

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет інформаційних технологій

Кафедра інтелектуальних технологій

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня «магістр»

НА ТЕМУ:

Система діагностики пульсу людини на основі аналізу відео з використанням методів машинного навчання

Галузь знань: 12 - «Інформаційні технології»

Спеціальність: 122 «Комп'ютерні науки»

Освітньо-наукова програма «Технології штучного інтелекту»

Виконав: студент 2 курсу магістратури
групи ТШІ-21

Пащенко Роман Борисович

Науковий керівник:

к.т.н., доцент

Іларіонов Олег Євгенович

Кваліфікаційна робота допущена до захисту
рішенням кафедри *інтелектуальних технологій*

Протокол № _____ від «___» _____ 2021 р.

Зав. кафедри _____

Робота не містить елементів плагіату _____ Пащенко Р.Б.

Київ 2021

РЕФЕРАТ

Робота містить 63 сторінок тексту, 24 рисунки, 2 таблиці, 36 посилань на літературні джерела та 1 додаток.

Актуальність. Сучасні перспективи дистанційних неінвазивних діагностичних систем базуються на оптичних методах дослідження у галузях оптоелектроніки та нанотехнологій.

Оскільки, оптичним методам притаманні ряд значних переваг, а саме: безболісність, екологічність, нешкідливість, дешевизна, неінвазивність, компактність, висока швидкість передачі інформації, а також можливість вимірювання показників практично у будь-якій точці поверхні тіла, що дозволить використовувати різні модифікації оптоелектронних приладів для розв'язання цілого ряду спеціальних задач, пов'язаних із дослідженням показників локального стану поверхні тіла людини, що базуються на процесі оптичної реєстрації досліджуваних ділянок шкіри.

Наразі значною проблемою є проведення швидкої експрес-діагностики, особливо за умов сучасної пандемії.

Тому актуальним є вдосконалення методів і пристроїв діагностики, які дозволять провести загальне, регулярне обстеження та забезпечити доступність і масовість при загальному регулярному діагностуванні широких верст населення. Оптичні прилади візуальної реєстрації досить поширені на індивідуальному рівні, так, наприклад, прилади для оптичної тканинної оксиметрії - аналізатори об'ємного капілярного кровонаповнення м'яких біологічних тканин.

Метою даної роботи є дослідження проблеми розробки системи діагностики пульсу людини на основі аналізу відео з використанням методів машинного навчання та надати алгоритмічні частини у вигляді коду, розробка програмної системи та проведення експериментів.

Об'єктом дослідження є процес дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Предметом дослідження є практичні прийоми та алгоритми апаратного та програмного розв'язання задачі розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Новизна отриманих результатів. Вперше було запропоновано та реалізовано комплексний алгоритм розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Апробація результатів дослідження. Практично опрацьовано запропонований комплексний алгоритм розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини. 1 стаття була опублікована в журналі «Наука онлайн».

Структура та обсяг роботи. Кваліфікаційна робота на здобуття освітнього ступеня «магістр» складається з вступу, трьох розділів та висновків.

У *вступі* до магістерської роботи надано загальну характеристику роботи, визначено актуальність теми досліджень, сформульовано мету досліджень та наукову новизну отриманих результатів.

У *першому* розділі «Огляд літератури» проаналізовано сучасні відомості про стан питання, що досліджується, обґрунтовано актуальність досліджень, розглянуто ретроспективу рішень питання розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

У *другому* розділі надано методичні аспекти досліджень, що проводились, сформульовано методичну основу питання розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

У *третьому* розділі описано алгоритм програмного застосунку та приведено його інтерфейс, логічно обґрунтовано вибір архітектури та

методів розробки системи,. Проведено тестування програмного застосунку та зроблено аналіз отриманих результатів.

У висновках представлені підсумки проведених досліджень.

Ключові слова: розпізнавання облич людей, розпізнавання ЧСС людей, обробка зображень.

ABSTRACT

The thesis consists of 63 pages, 24 figures, 2 tables, 36 literature sources, and 1 annex.

Relevance. Modern prospects for remote non-invasive diagnostic systems are based on optical research methods in the fields of optoelectronics and nanotechnology.

Since optical methods have a number of significant advantages, namely: painlessness, environmental friendliness, harmlessness, cheapness, non-invasiveness, compactness, high speed of information transfer, as well as the ability to measure almost anywhere on the body surface, will use various modifications of optoelectronics devices to solve the whole a number of special tasks related to the study of indicators of the local state of the surface of the human body, based on the process of optical registration of the studied areas of the skin.

Now a big problem is the rapid diagnosis, especially in a modern pandemic.

Therefore, it is important to improve the methods and devices of diagnosis, which will allow for a general, regular examination and to ensure the availability and mass in the general regular diagnosis of the general population. Visual registration optics are quite common at the individual level, for example, devices for optical tissue oximetry - analyzers of volumetric capillary blood supply to soft biological tissues.

The **purpose** of the work is the study of the problem of developing a system of human heart rate diagnostics based on video analysis using machine learning methods and provide algorithmic parts in the form of pseudocode, software system development and experiments.

The **object of the study** is the process noninvasive remote counting pulse rate (HR) person.

The **subject of the study** practical techniques and algorithms for hardware and software solution of the problem of human face recognition in real time for the purpose of remote non-invasive calculation of human heart rate (HR).

The scientific novelty of the results. For the first time, a complex algorithm for recognizing a person's face in real time for the purpose of remote non-invasive calculation of human heart rate (HR) was proposed and implemented.

Approbation of research results. The proposed complex algorithm for real-time human face recognition for the purpose of remote non-invasive calculation of human heart rate (HR) is practically developed. 1 article was published in the journal "Nauka online".

The structure and scope of thesis. Qualifying work consists of an introduction, three chapters and conclusions.

In the introduction to the master's thesis the general characteristic of work is given, the urgency of a research theme is defined, the purpose of researches and scientific novelty of the received results is formulated.

The first chapter "Literature Review" analyzes current information about the state of the research, substantiates the relevance of research, and considers a retrospective of solutions to the recognition of human face in real time for remote non-invasive calculation of human heart rate (HR).

The second chapter presents the methodological aspects of the conducted research, formulates the methodological basis of the question of human face recognition in real time for the purpose of remote non-invasive calculation of human heart rate (HR).

The third chapter describes the algorithm of the software application and its interface, logically substantiates the choice of architecture and methods of system development. The software application was tested, and the obtained results were analyzed.

The conclusions present the results of the research.

Keywords: face recognition, human heart rate recognition, image processing.

ЗМІСТ

РЕФЕРАТ	2
ABSTRACT	5
Перелік умовних позначень і скорочень	9
Вступ	10
РОЗДІЛ 1. АНАЛІЗ СУЧАСНОГО СТАНУ ПИТАННЯ ТА ОБҐРУНТУВАННЯ ВИБОРУ ТЕМИ ДОСЛІДЖЕННЯ	13
1.1. Аналіз сучасного стану питання та обґрунтування вибору теми дослідження	13
1.2. Алгоритми і моделі задачі дистанційного визначення ЧСС людини	16
1.3 Постановка задачі	17
РОЗДІЛ 2. СТРУКТУРА АЛГОРИТМУ ДІАГНОСТИКИ ЧСС ПО ВІДЕО	19
2.1. Основні поняття процесу дистанційного визначення ЧСС людини	19
2.2. Модель відбиття світла шкірою людини	21
2.3. Порівняння алгоритмів отримання rPPG сигналу з даних відеоінформації	25
РОЗДІЛ 3. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ	44
3.1. Розробка алгоритму програми	44
3.2. Опис програмного застосунку	46
3.3. Опис даних для тестування	53
3.4. Аналіз результатів роботи програмного застосунку	56
Висновки	58
Література	59
Додатки	64

Перелік умовних позначень і скорочень

англ. – англійська

БД– база даних

год. – година

ЧСС – частото серцевих скорочень

ЧД – частота дихання

теор. – теоретичний

ІС – інформаційна система

СШІ – система штучного інтелекту

ШП – шумопониження

ЧП – частота пульсу

уд / хв - ударів в хвилину

ЗВО – заклади вищої освіти

ССС – серцево-судинна система

rPPG - Remote photoplethysmography, Дистанційна фотоплетизмографія

ОІ - область інтересу

АМЗВ - активна модель зовнішнього вигляду

Вступ

Важливість вчасної діагностики змін фізіологічного стану людей не викликає сумнівів. Кількість чинників, що можуть справляти вплив на здоров'я людей дуже велика і діагностика ЧСС може бути одним з перших показників зміни стану здоров'я. Можливості дистанційної неінвазивної діагностики відкривають перспективи автоматичного моніторингу стану людей в режимі реального часу.

Актуальність. Сучасні перспективи дистанційних неінвазивних діагностичних систем базуються на оптичних методах дослідження у галузях оптоелектроніки та нанотехнологій.

Оскільки, оптичним методам притаманні ряд значних переваг, а саме: безболісність, екологічність, нешкідливість, дешевизна, неінвазивність, компактність, висока швидкість передачі інформації, а також можливість вимірювання показників практично у будь-якій точці поверхні тіла, що дозволить використовувати різні модифікації оптоелектронних приладів для розв'язання цілого ряду спеціальних задач, пов'язаних із дослідженням показників локального стану поверхні тіла людини, що базуються на процесі оптичної реєстрації досліджуваних ділянок шкіри.

Тому актуальним є вдосконалення методів і пристроїв діагностики, які дозволять провести загальне, регулярне обстеження та забезпечити доступність і масовість при загальному регулярному діагностуванні широких верст населення. Оптичні прилади візуальної реєстрації досить поширені на індивідуальному рівні, так, наприклад, прилади для оптичної тканинної оксиметрії - аналізатори об'ємного капілярного кровонаповнення м'яких біологічних тканин.

Метою даної роботи є дослідження проблеми розробки системи діагностики пульсу людини на основі аналізу відео з використанням методів машинного навчання та надати алгоритмічні частини у вигляді псевдокоду, розробка програмної системи та проведення експериментів.

Для реалізації мети були поставлені такі завдання:

- Провести захват зображення і пошук обличчя або обличь на ньому за допомогою згорткових нейронних мереж;
- Порівняти ефективність алгоритмів отримання інформації щодо ЧСС на основі відеоінформації;
- Розробити алгоритм для діагностики ЧСС людини на основі відеоінформації та його програмну реалізацію.

Методика дослідження ґрунтується на основі методів математичного моделювання, пульсометрії та штучного інтелекту.

Об'єктом дослідження є процес питання дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Предметом дослідження є практичні прийоми та алгоритми апаратного та програмного розв'язання задачі розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Новизна отриманих результатів. Запропоновано та реалізовано комплексний алгоритм розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини.

Апробація результатів дослідження. Практично опрацьовано запропонований комплексний алгоритм розпізнавання обличчя людини у режимі реального часу з метою дистанційного неінвазивного підрахунку частоти пульсу (ЧСС) людини. 1 стаття була опублікована в журналі «Сучасна ветеринарна медицина».

