка більшості завдань аналізу зображення особи. Сегментація обличчя вважається важливим та проміжним етапом для подальшого аналізу зображень людської особи. Це включає програми з області ідентифікації та розпізнавання на основі біометричних даних, індексації людини і управління роботами, аж до розуміння медицини і психічного стану. Сегментація обличчя також відіграє вирішальну роль у розробці різних інтелектуальних середовищ.

Декілька факторів навколишнього середовища сприяють надійній сегментації обличчя та впливають на продуктивність ефективної системи сегментації особи. До них відносяться, але не обмежуються ними; оклюзії, зміна умов освітлення, шум у різних формах, зміна міміки та положення голови тощо. кількість доступних наборів даних для сегментації обличчя мінімальна.

Декілька програм комп'ютерного зору виразно і вичерпно покладаються на надійні вихідні дані сегментації обличчя. Таким чином, нижче наведені типові завдання аналізу зображення людини, які значною мірою залежать від

точної сегментації обличчя.

Збереження та завершення ідентичності особи: через неправильну природу зображень обличчя завершення особи є досить складним завданням. Завершення особи відноситься до завдання заповнення тих областей, які були пропущені з тієї чи іншої причини. Ці пропущені області заповнені реалістичним синтезованим вмістом. У класичних методах спочатку проводиться пошук локальної інформації, щоб знайти деякі існуючі закономірності на зображенні особи.

Усунення розмиття обличчя: з розвитком та інноваціями у підходах до аналізу особи методи усунення розмиття зображення людини прийняли нові напрямки. Проблема усунення розмиття обличчя вирішується за допомогою семантичних сигналів між різними областями зображення людини. Оскільки людське обличчя має різні семантичні компоненти (очі, ніс, рот та підборіддя), семантичні сигнали надають достатню інформацію для відновлення зображення.

Вилучення орієнтирів особи: Орієнтири особи відіграють важливу роль в аналізі зображення обличчя людини. Як правило, орієнтири особи включають важливі області обличчя, такі як ніс, рот, брови та очі. Це набір високорівневих ознак, які легко розрізнити неозброєним оком. Зазвичай, орієнтири особи можна знайти з допомогою традиційної парадигми машинного навчання. Він включає навчання моделей машинного навчання рис обличчя з використанням всеосяжного набору даних. Однак ці методи показали нижчу продуктивність у необмежених умовах, наприклад, при зображеннях, що перекриваються, а також у диких, низьких і неоптимізованих умовах освітлення. Для подолання цієї проблеми було доведено, що вилучення орієнтирів особи у вигляді семантичної сегментації є оптимальним способом.

Заміна обличчя: перенесення однієї особи з вихідного зображення на особу, що з'являється на цільовому зображенні, для створення невідредагованих і реалістичних результатів зазвичай називається заміною обличчя.

Прикраса обличчя: макіяж робить людей більш привабливими і

красивими. На ринку є кілька комерційних систем макіяжу для особи. Наприклад, віртуальна зачіска забезпечує віртуальний підхід, що дозволяє спробувати різні зачіски. Деякі косметичні елементи, такі як губна помада, підводка для очей тощо, використовуються для прикраси при макіяжі.

Оцінка пози голови: оцінка пози голови передбачає орієнтацію голови зображення обличчя. Данна задача, в основному, використовується у аналізі відео.

Вираз обличчя, вік, раса та розпізнавання статі. У психологічній літературі стверджується, що існує тісна взаємодія між частинами обличчя та кількома іншими прихованими змінними в особах. Ці приховані змінні включають вираз обличчя, інформацію про вік, пол і расу і т. д. Сегментація особи вносить інформацію про різні частини обличчя, що допомагає зоровій системі людини ідентифікувати ці змінні. У літературі повідомляється про деякі чудові методи, які сегментують зображення обличчя на різні семантичні частини, такі як ніс, очі, рот, волосся, шкіра та фон. Ця щільна інформація семантичного класу потім використовується для моделювання структури для вищезазначених завдань аналізу особи.

Сегментація портрета: Сегментація портрета – це мистецтво фотографії та живопису. Художники прагнуть виділити портретний живопис і домінувати над його оточенням, наприклад, роблячи його різкішим і яскравішим. Сучасний цифровий світ може обробляти портретне зображення за допомогою фотографії.