Відповідно до поставленої мети дослідження визначено структуру роботи. Вона складається зі вступу, трьох розділів, __ підрозділів, висновків, списку використаних джерел із __ найменувань, __ додатків. Загальний обсяг роботи __ сторінок.

РОЗДІЛ 1. АНАЛІЗ СУЧАСНОГО СТАНУ ПИТАННЯ ТА ОБҐРУНТУВАННЯ ВИБОРУ ТЕМИ ДОСЛІДЖЕННЯ

1.1. Аналіз сучасного стану питання та обґрунтування вибору теми дослідження

Сучасна суспільна медицина – це, насамперед, попередження захворювань завдяки ранній діагностиці. Лікування захворювань на ранніх стадіях розвитку патології безперечно є більш ефективним для пацієнта [33]. Такий підхід передбачає впровадження нових технологій у процес діагностування і переходу до виявлення початкових змін, які можуть спричинити хворобу.

Неінвазивна дистанційна діагностика організму не матиме собі рівних у безпеці та зручності для пацієнтів, в тому числі для літніх людей, дітей, вагітних жінок. Сучасні дослідження у галузях оптоелектроніки та нанотехнологій вбачають перспективи дистанційних неінвазивних діагностичних систем, що базуються, в тому числі, і на оптичних методах.

Загальновідомо, що «частота пульсу — це кількість пульсових хвиль за 1 хвилину. У здорової людини кількість пульсових хвиль відповідає кількості серцевих скорочень і дорівнює 60—80 за 1 хвилину. Частоту серцевих скорочень понад 80 називають тахікардією, а частоту серцевих скорочень менше ніж 60 ударів за 1 хвилину — брадикардією» [31]. Поява тахікардії, брадикардії і аритмії свідчить про певні проблеми із здоров'ям у людини і вимагають подальшого медичного втручання. Фізіологічно нормальна мінливість частоти пульсу залежить від великої кількості різноманітних чинників:

1. Вікова мінливість: найвища частота серцевих скорочень спостерігається у перші роки життя. Відразу після народження ЧСС у немовля в нормі може коливатися до 190 ударів за хвилину. У дітей віком 1-3 місяці нормальна частота пульсу коливається в межах 145-170 ударів/хвилину [25];

2. Мінливість внаслідок фізичного навантаження виражається як прискорення пульсу за виконання інтенсивної фізичної роботи [33];
3. Статева мінливість ЧСС: частота серцевих скорочень у чоловіків, за рівних умов, як правило, на 5—10 ударів за хвилину менша, ніж у жінок[31];
4. Емоційна мінливість: при станах збудження нервової системи ЧСС збільшується, при гальмівних станах – стає більш повільною, наприклад, під час сну ЧСС зменшується на 10-20 ударів в хвилину[22];
5. Мінливість ЧСС з урахуванням стану тренуваності людини: у більш тренуваних людей пульс відносно повільніший [22].

Однією з основних причин прискореного пульсу людини лікарі визнають компенсаторні фізіологічні аспекти, які спостерігаються за підвищення температури тіла. Наприклад, відповідно до правила Лібермейстера, підвищення температури тіла людини на 1°C спричинює підвищення частоти пульсу на 8—10 ударів за 1 хвилину [28].

Таким чином, показник ЧСС використовується в медичній та спортивній практиці як фізіологічний показник нормального ритму серцебиття і є важливим елементом первинного діагностики фізіологічного стану організму та диференціації хворобливих станів [24].

Так, з огляду на основну із найактуальніших проблем сьогодення, встановлено, що інфікування COVID-19 у 67% людей призвело до вірогідного підвищенні ЧСС за кілька днів до появи клінічних симптомів [4].

Відносна варіація кольору шкіри (яка не фіксується візуально людиною), може бути використана для визначення ЧСС [30]. Також є дані щодо можливості візуальної реєстрації рухів з малою просторовою амплітудою, які важко або неможливо побачити людям (дихальні рухи, пульсація судів) [12]. Тенденції та результати цих досліджень свідчать

про можливість фіксації подібних візуальних та механічних сигналів та наступного комплексного аналізу просторової та часової обробки відео з метою посилення варіації досліджуваних чинників (кольору шкіри на видимих ділянках лиця) в межах смуги частот, пов'язаних з ЧСС, або посилити рухи грудної клітини, щоб зафіксувати дихальні рухи та проаналізувати частоту дихання (ЧД) у людини, яка рухається, використовуючи лінійну модель просторового наближення [12].

1.2. Алгоритми і моделі задачі дистанційного визначення ЧСС людини

Загалом, для отримання пульсу з зображення шкіри необхідно розуміти, з яких складових воно складається. Однією з моделей є наступна:

Джерело світла має постійний спектр, але його інтенсивність може змінюватися. Також, інтенсивність світла залежить від відстаней від шкіри до записуючого пристрою (камери) та до джерела світла.

Колір шкіри, видимий на камері - змінюється з часом. Зміни які впливають на інтенсивність відображеного світла викликані рухом тіла, а зміни, які змінюють колір шкіри - викликані пульсом.

Реалізація методів предикторної діагностики запропонована в розробці Інституту кібернетики ім. В.М.Глушкова НАНУ. Колективом вчених інституту створено апаратні і програмні складові системи, призначені для пульсової діагностики, яка, по своїй суті, є цілим вимірювальним комплексом з широкими функціональними і інформаційними можливостями. Головними предикторами захворювань, які дозволяє визначити комплекс - є ендотеліальна дисфункція, зміна швидкості розповсюдження пульсових хвиль, порушення часових та амплітудних параметрів контура пульсових хвиль. «Інформаційно-вимірювальний комплекс пульс-кардіологічної діагностики» [26] призначений для вирішення низки завдань, в тому числі для моніторингу

стану серцево-судинної системи у режимі реального часу; для наочного відображення динаміки частоти пульсу; для аналізу кривих показників пульсу; діагностики дисфункції епітеліального шару за його кольором; визначення вегето-судинної дистонії за швидкістю розповсюдження пульсових хвиль; визначення стану окремих судин; діагностики порушення ритму серця та візуалізації результатів; фіксації порушень психоемоційного стану за допомогою визначення показника ЧСС та відстеження динаміки змін психоемоційного стану окремих індивідів; - прогнозу виникнення кардіо-васкулярних станів; та ін.. [14].

Проблема нерозповсюдження вірусу COVID-19 при забезпеченні навчального процесу піднімається в роботі [15]. В роботі йдеться про впровадження експрес-діагностики, скринінгу і моніторингу функціонального стану людей з метою визначення та профілактики багатьох дисфункцій з метою запобігання та попередження серйозних захворювань. Виходячи із ситуації, що склалась у світі, та неможливості контролю за станом усіх присутніх в закладах вищої освіти (ЗВО) в період пандемії COVID-19, пропонується розробка комплексної інформаційної системи, складові модулі якої у вигляді спеціальних датчиків будуть частково знаходитись в студентських аудиторіях.

Метою цієї роботи також є розробка технології неінвазивної пульсової експрес-діагностики стану серцево-судинної та інших органів і систем людини, апаратна частина якої може бути використана як динамічний засіб визначення функціонального стану людини в місцях скупчень великої кількості людей, наприклад, студентів в аудиторіях. Вищезазначена система представляє собою поєднаний програмно-апаратний комплекс швидкого визначення стану людини і одночасного дослідження хімічного складу повітря в аудиторіях. Цей комплекс надасть змогу проводити дослідження функціонального стану серцево-судинної системи в режимі реального часу з можливістю визначення основних симптомів клінічно хворих на корона вірусну інфекцію, так, за комплексом

ознак, і станів носіїв хвороби без явно виражених симптомів. Планується апаратне дистанційне визначення стану здоров'я за основними показниками: температура тіла; насичення киснем крові; частота серцевих скорочень. Дослідження якості повітря в приміщеннях передбачають вимірювання температури, вологості; швидкості руху повітряних мас; складу повітря за концентрацією аероіонів.

1.3 Постановка задачі

Покращити один з існуючих на даний момент алгоритмів визначення пульсу на основі відеоінформації. Досягнення такої мети передбачає виконання наступних завдань:

- 1) Провести аналіз існуючих алгоритмів визначення пульсу на основі відеоінформації.
- 2) Проаналізувати слабкі і сильні місця відомих алгоритмів та посилити один з них на основі такого аналізу.
- 3) Розробити програмну реалізацію нового алгоритму.
- 4) Провести експериментальне дослідження та оцінити новий алгоритм визначення пульсу на основі відеоінформації.

РОЗДІЛ 2. СТРУКТУРА АЛГОРИТМУ ДІАГНОСТИКИ ЧСС ПО ВІДЕО

2.1. Основні поняття процесу дистанційного визначення ЧСС людини

Структура алгоритму діагностики ЧСС з використанням відео фіксації представлена на рис.2.1.

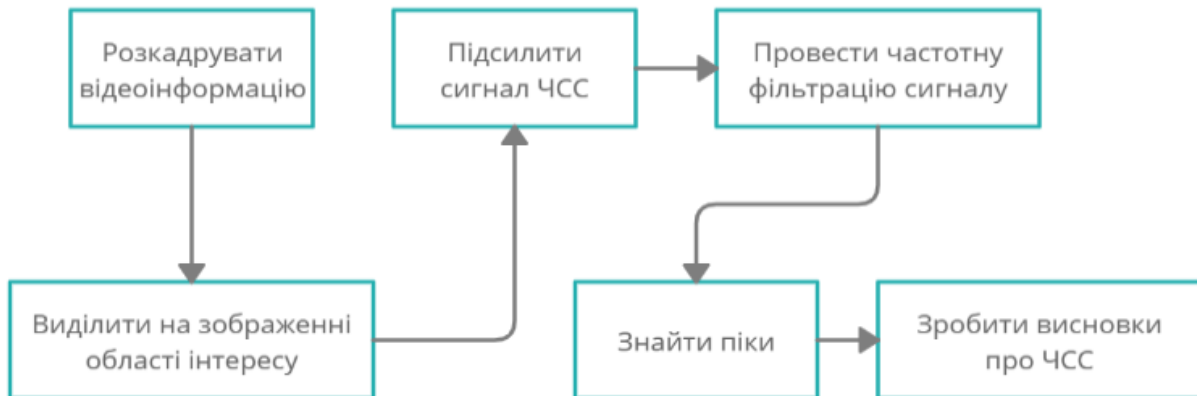


Рисунок 2.1. Структура алгоритму діагностики ЧСС на відео.

Роздивимось кожний крок детальніше.

Розкадрування – процес, під час якого з відео виділяються окремі кадри, або фрейми. В подальшому розглядається їх послідовність

Цей процес можна виконати за допомогою існуючих бібліотек будь якої сучасної мови програмування, або за допомогою відкритої інформації про формат зберігання відео.

Областями інтересу (ОІ) називаються, в нашому випадку, такі області зображення, на яких знаходиться лише шкіра. Це пов'язано з тим, що серцеві скорочення можна відслідкувати за змінами кольору шкіри, однак вони ніяк не проявляють себе на інших частинах тіла (таких як волосся, борода, брови, очі, тощо). Найчастіше виділяють верхню частину обличчя, виключаючи очі, хоча іноді також беруть до уваги інші ОІ, такі як лоб і щоки.

Так як задача знаходження обличчя на зображенні є дуже глибоко розглянутою в останні роки, то логічно використати напрацювання з цієї області для знаходження областей інтересу. В нашому випадку доцільно використати алгоритми, які будують так звану активну модель зовнішнього вигляду, або АМЗС.



Рисунок 2. 2 - Приклад активної моделі зовнішнього вигляду обличчя

Ця модель використовується для отримання кластерів пікселів шкіри і обчислення орієнтації голови. Це дозволяє нам надійно відслідковувати пікселі в цій ОІ протягом всього відео. Орієнтація голови використовується для вимірювання і відстеження нахилу, крену і відхилення кутів голови.

Для підсилення сигналу ЧСС, або, як ще кажуть, виділення сигналу rPPG (remote photoplethysmography), використовують низку алгоритмів, які описані в розділі 2.2 .

В загальному, з усіх пікселів ОІ виділяють деяку інформацію, найчастіше - середні значення по кожному з каналів R, G та B. Ці усереднені значення створюють 3-канальний сигнал. З нього отримують одновимірний rPPG сигнал, який підсилює значення пульсової компоненти.

Для частотного аналізу найчастіше використовують швидке перетворення Фур'є. Таким чином вираховуються і видаляються шуми,

спричинені ритмічними рухами (такими як крокування або виконання вправ) або записуючим пристроєм.

Після того як очищений від непотрібної інформації rPPG сигнал було отримано, ми можемо використати алгоритми пошуку піків сигналу для отримання кількості серцевих скорочень за період зйомки і апроксимувати його до значення ЧСС за хвилину.

2.2. Модель відбиття світла шкірою людини

Найчастіше, для визначення ЧСС людини на основі відеоінформації використовують модель дихроматичного відбиття. Вона стверджує, що якості відбитого від тіла світла залежать від трьох компонентів - від відбиття від поверхні тіла (дзеркального відбиття), відбиття від внутрішніх компонент (дифузного відбиття) і від інтенсивності освітлення.

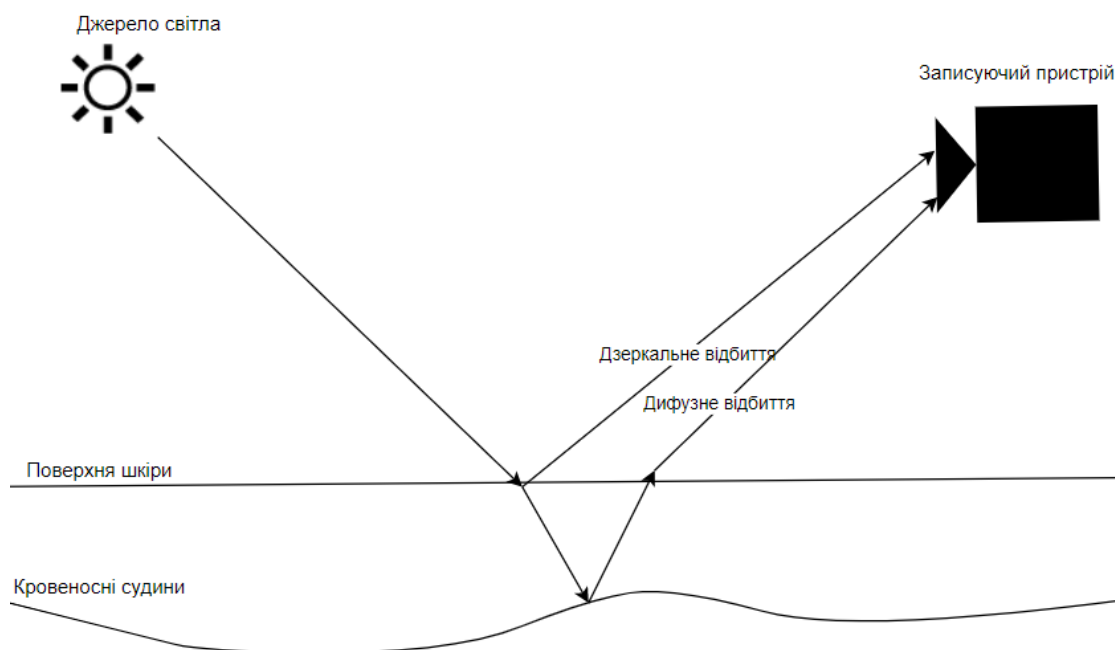


Рисунок 2.3 - Модель відображення шкіри, що містить дзеркальні та дифузні відображення, де лише дифузне відображення містить пульсуючу інформацію

В нашому випадку також додаються шуми, накладені записуючим пристроєм, які від інтенсивності освітлення вважаються незалежними. Таким чином, можна прийти до висновку, що відображення кожного пікселя на шкірі в записанному кадрі може бути описане наступною функцією (2.1), залежною від часу:

$$C_k(t) = I(t) (v_s(t) + v_d(t)) + v_n(t) \quad (2.1)$$

де:

$C_k(t)$ - це стан k-го пікселя шкіри;

$I(t)$ - визначає рівень яскравості освітлення, включає в себе зміни, накладені змінами яскравості джерела світла і змінами відстаней між джерелом, тілом та записуючим пристроєм;

$I(t)$ в дихроматичній моделі модулюється двома компонентами: дзеркальним і дифузним відбиттям ($v_s(t)$, $v_d(t)$ відповідно). Залежність в часі впливає з рухів тіла і пульсу (відповідно);

$v_n(t)$ - шуми, накладені камерою.

Дзеркальне відображення - відбиття світла від поверхні тіла, не містить жодної пульсуючої інформації. З цього випливає, що його спектральний склад співпадає з спектральним складом джерела світла. Воно залежить від часу через те, що рухи тіла змінюють геометричне розташування поверхні тіла відносно джерела світла і камери. Його можна записати наступним чином (2.2):

$$v_s(t) = u_s(s_0 + s(t)) \quad (2.2)$$

де:

u_s - визначає одиничний вектор кольору спектра світла;

s_0 і s_t - стаціонарні та змінні частини дзеркального відображення. $s(t)$ - спричинений рухом.

Дифузне відбиття визначається поглинанням і розсіюванням світла в поверхні тіла. Гемоглобін і меланін, які знаходяться в шкірі - спричиняють

специфічну хроматику v_d . Також, v_d змінюється через зміну кількості крові. Ця складова - залежить від часу (2.3). Таким чином:

$$v_d = u_d d_0 + u_p p(t) \quad (2.3)$$

де:

u_d - визначає одиничний вектор кольору шкіряного покриву;

d_0 - сила відбиття (яка є незмінною);

u_p - визначає силу пульсуючих змін в RGB; $p(t)$ - сигнал пульсу.

Підставляючи (2. 2) і (2. 3) в (2.4), отримуємо:

$$C_k(t) = I(t) (u_s(s_0 + s(t)) + u_d d_0 + u_p p(t)) + v_n(t) \quad (2.4)$$

Незмінні в часі частини, які відображають степінь відбиття різними частинами шкіри - можна поєднати в одне (2.5):

$$u_c c_0 = u_s s_0 + u_d d_0 \quad (2.5)$$

де:

u_c визначає одиничний вектор кольору, а c_0 - силу відбиття.

Також, помітимо, що можна визначити 2.6:

$$I(t) = I_0(1 + i(t)) \quad (2.6)$$

Розбиваючи $I(t)$ на незмінну частину I_0 і залежну від часу - $I_0 i(t)$. (нагадаю, що зміни інтенсивності світла, які спостерігається камерою - пропорційні до рівня інтенсивності).

Таким чином, можна переписати (2. 4) в наступному вигляді (2.7):

$$C_k(t) = I_0(1 + i(t)) (u_c c_0 + u_s s(t) + u_p p(t)) + v_n(t) \quad (2.7)$$

Тут $i(t), s(t), p(t)$ - сигнали з нульовим середнім значенням.

Насправді, дзеркальний сигнал - є найбільшою з компонент і може з легкістю сильно перевищувати всі інші. Припускається, що є засоби для відкидання областей на шкірі, в яких дзеркальний сигнал є найбільшим. Отже, ми розглядаємо лише ті пікселі, в яких u_d лише в малому степені визначається диффузним відбиттям.

Звідси - загальне завдання полягає в тому, щоб видобути з $C_k(t)$ значення $p(t)$.

Така модель використовується в якості основи для різних методів підрахунку пульсу. Більшість таких методів використовує середнє значення кольору шкіри, створюючи тимчасові RGB-сигнали для визначення пульсу. Такий крок, в першу чергу, зменшує значення похибки, яке накладає на підрахунки камера. Потрібно мати на увазі, що для використання такого кроку потрібно використовувати достатню кількість пікселів. Тоді можна переписати (2.7) у вигляді (2.8):

$$C(t) = I_0(1 + i(t)) (u_c c_0 + u_s s(t) + u_p p(t)) \quad (2.8)$$

де:

$C(t)$ - визначає середнє значення кольору в момент часу t . Також, при застосуванні цього кроку вважається, що колір пікселя не залежить від його положення на тілі.

Можна спростити (2.8) наступним чином (2.9):

$$\begin{aligned} C(t) &= u_c I_0 c_0 + u_s I_0 s(t) + u_p I_0 p(t) + u_c I_0 c_0 i(t) + \\ &+ u_s I_0 s(t) i(t) + u_p I_0 p(t) i(t) \approx \\ &\approx u_c I_0 c_0 + u_c I_0 c_0 i(t) + u_s I_0 s(t) + u_p I_0 p(t) \end{aligned} \quad (2.9)$$

Використовуючи той факт, що змінні частини набагато менші за постійні, а отже значення складових, в яких використовується їх добуток (наприклад, $s(t)i(t)$) - не будуть надавати результату помітних змін.

З (2. 9) легко бачити, що $C(t)$ - це просто лінійна комбінація трьох джерел-сигналів - $i(t), s(t), p(t)$. Це означає, що ми можемо розділити ці сигнали використовуючи лінійне відображення. Таким чином, завдання визначення пульсу з відео перетворюється на завдання визначення системи проєкції для розбиття $C(t)$.

2.3. Порівняння алгоритмів отримання rPPG сигналу з даних відеоінформації

Сімейство методів, заснованих на сліпому розділенні сигналів (Blind Source Separation, BSS) включає в себе, наприклад такі алгоритми, як ICA та PCA і може бути, в загальному, описані наступним чином (2.10):

$$Y(t) = WC(t) \quad (2.10)$$

де:

$Y(t)$ - факторизовані джерела-сигнали, що складаються з пульсу і шуму;

W - матриця, яка може бути оцінена різними методами (метод головних компонент, аналіз незалежних компонент, тощо).