Розпізнавання осіб. Десять років тому розпізнавання осіб було складною сферою комп'ютерного зору. Завдяки нещодавнім розробкам у галузі методів комп'ютерного зору розпізнавання осіб у цей час є зрілою областю досліджень із безліччю оптимізованих алгоритмів. У SOA розпізнавання обличчя також досліджується за допомогою сегментації осіб. У цих підходах зображення обличчя спочатку сегментується на виділення та непомітні області, а потім виконується процес розпізнавання обличчя

З урахуванням прогресу глибоких нейронних мереж за останні кілька

років у різних галузях, найкращі результати для семантичної сегментації найчастіше досягаються за допомогою згорткових нейронних мереж, у тому числі коли дані слабо розмічені. Справді, проблема низького рівня розміченості даних у семантичній сегментації є досить важливою, оскільки для кожного пікселя визначити його належність з високою точністю — завдання, яке потребує високих витрат часу та не завжди високу точність. Однак, поєднання добре розмічених даних зі слабо розміченими даними (наприклад, з точністю до рамок, що обмежують) покращує продуктивність моделі. Для завдання сегментації добре себе показали FCN (fully-convolutional networks) — повнозвертальні мережі, що дозволяють працювати із зображеннями довільного розміру, а на виході видавати теплову карту знаходження класів на зображенні через серію згорток.

4.2 Бази даних для сегментації

Процес парсингу – це автоматичне вилучення великого масиву даних із веб-ресурсів, яке виконується за допомогою спеціальних скриптів. Існують доступні набори даних сегментації осіб, які можна використовувати як орієнтири для навчання алгоритмів, які будуть виконувати сегментацію обличчя.

HELEN складається з зображень осіб, зібраних за допомогою Flickr (Flickr – фотохостинг, призначений для зберігання та подальшого використання користувачем цифрових фотографій та відеороликів), які відрізняються великою різноманітністю порівняно з іншими наборами даних SOA. Алгоритм виявлення осіб було запущено для отримання інформації про частини особи. Усі зображення якісні, оскільки неякісні після збору відфільтровувалися. Ці зображення були анотовані вручну за допомогою Amazon Mechanical Turk, що допомагає якимось чином легко знаходити частини обличчя, такі як очі, ніс, рот і т.д. База даних складається з 2330 зображень осіб, з 11 мітками. Ці ярлики включають фон, шкіру, ніс, ліву брову, праву брову, ліве око, праве око, верхню

губу, нижню губу, внутрішню частину рота та волосся. Вихідна база даних розділена на 2000 навчальних, 230 перевірочних та 100 перевірочних зображень. Також існує покращена база даних HELEN*.

FASSEG містить близько 500 зображень осіб, які були вручну розділені на шість класів осіб, включаючи ніс, волосся, рот, шкіру, очі та фон. База даних також складається з чотирьох підмножин зображень осіб. Три підмножини є фронтальними зображеннями, а саме frontal 01, 02 і 03, тоді як четверте підмножина, а саме Multipose01, складається з зображень обличчя в декількох позах від -90° до $+90^\circ$.

FASSEG V-2 містить 150 зображень осіб, взятих із іншої бази даних Siblings. Зображення в цій версії являють собою фронтальні зображення з високою роздільною здатністю, зняті в різних орієнтаціях, мінливих умовах освітлення і з різними виразами обличчя.

FASSEG V-3 містить 294 зображення особи в декількох позах, які також були вручну анотовані за допомогою програмного забезпечення для редагування зображень. Ці зображення мають різні пози в діапазоні від -90° до $+90^\circ$. Різниця між двома позами становить 30° . Ці зображення спочатку були взяті з бази даних Pointing'04.

LFW База даних позначених осіб у дикій природі (LFW) з частинами, поміченими (PL), яка також відома як (LFW-PL), не надає мітки класів для всього зображення обличчя. LFW-PL містить 2927 зображень осіб, які були вручну розбиті на три класи: шкіра, фон та волосся. Зображення були зібрані з Інтернету, а розпізнавання обличчя було виконано вручну. Усі зображення збираються у необмежених умовах. LFW розділений на 1500 навчальних, 520 перевірочних та 927 тестових зображень.

Multi-Face база даних також була зібрана у необмежених умовах. Multi-Face містить кілька обличчя на одному зображенні. Це велика база даних з 9645 зображень осіб. Розмітка по пікселям виконувалася для трьох класів: шкіра, фон та волосся. Вихідні зображення бази даних розділені на 9045 навчальних, 200

перевірочних та 200 тестових зображень. Розмір зображення був збережений досить великим (512×512), щоб підтримувати різний контент обличчя у стані.