Істотним для цих методів є припущення, що $i(t), s(t), p(t)$ є між собою незалежними і некорельованими.

Методи, засновані на сліпому розділенні сигналів - зазвичай вважають, що найбільш періодичний з результуючих сигналів і є пульсом. Нажаль, виходячи з такого припущення, вони нездатні впоратись з ситуаціями, коли рухи також є ритмічними (наприклад, в залі, під час крокування, таців, тощо).

Також, додаткові обмеження накладаються вибором методу для отримання матриці W [11-18]

Однак найбільшою проблемою BSS-методів є те, що вони розглядають дану задачу як самостійну задачу про розділення сигналів,

ігноруючи властивості шкіри в контексті відбивання світла. Наприклад, ігноруючи факт, що дані визначають колір.

Методи, засновані на моделях (model-based methods).

На відміну від BSS-методів, які ігнорують той факт, що вхідні сигнали задають деякі кольори, методи, засновані на моделях використовують інформацію про кольори різних компонент для контролю факторизації. Однак, потрібно зазначити, що обидва підходи мають спільний крок - знищення залежності $C(t)$ від середнього кольору відбитого від шкіри світла, включаючи колір джерела світла і внутрішній колір шкіри. Для цього проводять тимчасову нормалізацію, розділяючи вхідний RGB сигнали на їх тимчасові середні. Зазначимо, що таке ділення не впливає на змінну в часі компоненту сигналу.

Як було зазначено вище, на результуючий вектор-колір дуже сильно впливає значення постійного компонента (2.11).

$$\underline{C}(t) \approx u_c I_0 c_0 \quad (2.11)$$

звідси визначаємо нормалізуючу, діагональну матрицю N (2.12), таку, що:

$$N \underline{C}(t) \approx N u_c I_0 c_0 = \underline{1} \quad (\underline{1} = (1,1,1)^T) \quad (2.12)$$

І N використовується для тимчасової нормалізації $C(t)$ (2.13):

$$\begin{aligned} C_n(t) = NC(t) &= Nu_c I_0 c_0 + Nu_c I_0 c_0 i(t) + Nu_s I_0 s(t) + Nu_p I_0 p(t) = \\ &= \underline{1}(1 + i(t)) + Nu_s I_0 s(t) + Nu_p I_0 p(t) \end{aligned} \quad (2.13)$$

Причому зазначимо, що $\underline{1}(1 + i(t))$ - включає в себе компоненту інтенсивності світла.

$Nu_s I_0 s(t)$ - компоненту дзеркального відображення.

$Nu_p I_0 p(t)$ - компоненту пульса.

В [8] також зазначено важливі властивості даних компонент:

Інтенсивність - найбільша з компонент, як правило, рівномірно розподілена по всім трьом RGB-каналам.

Компонента дзеркального відображення - в умовах, коли джерело світла генерує біле світло (2.14):

$$Nu_s I_0 = N \underline{1} I_0 \quad (2.14)$$

Тобто масштабується оберненим значенням кольору шкіри. В інших умовах, ця компонента залежить від спектру світла і чутливості камери (помітимо, що N залежить від тих самих значень, але окрім них ще і від властивостей шкіри).

Компонента пульсу - майже не залежить від кольору шкіри. Було також показано, що найбільша амплітуда коливань належить зеленій компоненті кольору.

Проаналізуємо окремо RBV і CHROM, щоб роздивитися, як саме вони використовують властивості фізіології і відбиття людської шкіри і для отримання більш точних результатів.

Дані методи використовують сигнал $C_n(t)$ з видаленим зміщенням (2.15):

$$\underline{C}_n(t) = 1 - C_n(t) = \underline{1} i(t) + Nu_s I_0 s(t) + Nu_p I_0 p(t) \quad (2.15)$$

Тут $\underline{C}_n(t)$ - це значення сигналу кольору з нульовим середнім.

Даний метод напряму отримує значення пульсу з пульсуючої компоненти, обмежуючи всі кольорові варіації пульсуючим напрямком. Це досягається за допомогою відображення $\underline{C}_n(t)$ на одну вісь z , щоб створити наближення $\underline{p}(t)$, пропорційне до $p(t)$ (2.16):

$$\underline{p}(t) = \underline{C}_n^T(t) z = kp(t) \quad (2.16)$$

де:

z - вектор проєкції 3×1 , що містить комбінуючі ваги;

k - коефіцієнт пропорційності ($k \neq 0$).

Далі ми припускаємо, що $p(t)$ (і, отже, $\underline{p}(t)$) є некорельованим з іншими джерелами сигналу (2.17):

$$E\{p(t) \cdot i(t)\} = E\{p(t) \cdot s(t)\} \approx 0 \quad (2.17)$$

Тепер, розглядаючи очікуване значення $E\{\underline{C}_n(t) \cdot \underline{p}(t)\}$, маємо (2.18):

$$E\{\underline{C}_n(t) \cdot \underline{p}(t)\} = E\{\underline{C}_n(t) \underline{C}_n^T(t)\} \cdot z = k \cdot E\{\underline{C}_n(t) p(t)\} \approx k \cdot N \cdot u_p \cdot I_0 \cdot E\{p(t) \cdot p(t)\} \quad (2.18)$$

Таким чином, RBV передбачає попередньо-відомий вектор u_{pbv} (2.19-2.22), який задовольняє:

$$u_{pbv} = N \cdot u_p \cdot I_0, \quad (2.19)$$

Тоді можна записати:

$$E\{\underline{C}_n(t) \underline{C}_n^T(t)\} \cdot z = k \cdot u_{pbv} \cdot E\{p(t)^2\} \quad (2.20)$$

А вектор z може бути виведено через:

$$z = E\{\underline{C}_n(t) \underline{C}_n^T(t)\}^{-1} k \cdot u_{pbv} \cdot E\{p(t)^2\} \quad (2.21)$$

RBV використовує часову матрицю коваріації

$$\underline{\Sigma} = \underline{E\{\underline{C}_n(t) \underline{C}_n^T(t)\}} \quad (2.22)$$

де:

$\underline{\{\cdot\}}$ - тимчасовий оператор усереднення для виведення коваріації у часовій області і приймає k такий, щоб z мав довжину 1.

Таким чином, беручи до уваги (2.21) та (2.22), отримуємо (2.23):

$$z \propto \underline{\Sigma}^{-1} \cdot u_{pbv} \quad (2.23)$$

Так як $\underline{\Sigma}$ визначається з відео, задача PBV методу полягає лише в знаходженні вектора u_{pbv} .

Перевага методу PBV очевидна: при виконанні припущення (2.19), розрахункова вісь проєкції є оптимальною для отримання пульсу.

Однак, даний метод має декілька обмежень. Так, не можна визначити z , якщо $rank(\underline{\Sigma}) < 3$. Коли ранг матриці одиничний чи близький до одиниці - значення занадто кероване шумом. Якщо ж ранг = 1, то будь-яке значення z буде вирішенням. Це типово виходить, коли $i(t) = s(t) = 0$, тобто в ідеальних умовах.

По-друге, цей метод потребує точного знання про u_{pbv} , тобто передбачаючи здатність методу залежить від того, чи дійсно (2.19), і швидко падає, коли це не так. Отже, результат алгоритму чутливий до певного параметра налаштування PBV, а PBV, у свою чергу, визначається (і, отже, обмежується) певним налаштуванням запису, залежно від спектру світла і камери.

Розвитком попереднього методу можна вважати CHROM.

На відміну від прямолінійного однокрокового рішення PBV, CHROM запроваджує гнучкість при оцінці напрямку проєкції та зменшення чутливості до попередніх знань, що використовуються для видобування імпульсів.

Даний метод на початку зменшує розмірність завдання розділення усуваючи дзеркальну компоненту. Це досягається при розгляданні лише сигналів кольору, які ми опишемо як проєкцію $\underline{C}_n(t)$ на площину, ортогональну напрямку зміни кольору дзеркальної компоненти.

Для правильної роботи в умовах змінного кольору світла, метод передбачає стандартизований колір тону шкіри, який включає автоматичний баланс білого на зображенні. Таким чином, ми впроваджуємо колір тону шкіри як матрицю відображення (2.24):

$$M^{-1}u_{skin} = 1 \quad (2.24)$$

де:

u_{skin} - вектор 3×1 , який позначає колір шкіри під білим світлом (його можна отримати з масштабного експерименту [9]);

M - діагональна матриця відображення, яка використовується для відображення $\underline{C}_n(t)$ наступним чином (2.25, 2.26):

$$M\underline{C}_n(t) = M\underline{1} i(t) + MNu_s I_0 s(t) + MNu_p I_0 p(t) \quad (2.25)$$

Тут $Nu_s I_0$ наближено відображений в напрямку білого світла, тобто:

$$MNu_s I_0 = \kappa \underline{1}, \quad (2.26)$$

де κ - фактор пропорційності.

Наступний крок CHROM полягає в проєкції $M\underline{C}_n(t)$ на площину, ортогональну $\underline{1}$, щоб виключити залежність від дзеркальних варіацій (після корекції тону шкіри), (2.27).

$$S(t) = P_c \cdot M \cdot \underline{C}_n(t) = P_c M \underline{1} i(t) + P_c MNu_p I_0 p(t) \quad (2.27)$$

маючи на увазі, що:

$$P_c MNu_s I_0 \approx \kappa P_c \underline{1} = 0, \quad (2.28)$$

де:

P_c - матриця початкової проєкції 2×3 , що використовується CHROM та складається з осей проєкції в рядках, які визначають площину у часово нормалізованому просторі RGB.

Зазначимо, що $P_c \cdot M$ - результуюча матриця, яку використовує даний метод. Така матриця має чудову властивість - вона створює два розділених

сигнали в $S(t)$, в яких зміни, спричинені рухом і пульсом знаходяться в фазі\протифазі [3].

Ця властивість дозволяє досить просто наближено оцінити $\underline{p}(t)$, який надає наближене значення $p(t)$ (2.29). Цей процес називають альфатюнінгом.

$$\underline{p}(t) = S_1(t) - \alpha S_2(t) \quad (2.29)$$

де:

$$\alpha = \frac{\sigma(S_1)}{\sigma(S_2)},$$

σ - стандартне відхилення;

S_i - визначає i -тий прогнозований сигнал.

Коли пульсуюча компонента домінує - S_1 і S_2 знаходяться в протифазі, що сильно допомагає спростити (2.29) в (2.30):

$$\underline{p}(t) = 2 S_1(t) \propto p(t). \quad (2.30)$$

Коли компонента, спричинена рухом, рівняння виключає фазові компоненти руху, щоб оцінити $p(t)$. Лише у випадках, коли сила змін, спричинених рухом і пульсом відносно збалансована, таке наближення не є оптимальним.