Figaro1K містить 1050 зображень. Усі зображення були анотовані лише класу волосся. Кожне зображення позначено однією з семи зачісок, включаючи хвилясті, прямі, кучеряві, ексцентричні, коси, короткі та дреди. Оскільки база даних була спеціально розроблена для сегментації волосся, обличчя часто не видно на зображенні. У деяких випадках одного кандидата захоплюють кілька зразків. З цієї причини цю базу даних можна використовувати лише для сегментації волосся.

CelebA – це велика база даних, що містить понад 200 000 зображень. Кількість учасників бази також досить велика (10 000). Ці зображення було знято у необмежених умовах. На зображенні CelebA можна побачити кілька неправильних позначень класу волосся. Тільки чотири класи волосся, включаючи чорне, каштанове, сиве та світле, визначені для тону волосся.

На основі цих баз даних можна реалізувати процес сегментації особи різними шляхами.

4.3 Геометричні методи

Переваги та недоліки: ці методи мають деякі переваги, порівняно з іншими описаними методами. Наприклад, геометричні методи стійкі до пози, виразу обличчя, варіацій переміщення та обертання та умов освітлення. Для цих методів потрібно дуже мало інформації; отже, обчислювальні витрати набагато менші. Методи аналізу осіб на основі ASM також мають низку серйозних недоліків. Ці методи ігнорують інформацію про текстуру шкіри та інших частин обличчя, яка є важливою підказкою сегментації різних частин обличчя. В результаті релевантна інформація втрачається на етапі отримання ознак.

Геометричні методи засновані на початковій локалізації деяких точок особи на зображенні обличчя та подальшому аналізі обличчя. Ці методи залежать від активного моделювання форми (ASM), яке використовується для

статистичного моделювання форми різних частин особи. Ці методи також ґрунтуються на розпізнаванні підказок серед різних частин особи. Методи на основі ASM аналогічні тому, як людський мозок ідентифікує різні частини обличчя. У геометричних методах частини обличчя спочатку локалізуються, та потім проходять попередню обробку. Після попередньої обробки інформація про орієнтири витягується та застосовується афінне перетворення, що гарантує, що зіниці на зображенні не будуть порушені щодо розташування після вирівнювання. В результаті складові особи мають однакові розміри на різних зображеннях обличчя.

Метод сегментації обличчя за допомогою орієнтирів особи було представлено Segundo et al. [3]. Платформа була розроблена для автоматичного процесу розпізнавання осіб з інформацією про глибину. Метод сегментації поєднує в собі виявлення країв, аналіз форми та кластеризацію для виділення областей обличчя, тоді як виявлення орієнтирів поєднує інформацію про кривизну поверхні для пошуку очей та носа. У роботі також розглядався вплив сегментації особи на розпізнавання осіб.

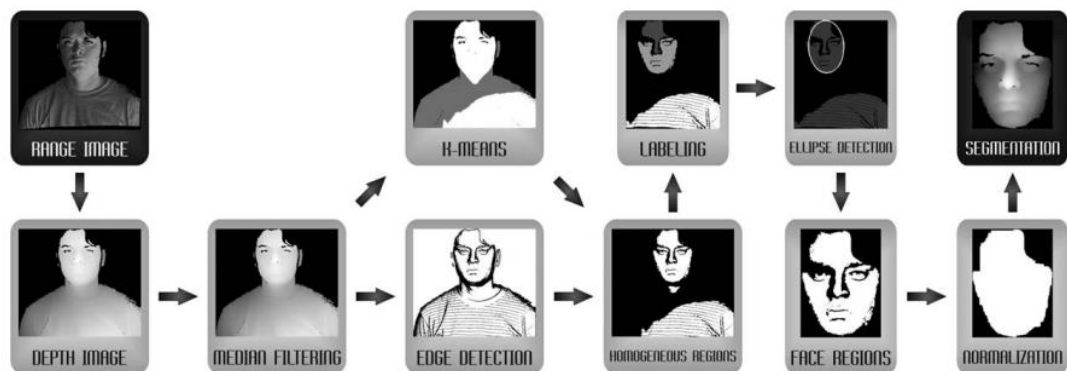


Рисунок 4.1 – Діаграма представленого алгоритму сегментації

Метод семантичного розбору особи, запропонований у [4], заснований на конкретній інформації про позу, яка закодована в наборі ключових точок або орієнтирів. Платформа об'єднала аналіз особи та оцінку орієнтирів особи в одній моделі. Запропонована структура була заснована на глибоких згорткових нейронних мережах (CNN). Модель була розроблена на стандартних наборах

даних LFW та HELEN і показала найкращі результати порівняно з іншими процесами SOA.