Вкажемо також, що в [9] були отримані оптимальні значення для матриці $P_c = \begin{pmatrix} 1 & -1 & 0 \\ 0.5 & 0.5 & -1 \end{pmatrix}$, яка відображається на вектор стандартизованого кольору шкіри $u_{skin} = [0.77, 0.51, 0.38]^T$, в результаті отримуємо (2.31):

$$P_c M = \begin{pmatrix} 3 & -2 & 0 \\ 1.5 & 1 & -1.5 \end{pmatrix} \quad (2.31)$$

Таким чином, перевага методу CHROM полягає в тому, що він є відносно стійким до не білого освітлення. Однак він вимагає провести ціле дослідження, щоб компенсувати дзеркальне відбиття світла $Nu_s I_0$ за

допомогою вектора кольору шкіри u_{skin} . Інакше він буде демонструвати залишки дзеркального відбиття світла на розкладених сигналах.

Метод, який базується на даних (2SR, Spatial Subspace Rotation) був розроблений в роботі [16].

Як видно, усі представлені раніше методи використовують прийом усереднення кольору шкіри, отримуючи $(\underline{R}, \underline{G}, \underline{B})$. Вони не враховують просторовий розподіл пікселів шкіри в RGB просторі. 2SR пропонує взяти його, враховуючи що цей розподіл може бути просто виведений з просторового співвідношення RGB (2.32):

$$C = \frac{V^T V}{N} \quad (2.32)$$

де:

N - кількість пікселів шкіри;

V - матриця пікселів шкіри на кадрі розміру $N \times 3$. Наприклад, кожний рядок задає один піксель, а кожний стовпчик - один з каналів RGB;

C - 3×3 симетрична, кореляційна матриця невід'ємних значень.

Звернімо увагу, що C відрізняється від матриці коваріації, в якій віднімається середнє значення V . Причина в тому, що $(\underline{R}, \underline{G}, \underline{B})$ має важливу інформацію про пульс при поєднанні декількох спостережень в часі, що використовується іншими методами для визначення ЧСС. При видаленні цього значення з V , пульсуюча компонента також буде знищена з C , що робить неможливим визначення пульсу. При розкладі C ми можемо отримати підпростір пікселів шкіри (2.33):

$$C \cdot U = \lambda \cdot U \Rightarrow \det(C - \lambda \cdot I) = 0 \quad (2.33)$$

\det - визначник матриці;

U, λ - власні вектор і значення матриці, відповідно.

Зауважимо, що C - матриця 3×3 з рангом 3, так як в реальних умовах дуже сумнівно отримати однакові значення для каналів RGB.

Звідси, можна розкрити C (2.34):

$$C = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \lambda_3 u_3 u_3^T, \quad (2.34)$$

де:

u_i - i -тий вектор стовпець U ;

u_i^t - i -тий діагональний елемент матриці U_t .

В RGB-просторі ми визначимо U , як нову систему координат кольорів пікселів. Тут перший власний вектор u_1 - вектор шкіри домінує в кластері пікселів шкіри, що є найменшою квадратною оцінкою, яка є надійною стійкою до просторових викидів. u_2, u_3 - наступні напрямки, ортогональні до u_1 (рисунок 2.4).

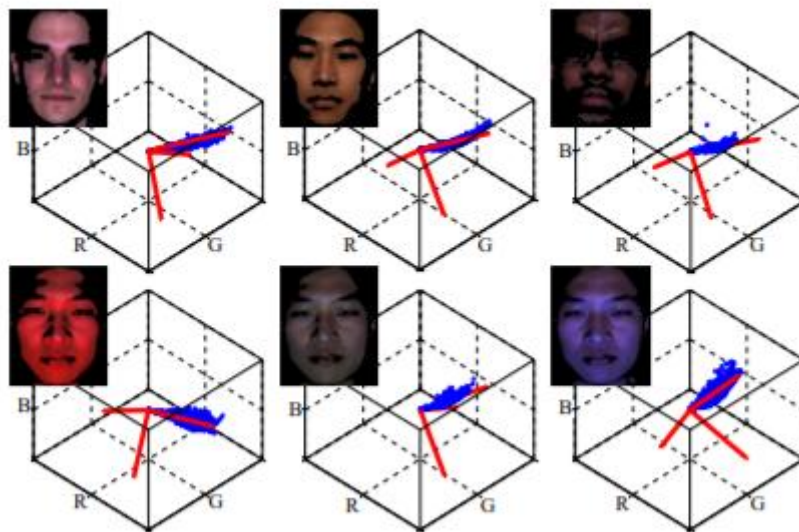


Рисунок 2.4- Приклад підпросторів пікселів шкіри, сформульованих за різних обставин.

Підпростір (наприклад, власний вектор, позначений червоним) може адаптуватися до різних розподілів в просторі RGB пікселів шкіри (наприклад, розсіювання, позначені синім кольором) у різних тонах шкіри та умовах освітленості. Основний напрямок підпростір знаходиться вздовж основної дисперсії скупчення пікселів шкіри.

Звернімо увагу на два важливих моменти на цьому рисунку:

1. Так як значення пікселів завжди додатні - вектор не може вказувати у від'ємному напрямку RGB простору.
2. Шкіра людини відносно однорідна за кольором (у межах певного діапазону відтінків), але змінюється за інтенсивністю через тіні та дзеркальні відбиття.

Виходячи з різних рівнів інтенсивності, розподіл пікселів шкіри у просторі RGB може бути або еліпсом, що виникає з RGB джерела, або компактна сфера з однорідними відбиттями шкіри. В обох випадках u_1 вказуватиме на кластер пікселів шкіри з RGB походження. Так як шкіра людини не барвиста, розподіл пікселів шкіри ніколи не може бути хмарою розподіленою повсюди в просторі RGB.

Надалі в цьому методі вектор усереднених значень RGB каналів пікселів шкіри замінюються вектором U .

Зазначимо, що даний метод (на відміну від попередніх) потребує добре визначеної маски шкіри, інакше він буде змішувати при аналізі кольори шкіри і фону, що призведе до істотного зниження точності.

В часовому просторі пульсуюча кров викликає коливання в каналах RGB і, таким чином, змінює підпростір пікселів шкіри. Так як цей підпростір будувався без нормалізації, ми не можемо прямо дивитися на зсув середнього значення RGB в часі, щоб вимірювати пульс, так як це значення є пропорційним до яскравості освітлення. Даний метод модулює часову залежність між двома підпросторами як миттєвий поворот і масштабування:

по-перше, кут між власними векторами визначається різними вкладками PPG компонент в RGB канали. По-друге, зміни в власних значеннях мають відношення до коливань кольору шкіру.

Так як підпростір U складається з власних векторів з одиничною нормою, її часові зміни можуть створювати лише повороти. Таким чином, введемо поняття часового кроку довжини l , щоб аналізувати обертання

підпростору. Роздивляючись підпростір в першому кадрі часового кроку U_τ як вихідне значення, кут повороту між $U_{t,t < l}$ і U_τ буде (2.35):

$$R = U_\tau^T U_t = (u_1^{tT} \ u_2^{tT} \ u_3^{tT}) \cdot (u_1^{\tau T} \ u_2^{\tau T} \ u_3^{\tau T}), \quad (2.35)$$

де:

R - матриця повороту;

u_i^t - i -тий стовпець U_t .

Так як L2 норма власних векторів нормалізована до 1, значення в R відповідають косинусам кутів між власними векторами (2.36):

$$\cos(\theta_{ij}) = \frac{u_i^{tT} \cdot u_j^{\tau T}}{\|u_i^t\| \cdot \|u_j^{\tau T}\|} \quad (2.36)$$

де:

i, j - визначають індекс стовпця,

θ_{ij} - кут обертання між u_i^t і $u_j^{\tau T}$, $\|\cdot\|$ - L2 норму,

$\|u_i^t\|$ і $\|u_j^{\tau T}\|$ дорівнюють одиниці.

Помітимо також, що власні вектори U - ортогональні, а кути θ_{ij} близькі до нуля (це зміни в часі за дуже малий час). Тоді, (2.36) може бути переписана як (2.37):

$$R = \begin{pmatrix} \cos(\theta_{11}) & \sin(\theta_{12}) & \sin(\theta_{13}) \\ -\sin(\theta_{12}) & \cos(\theta_{22}) & \sin(\theta_{23}) \\ -\sin(\theta_{13}) & -\sin(\theta_{23}) & \cos(\theta_{33}) \end{pmatrix} \quad (2.37)$$

Так як зміни кольору, спричинені пульсом мінорні - кути θ_{ij} - близькі до нуля.

Таким чином, $\cos(\theta_{ij})$ - діагональні елементи - варіюються коло 1. Це не може точно відображати напрям обертання (завжди додатне). Таким чином, відносний вклад PPG компонентів в RGB канали стає дещо розпливчатим через беззнакове обертання. До того ж, косинусові

компоненти є менш чутливими до пульсу. На відміну від них, не діагональні синусові компоненти не мають проблем з напрямком і демонструють більш сильну змінність, змінюючись у значеннях, близьких до нуля.

На рисунку 2.5 показаний приклад значень синусів та косинусів в часі. Стандартне відхилення $\cos(\theta_{11})$, $\sin(\theta_{12})$, $\sin(\theta_{13})$, складають, відповідно:

$4.3 \cdot 10^{-6}$, $1.6 \cdot 10^{-3}$, $1.9 \cdot 10^{-3}$. Тому лише не діагональні елементи матриці будуть використовуватися.

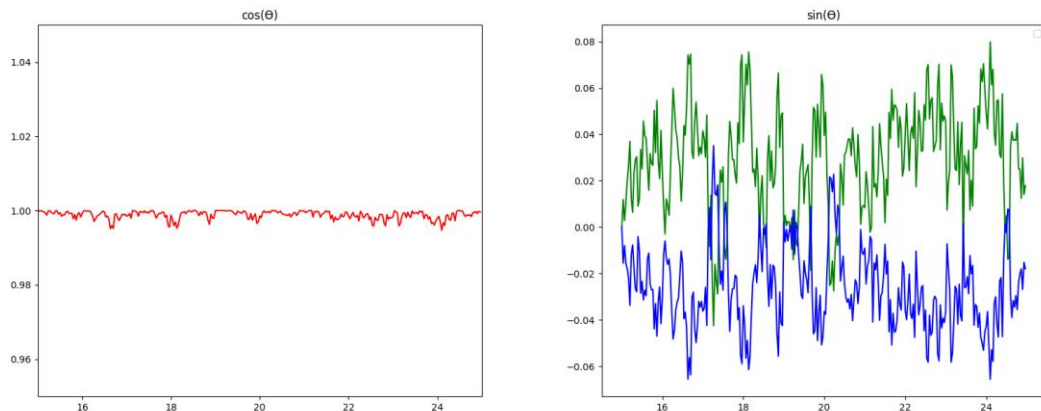


Рисунок 2.5. Значення синусів та косинусів в матриці обертання. Використовується відео з 75 кадрами. Записаний об'єкт має світлу шкіру і залишається непорушним протягом запису. Використовується запис першого рядка матриці

Пульсуючі зміни кольору шкіри, спричинені рухом крові, які, по факту, є обертанням вектора кольору шкіри u_1 в просторі RGB. Тому, нас цікавлять лише часові зміни вектора u_1 . Однак, як було показано раніше, обертання між u_1^t і u_1^t (тобто, $\cos(\theta_{11})$) не можна ефективно відслідкувати. Тому, для вимірювання обертання будемо використовувати кути між даним вектором u_1^t і ортогональною площиною $(u_2^t \ u_3^t)$ (2.38):

$$R' = (u_1^{tT}) \cdot (u_2^{tT} \ u_3^{tT}) = (u_1^{tT} \cdot u_2^{tT} \ u_1^{tT} \cdot u_3^{tT}), \quad (2.38)$$

Що, по факту, співпадає з значеннями $\sin(\theta_{12})$ і $\sin(\theta_{13})$. Щоб зрозуміти ефективність цього кроку, представимо більш розгорнуті

результати на рисунку 2.3.4. У випадку нерухомості, коливання, спричинені пульсом - є єдиним джерелом сигналу, яке викликає періодичні обертання підпростору кольору шкіри, що створює два періодичних сигнали, які знаходяться у фазі. У випадку з рухом, один з напрямків в R' буде якимось чином спотворений через рухи голови, що вже має відношення до спектру світла. Інший напрямок - ортогональний і все ще, в більшості, залежить від пульса.