4.4 Цілісні методи

Ці методи мають деякі переваги проти іншими конкурентними методами. Наприклад, це прості методи із простою стратегією реалізації. Ці методи не вимагають будь-яких негативних навчальних даних на етапі навчання. Так само шаблони для етапу навчання можуть бути додані в будь-який час для розширення, що дозволяє архітектурі при необхідності адаптуватися до змін умов. Ці методи підходять як для зображень з високою, так і низькою роздільною здатністю. Є і недоліки. Наприклад, цілісні підходи припускають, що система вже розпізнає головну частину. Помилка локалізації, що виникає, знижує точність системи. Потрібна велика кількість навчальних даних, що робить систему обчислювально дорогою.

У цілісних методах ознаки витягуються із статистичної інформації, розглядаючи зображення як одномірний вектор. Ці методи завжди передбачають певний взаємозв'язок між властивостями 2D зображення обличчя. Статистичні методи навчання використовуються для побудови класифікатора за великою кількістю зображень.

У методі, запропонованому Уоррелом і Прінсом [5], аналіз осіб розглядався з погляду маркування зображень. Спочатку навчався попиксельний унарний класифікатор, а потім оцінювалося щільне маркування зображень обличчя для чотирьох частин особи. Передбачувані частини особи включали рот, волосся, вуса та капелюх. Пропонований метод враховував великомасштабні варіації як форми, і характеристик зовнішнього вигляду на зображеннях різноманітних осіб. Для класифікації автори використали класифікатор Adaboost.

У роботі [6] запропонували алгоритм сегментації особи, зокрема, для розпізнавання електронних документів, що засвідчують особу. Запропонована

модель сегментувала зображення обличчя на п'ять областей: волосся, шкіра, фон, плечі та оббивка. Пропонований алгоритм складався з двох кроків: пересегментація та маркування осіб. Спочатку зображення сегментувалося на однорідні області, та був маркування областей. Для сегментації було адаптовано стратегію сегментації за середнім зрушенням, тоді як маркування використовувався класифікатор Adaboost.

Інший метод, представлений у [7], призначений для розбору особи та покращення обличчя в єдиній структурі. Перша функція давала рекомендації щодо макіяжу обличчя. Візуально схожа особа була знайдена з бази зображень на першому етапі. Друга функція переносила еталонний макіяж на бажане обличчя. Було зафіксовано п'ять критеріїв переносу з еталонної моделі на бажану модель.

4.5 CRF

Умовні випадкові поля (CRF – Conditional random fields) часто використовуються для маркування зображень і за допомогою цих методів добре моделюються сусідні області. Однак CRF має обмеження при роботі зі складними сценаріями. Крім CRF, обмежена машина Больцмана (RBM – restricted Boltzmann machine) використовується для моделювання глобальних форм, які створюються моделями сегментації. У деяких недавніх методах CRF також поєднуються з архітектурами глибокого навчання, що дає перспективні результати. У цьому підрозділі обговорюються різні методи, у яких CRF, CRF поєднують RBM та CRF у поєднанні з архітектурами глибокого навчання.

Алгоритм аналізу особи за допомогою CRF був запропонований Ханом та співавт. [8]. У запропонованій моделі кожному вузлу відповідав суперпіксель, тоді як сусідні суперпікселі з'єднані ребрами з вузлами. Набір міток було розширено до шести класів, включаючи рот, ніс, волосся, шкіру, фон та очі. Запропонована модель була заснована на трьох наборах даних FASSEG, Figaro та LFW.

Розріджені FCN для аналізу осіб були запропоновані Dong et al. [9]. У порівнянні з іншими методами, FCN продемонстрували сильні можливості у навчанні репрезентації, особливо для семантичної сегментації, повністю використаної авторами. Оскільки FCN в основному страждали від надмірності проблем, які були вирішені авторами за допомогою спеціальної стратегії. Групове ласо та регуляризація внутрішньогрупового ласо використовувалися для розрідження згорткових шарів повністю згорткових мереж.