В додаток до обертання підпростору, власні значення, відповідні дисперсії енергії власних векторів - також піддаються змінам, спричиненим пульсом, що також може бути використано.

Так як λ_i , витягнуті з C задають потужну дисперсію, ми квадратуємо їх і таким чином знайдемо їх зміни масштабу (2.39):

$$S = \sqrt{\lambda_1^t} \cdot \text{diag}((\lambda_2^T \ 0 \ 0 \ \lambda_3^T)^{-1}) = \left(\sqrt{\lambda_1^t / \lambda_2^T} \ \sqrt{\lambda_1^t / \lambda_3^T} \right), \quad (2.39)$$

де:

diag - визначає лише діагональні значення матриці.

Сигнали, які спродуковані цією матрицею - фактично, відповідають коливанням кольору вимірюваної шкіри. Однак, якщо відбитий шкірою сигнал також включає в себе зміни інтенсивності освітлення (в тому числі спричинені рухами), S - також може бути під впливом. Наприклад, сигнали отримані з об'єкта, що обертається - перебувають також під впливом поворотів голови і змінюються в більшому діапазоні (± 10), на відміну від аналогічних сигналів від нерухомого об'єкту (± 4). Тому, ми не можемо використовувати лише власні значення для вимірювання пульсу. Так як (2.3.35) було складено з використанням вихідного зображення, ми можемо обмежити її зміни лише напрямками, які співпадають з напрямками обертання підпростору кольору шкіри, комбінуючи вирази (2.38) і (2.39) в (2.40):

$$SR = S^T \times R' = \left(\sqrt{\frac{\lambda_1^t}{\lambda_2^T}} \sqrt{\frac{\lambda_1^t}{\lambda_3^T}} \right) \overset{\text{scaling}}{\times} (u_1^{tT} \cdot u_2^{tT} u_1^{tT} \cdot u_3^{tT}) \overset{\text{rotation}}{\times}, \quad (2.40)$$

де;

\times - означає операцію поелементного множення.

Інтуїтивне значення формули (2.40) - вектор доміантного кольору шкіри u_1^t - спроектований і масштабований на ортогональну площину $(u_2^T u_3^T)$.

Такий крок має кілька переваг:

По-перше, він збільшує амплітуду обертання підпростору, що означає підвищення коливаючої компоненти.

По-друге, він зменшує силу коливань в S , які не впливають на обертання підпростору кольору шкіри, наприклад - змін рівня освітленості шкіри.

Також, так як значення масштабуючої складової (відповідно до λ_1) набагато більші, ніж відповідні в складовій обертання, масштабуюча складова завжди додаткова і може керуватися обертанням до додатних або від'ємних напрямках.

Зважаючи на метод ковзного вікна [7], вихідний підпростір U_τ - константний в кожному з часових кроків. Щоб отримати часово-стійкий SR на декількох кроках, проаналізуємо SR в тому ж просторі і таким чином спроектуємо його у вихідний RGB простір (2.41):

$$SR' = SR \cdot (u_2^{T T} u_3^{T T}) = \left(\sqrt{\frac{\lambda_1^t}{\lambda_2^T}} \sqrt{\frac{\lambda_1^t}{\lambda_3^T}} \right) \odot (u_1^{t T} \cdot u_2^{t T} u_1^{t T} \cdot u_3^{t T}) \cdot (u_2^{T T} u_3^{T T}) = \sqrt{\frac{\lambda_1^t}{\lambda_2^T}} \cdot u_1^{t T} \cdot u_2^{t T} \cdot u_2^{T T} + \sqrt{\frac{\lambda_1^t}{\lambda_3^T}} \cdot u_1^{t T} \cdot u_3^{t T} \cdot u_3^{T T}. \quad (2.41)$$

Проблема довільного знаку власних векторів була вирішена після зворотної проєкції. На одному кроці декілька SR' , між вихідним кадром і

наступними оцінені и поєднані в \overrightarrow{SR} . Як видно на наступному графіку, перші два записи в \overrightarrow{SR} - знаходяться в антифазі. Так само, як і в методі CHMOD, ми підсилюємо пульсуючу компоненту, поєднуючи сигнали в антифазі наступним чином (2.42):

$$\vec{p} = \overrightarrow{SR}_1' - \frac{\sigma(\overrightarrow{SR}_1')}{\sigma(\overrightarrow{SR}_2')} \cdot \overrightarrow{SR}_2', \quad (2.41)$$

де :

\overrightarrow{SR}_1' - і-та компонента SR' ;

$\sigma(\cdot)$ - визначає оператор середньоквадратичного відхилення.

Таким чином, довготривалий пульс-сигнал \vec{P} може бути оцінений завдяки послідовному накладанню-додаванню кадрів (2.42):

$$\vec{P}^{t-l} = \vec{P}^{t-l} + (\vec{p} - \mu(\vec{p})), \quad (2.42)$$

де:

$\mu(\cdot)$ -оператор усереднення;

\vec{P} - одновимірний сигнал довжини K (кількості кадрів в відео), який ініціалізований нульовими значеннями та змінюється \vec{p} .

Сигнал пульсу може бути отримано навіть у випадку, коли вихідний RGB сигнал сильно порушений через рухи (рис. 2.6).

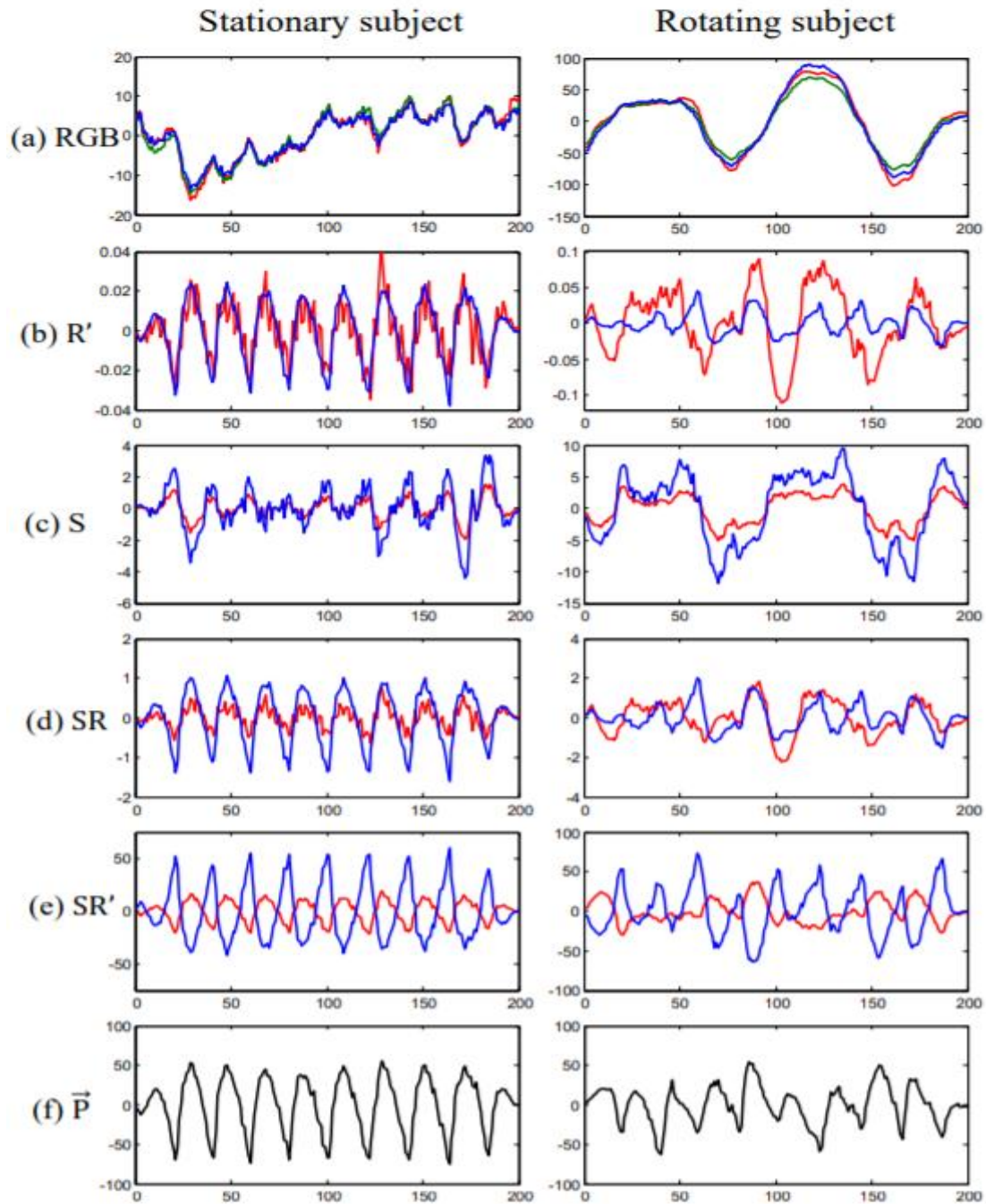


Рисунок 2.6 - Результати, отримані на кожному кроці алгоритму 2SR.

В лівій колонці рис. 2.6 суб'єкт зйомки був нерухомий, в правій - обертав головою, щоб показати простий і складний сценарії. Зазначимо, що в випадках, коли крок продукує два сигнали, ми позначаємо перший синім кольором, а другий - червоним. [20]

Порівняємо вищеописані алгоритми:

Сила **BSS**-методів полягає в тому, що вони дозволяють розглядати дану задачу чисто з математичної точки зору, не впроваджуючи жодних додаткових сутностей. В цьому ж криється їх проблема: вони розглядають дану задачу як самостійну задачу про розділення сигналів, ігноруючи властивості сигналу, шкіри, кольорів, тощо, що призводить до зниження точності результатів і неможливості використання цих методів у специфічних для даної задачі випадках (рухи, зміни освітлення, тощо).