4.6 Гібридні методи

Гібридні методи також відомі як багатозадачні методи. Оскільки сегментація особи тісно пов'язана з декількома іншими завданнями аналізу зображення обличчя, включаючи оцінку положення голови, розпізнавання виразу обличчя, розпізнавання статі тощо, деякі методи довели, що кращу продуктивність можна отримати за допомогою багатозадачної структури замість моделі з одним ізольованим завданням. Гібридні моделі спільно навчають кілька мереж, які враховують взаємодію між цільовим завданням сегментації обличчя та деякими іншими другорядними завданнями, такими як оцінка пози голови, розпізнавання виразу обличчя тощо. Такі методи вимагають розмічених даних із усіх завдань, пов'язаних із зображенням обличчя, і в таких випадках фаза навчання стає досить громіздкою, тому що виконується дедалі більше завдань.

Цікава гібридна модель, що враховує різні риси обличчя, була представлена Huang et al. [10]. Модель сегментації навчалася обмеженому наборі зображень. Запропонована багатозадачна модель використовувалася для розпізнавання облич у необмежених умовах. Структура також була розширена для оцінки пози голови з трьома простими варіантами, включаючи фронтальні, ліві та праві зображення профілю.

4.7 Методи глибокого навчання

Продуктивність традиційних методів машинного навчання була вражаючою для зображень, зібраних у контрольованих лабораторних умовах. Однак коли ці традиційні методи машинного навчання були піддані впливу зображень, зібраних у дикій природі, їхня продуктивність значно знизилася. На відміну від традиційних методів машинного навчання, базові методи глибокого навчання вивчають вищий рівень абстракції із вхідних даних зображення обличчя. Завдяки цим методам потреба у розробці ознак значно знижується. Поряд із перевагами ці методи мають і серйозні недоліки. Глибоке навчання - дуже складна процедура, яка вимагає від людини вибору. Наприклад, налаштування передавальної та активаційної функцій алгоритму навчання та сон. Оскільки ці методи з'явилися порівняно недавно, все ще необхідно встановити їх корисність для завдання сегментації особи.

В останні дні методи, що базуються на глибокому навчанні, показали відмінні результати в різних завданнях візуального розпізнавання. Вони покращують складніші сценарії комп'ютерного зору. Деякі обмеження у методах машинного навчання пом'якшуються із переходом цих методів на недавно представлені архітектури глибокого навчання. У цьому пункті різні методи сегментації людини, розроблені з допомогою глибокого навчання.

Халіл та ін [11] запропонували модуль розбору особи з використанням CNN та архітектури глибокого навчання. Продуктивність запропонованої структури оцінювалися за допомогою трьох наборів даних HELEN, LFW та FASSEG.

Лін та ін [12] запропонували метод розбору особи, натхненний фізіологічною системою зору людини. Це також була згорткова нейронна мережа, яка вирішувала дилему між областю обмеженого розміру, що цікавить, і областю периферійної інформації. Набір даних HELEN також був перемаркований, оскільки більшість областей на зображеннях HELEN неправильно позначені сусідніми областями. Автори також використовували LFW разом із HELEN для експериментів.

Сайто та ін [13] запропонували метод парсингу осіб, що враховує дике середовище. За словами авторів, продуктивність фреймворків для синтаксичного аналізу осіб знижувалася при впливі наборів даних, зібраних у неконтрольованих умовах. Запропонований алгоритм також враховував такі умови, як оклюзія, аксесуари, візуальні артефакти та інші фактори довкілля. У пропонованому методі спочатку маскували безособові області. Базова концепція алгоритму заснована на перепрофілюванні CNN, які спочатку були розроблені для загальної семантичної сегментації зображень. Продуктивність системи була покращена за рахунок конкретних стратегій збільшення даних та розробки кращих додаткових характеристик.