Найбільша перевага методу **PBV** в тому, що при виконанні деяких припущень про відео, він надає оптимальні для отримання пульсу результати. Однак, можна виділити декілька обмежень. По-перше, даний алгоритм погано працює в умовах, близьких до ідеальних, по-друге - його результати сильно чутливі до налаштувань камери і спектру світла.

Сила методу **CHROM** полягає в тому, що він є відносно стійким до не білого освітлення. Однак він вимагає провести ціле дослідження, щоб компенсувати дзеркальне відбиття світла за допомогою вектора кольору шкіри.

2SR метод показує точні результати, стійкий до змін освітлення і до обертання голови. Також, він використовує лише дані суб'єкта, тобто йому не потрібні підраховані наперед середні значення векторів кольору і тону шкіри. В той самий час, 2SR потребує великої кількості пікселів шкіри, не використовує фізіологічних властивостей шкіри і, іноді, має проблеми з ефективністю, що робить його не найвдалішим рішенням якщо потрібно виконувати rPPG в реальному часі.

РОЗДІЛ 3. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

3.1. Розробка алгоритму програми

Для покращення було обрано алгоритм сімейства BSS. Загальна ідея покращення полягає в тому, щоб використати ідеї алгоритму SR2 як попередню обробку даних перед використанням ICA (або PCA).

Тобто, замість використання в якості сигналу середні значення каналів R, G та B - будемо для початку вираховувати кути повороту між площинами, які будуються з власних векторів пікселів в просторі, тобто (2.3.34).

Завдяки цьому кроку, в BSS алгоритмі використовується інформація про походження сигналу, що позбавляє нас від однієї з найбільших його проблем.

Для реалізації алгоритму було обрано мову програмування python, так як більшість потрібних для реалізації компонент в ній уже представлені і добре протестовані в існуючих бібліотеках.

Розроблений алгоритм покращення BSS методу представлено на рисунку 3.1.



Рисунок 3.1 Алгоритм покращеного BSS

Тут, в елементі з назвою «Знайти області інтересу» закладено дії, описані на рисунку 3.2.



Рисунок 3.2 Алгоритм пошуку областей інтересу

3.2. Опис програмного застосунку

Для пошуку обличчя і точок інтересу на обличчі був використаний доступний для використання предиктор «81 Facial Landmarks Shape Predictor». Він знаходить на обличчі 81 проіндексовану ключову точку. Тобто, кожна точка на обличчі має свій номер незалежно від того, чи були знайдені інші точки.

Приклад роботи алгоритму представлений на рисунку 3.3:

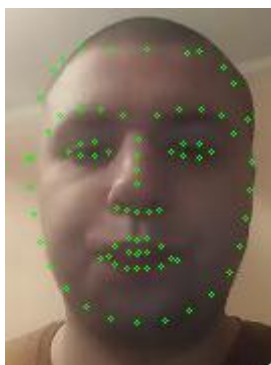


Рисунок 3.3 Приклад роботи алгоритму для пошуку ключових точок на обличчі

Як видно з алгоритму на рисунку 3.2, наступним кроком є видалення цілих зон, які не відповідають нашим потребам. Таким чином, мають залишитись лише лоб та щоки.

Для цього, евристичним шляхом були знайдені оптимальні послідовності точок, між якими знаходиться та частина обличчя, яка може містити очі, брови, бороду, тощо. Результат цього етапу алгоритму представлена на рисунку 3.4.



Рисунок 3.4 Зображення, після видалення зон, на яких можуть знаходитись непотрібні для подальшого аналізу елементи

Таким чином, ми позбуваємось цілих частин обличчя через те, що на них можливо знайти бороду або інші елементи, які вже були перераховані. В реальних умовах потрібно також мати на увазі, що на обличчі можуть знаходитись також бліки від світла, або занадто різкі тіні, як це видно на наступному рисунку 3.5:



Рисунок 3.6 Зображення, на якому присутні бліки.

Найчастіше подібні елементи знаходяться на опуклих частинах тіла, таких як надбрівні дуги, опуклості на лобі, на носі, щоках. Вони зумовлені занадто інтенсивним світлом. Такі частини відео не мають жодної корисної інформації для подальшого аналізу.

Їх можна прибрати, використовуючи фільтр. Його головна ідея полягає в тому, що в залежності від кольору шкіри, можна знайти деякий діапазон значень (який при представленні зображення в різних палітрах може відрізнитися), в котрому не надто освітлена шкіра буде знаходитись. Тож відфільтровуючи частини зображення, які занадто темні або світлі, отримуємо те, що зображено на рисунку 3.7.



Рисунок 3.7 Зображення, на якому вирізані

Після цього доречно виконати стиснення кадру. Для цього ми видаляємо всі стрічки і всі стовпчики зображення, на яких знаходяться лише чорні пікселі.



Рисунок 3.8 Стиснене зображення. В кожному стовпчику і в кожній стрічці знаходиться хоча б один кольоровий піксель.

В подальшому все одно аналізуються лише кольорові пікселі. На перший погляд може здатись, що в цьому кроці немає сенсу, адже для перевірки кожних стрічки і стовпчика ми все одно оглядаємо кожний піксель, однак це не так. Це пов'язано з тим, як саме реалізовані в мові програмування python деякі частини роботи з масивами. Вони включають в себе низько-рівневі оптимізації і, як правило, подібні алгоритми в python реалізовані на C/C++, що означає суттєве підвищення швидкодії.

Для подальшої роботи отримані пікселі використовуються як множина, їх розташування більш не має жодного значення.

Як видно з алгоритму, зображеного на 3.1, наступним кроком є знаходження матриці C , описаній в виразі (2.32):

$$C = \frac{V^T V}{N}$$

, яка, в свою чергу, використовується для знаходження матриці власних векторів U з виразу (2.33)

$$C \cdot U = \Lambda \cdot U$$

А з її допомогою ми можемо використати вираз (2.38) для знаходження R' , який в подальшому і буде аналізуватися за допомогою ІСА.

$$R' = (u_1^{tT} \cdot u_2^{tT} \ u_1^{tT} \cdot u_3^{tT})$$

Нагадаємо, що в матриці R' зберігаються лише синуси кутів, так як зміни їх значень в часі значно перевищують відповідні значення косинусів. Це продемонстровано на рисунку 2.5.

Зауважимо, що якщо в даному алгоритму не проводити знаходження матриці R' , а використовувати замість неї нормалізовані усереднені

значення кольору пікселів напряму (як це робиться в стандартних BSS алгоритмах), то можна отримати щось на кшталт зображеного на рисунку 3.9:

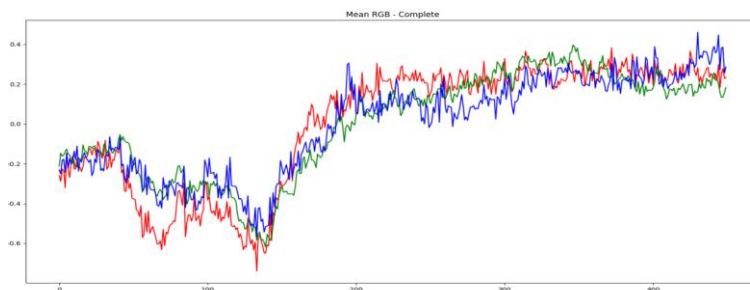


Рисунок 3.9 Усереднені значення каналів RGB на 15-секундному проміжку. Від кожного значення віднято середнє.

Розділивши за допомогою одного з BSS методів отриманий сигнал R' , отримаємо результат, зображений на рисунку 3.10.

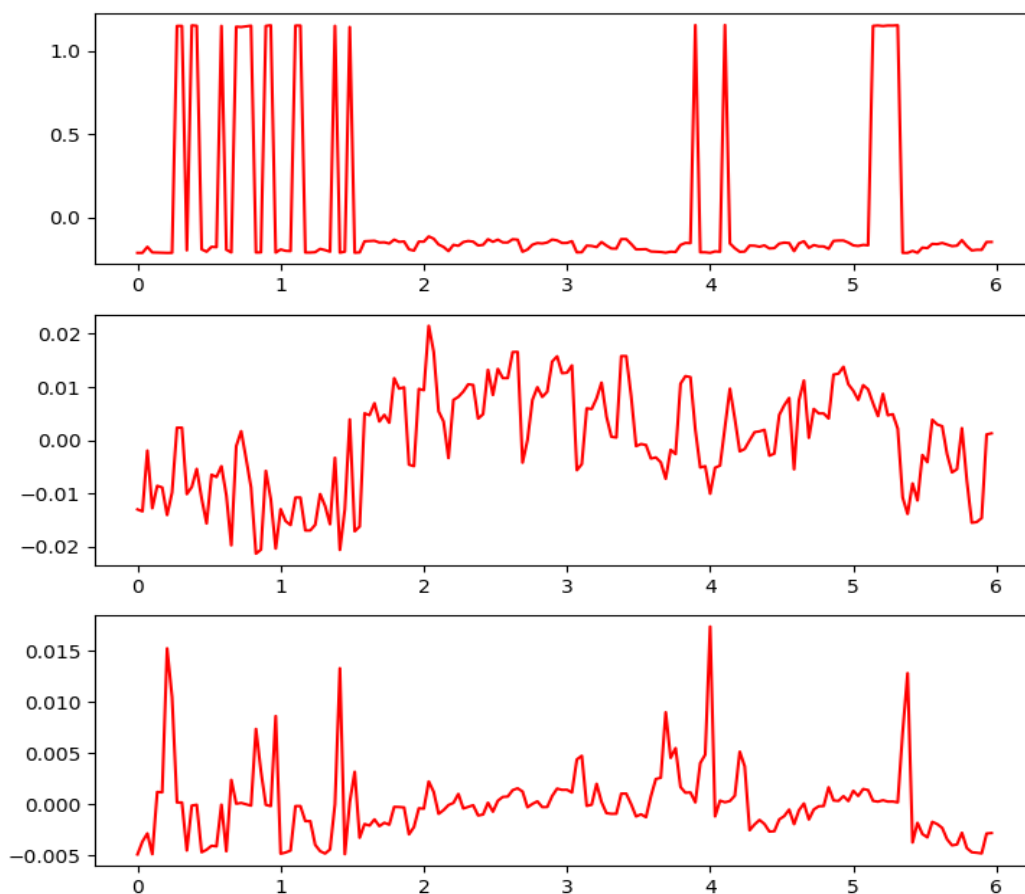


Рисунок 3.10 Результуючі компоненти, отримані за допомогою PCA. Для тесту аналізувався власноруч записаний відео файл.

Як видно з графіків на рисунку 3.10, найбільш пульсуюча компонента – друга.