ВИСНОВКИ

У рамках кваліфікаційної роботи було розглянуто популярні існуючі методи сегментації зображень та з огляду на зазначену природу адитивних завад можна зробити висновки, що методи виділення кордонів та метод зв'язкових компонентів можуть використовуватися для уявлення форми об'єкта, виключаючи багато інших важливих характеристик, необхідних для якісного подальшого використання сегментованого зображення, а в умовах дії адитивних завад навіть межі об'єктів не будуть виділені досить чітко. Найважливіша перевага порогових методів є їх обчислювальна ефективність, тож із цієї групи ефективність сегментації в умовах завад може показати хіба що пороговий метод із адаптивним порогом. Сегментація методами нарощування областей дають чіткий результат щодо меж об'єктів, проте основною проблемою методу вододілів є надмірна сегментація, оскільки всі межі і шуми подаються в градієнті, що робить необхідним процес видалення. Тож без попередньої обробки буде малоефективним у сегментації людей. Текsturні методи сегментації чудово показують себе на зображеннях гарної якості, та відсутності шумів. У випадку наявності адитивних завад, таких як гаусівський шум зображення просто не буде мати достатньої чіткості. Метод Grabcut може бути ефективним лише за умови наявності уважної людини, що ним оперує. Використання повнозвертальних нейронних мереж разом із базами попередньо розмічених зображень показують найкращі результати у більшості випадків сегментації. А з огляду на те, що деякі методи на їхній основі були спеціально розроблені з урахуванням таких адитивних завад як несприятливі погодні умови, нечіткість вихідного зображення через людський фактор, ці методи можна вважати найкраще пристосованими до роботи в умовах наявності адитивних завад.

ПЕРЕЛІК ПОСИЛАНЬ

1. Arulkumar Subramaniam, Jayesh Vaidya, Muhammed Abdul Majeed Ameen, Athira Nambiar, Anurag Mittal. Co-segmentation inspired attention module for video-based computer vision tasks, / Arulkumar Subramaniam, Jayesh Vaidya, Muhammed Abdul Majeed Ameen, Athira Nambiar, Anurag Mittal. // Computer Vision and Image Understanding. – 2022. – С. Volume 223.
2. Анатомія ока. Методи дослідження в офтальмології: навч. посібник для студентів мед. фак.-тів / Н. Г. Завгородня, Л. Е. Саржевська, О. М. Івахненко [та ін]. – Запоріжжя, 2017. – 76 с.
3. M. P. Segundo, L. Silva, O. R. P. Bellon and C. C. Queirolo, "Automatic face segmentation and facial landmark detection in range images", IEEE Trans. Syst. Man Cybern. B Cybern., vol. 40, no. 5, pp. 1319-1330, Oct. 2010.
4. A. S. Jackson and M. V. G. Tzimiropoulos, "A CNN cascade for landmark guided semantic part segmentation", Proc. Eur. Conf. Comput. Vis., pp. 143-155, 2016.
5. J. Warrell and S. J. D. Prince, "Labelfaces: Parsing facial features by multiclass labeling with an epitome prior", Proc. 16th IEEE Int. Conf. Image Process. (ICIP), pp. 2481-2484, Nov. 2009.
6. M. Subašić, S. Lončarić and A. Heđi, "Segmentation and labeling of face images for electronic documents", Expert Syst. Appl., vol. 39, no. 5, pp. 5134-5143, Apr. 2012.
7. S. Liu, X. Ou, R. Qian, W. Wang and X. Cao, "Makeup like a superstar: Deep localized makeup transfer network", arXiv:1604.07102, 2016, [online] Available: <http://arxiv.org/abs/1604.07102>.
8. K. Khan, N. Ahmad, K. Ullah and I. Din, "Multiclass semantic segmentation of faces using CRFs", TURKISH J. Electr. Eng. Comput. Sci., vol. 25, no. 4, pp. 3164-3174, 2017.

9. M. Dong, S. Wen, Z. Zeng, Z. Yan and T. Huang, "Sparse fully convolutional network for face labeling", *Neurocomputing*, vol. 331, pp. 465-472, Feb. 2019.
10. G. B. Huang, M. Narayana and E. Learned-Miller, "Towards unconstrained face recognition", *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, pp. 1-8, Jun. 2008.
11. K. Khan, M. Attique, R. U. Khan, I. Syed and T.-S. Chung, "A multi-task framework for facial attributes classification through end-to-end face parsing and deep convolutional neural networks", *Sensors*, vol. 20, no. 2, pp. 328, 2020.
12. J. Lin, H. Yang, D. Chen, M. Zeng, F. Wen and L. Yuan, "Face parsing with RoI tanh-warping", *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 5654-5663, Jun. 2019.
13. S. Saito, T. Li and H. Li, "Real-time facial segmentation and performance capture from RGB input", *Proc. Eur. Conf. Comput. Vis.*, pp. 244-261, 2016.