Після використання частотного аналізу, отримуємо наступне:

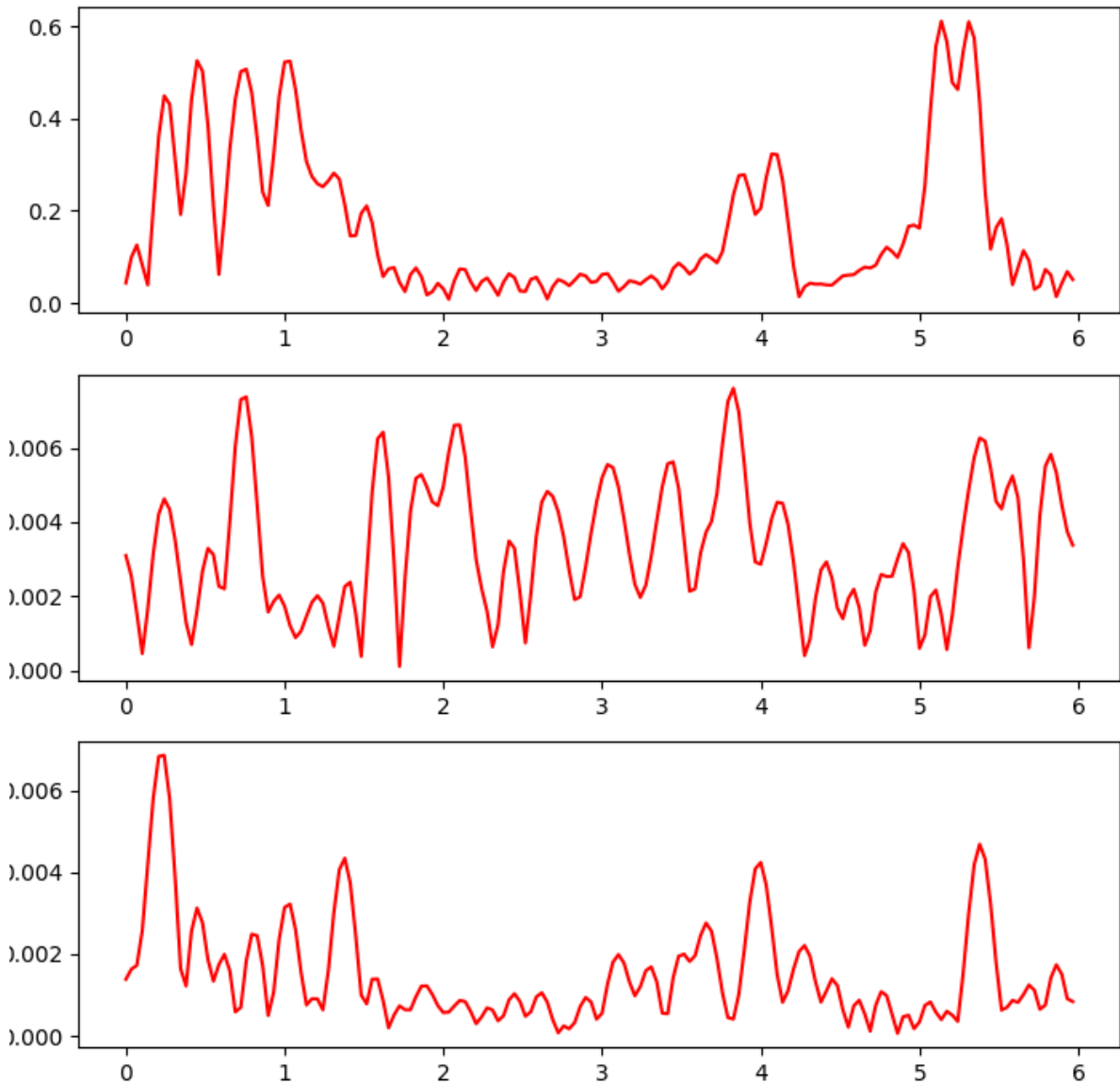


Рисунок 3.11 Результат використання частотного аналізу над відповідними графіками з рисунку 3.10

В даному випадку, мануальним підрахунком мій пульс складав близько 120-130 ударів на хвилину. Знайдемо для другої з отриманих компонент піки:

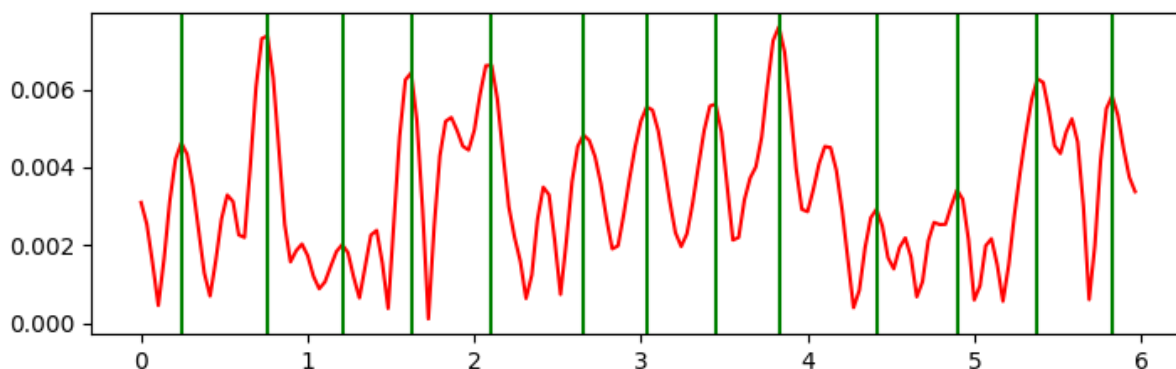


Рисунок 3.12 Піки, які було знайдено алгоритмом

Тут враховується той факт, що пульс людини не може бути нижчим за 50, або вищим за 250. Таким чином відфільтровуються додаткові, непотрібні піки, які можуть бути наслідками закладених в BSS алгоритми проблем.

Якщо підрахувати піки на рисунку 3.12, то їх вийде 13. Цей запис мав довжину в 6 секунд, тобто, з точки зору алгоритму, мій пульс відповідає 130, що досить близько до справжнього результату.

3.3. Опис даних для тестування

Для масштабної перевірки алгоритму було використано відкритий для таких цілей датасет [10]. В ньому представлені дані для тестування. Всього представлено 21 набір, які демонструють відеозаписи різних людей в умовах різного освітлення, в стані спокою та після навантажень. В кожному наборі представлено 2 відео - до та після попередньої обробки, яка полягає в вирізанні всього зайвого, що потрапляє в кадр, а також перших кількох секунд, коли людина ще не зайняла місце в кадрі. Представлені тести, в яких людина нерухомо зафіксована, а також тести, де людина обертає головою. Приклади кадрів з вказаних відео представлені на рисунках 3.13 - 3.17 .



Рисунок 3.13 Кадр з оригінального відео



Рисунок 3.14 Кадр з відео після обробки



Рисунок 3.15 Приклад інших умов освітлення



Рисунок 3.16 Приклад кадрів з відео, де людина рухається



Рисунок 3.17 Приклад кадрів з відео з людьми з різним кольором шкіри

Також, кожний тестовий набір супроводжується даними з пульсометра, які представлені як послідовність чисел, які описують час кожного удару пульсу людини.

Довжина кожного відео складає близько п'яти хвилин.

Для тестування, з кожного відео було обрано по декілька проміжків довжиною 10 секунд. Всього, таким чином, було обрано 90 проміжків, по яким і проводилося тестування алгоритму. Для кожного часового проміжку було обрано

В наступній таблиці приведені статистичні дані про обрані проміжки.

Таблиця 3.1 – описова статистика даних (за вказаний проміжок часу)

Середнє	11.4
Стандартна помилка	0.16
Медіана	11
Мода	10
Стандартне відхилення	1.49
Дисперсія	2.24

Дані розташовані в окремому каталозі та відсортовані по окремим підкаталогам.

3.4. Аналіз результатів роботи програмного застосунку

Над цими даними було запущено описаний в розділі 3.1 алгоритм і підраховані значення помилок. Приклад роботи над одним з відеофайлів:

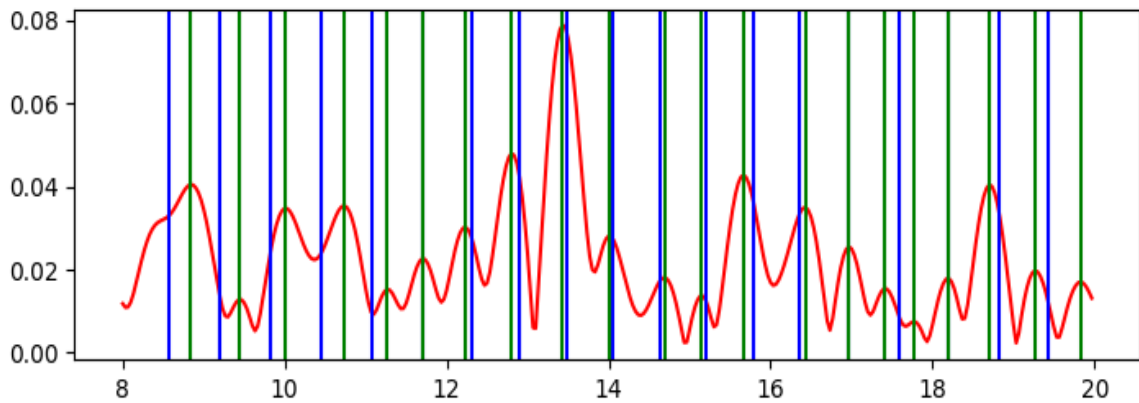


Рисунок 3.18 Графік даних після обробки оновленим ICA і FFT. Тут сині стовпці позначають момент, коли в людини був реальний удар серця, а зелені - момент, коли алгоритм з розділу 3.1 показав наявність удару серця.

Зауважимо, що в даному випадку при підрахунку похибок не враховувались реальні моменти часу, коли був здійснений удар серця, лише їх кількість, однак розташування також показує хороші результати, що видно з рисунку 3.18.

Використовуючи ідею, яка лежить в основі методу 2SR, вдалося покращити результати одного з BSS методів. Таким чином вдалося покращити PCA метод знизивши середню абсолютну помилку з 9.15 ударів на секунду до 7.09 ударів за секунду і дещо збільшити кореляцію з результатами.

При порівнянні на нормалізованих тестах були отримані наступні результати:

Таблиця 3.2 – отримані результати

Алгоритм	PCC, r	MAE
PCA	0.49	9.15
Оновлений алгоритм	0.53	6.96
CHROM	0.87	3.81

Так як оригінальний ІСА не тестувався на даному наборі даних, для порівняння було обрано результати з інших датасетів [30].

Висновки

На основі аналізу відео з використанням методів машинного навчання були розроблені елементи системи діагностики пульсу людини, які дозволяють отримати результат з точністю до семи серцевих скорочень на хвилину. Це показує, що можливо покращувати окремі алгоритми, компенсуючи їх слабкі сторони за допомогою ідей, представлених в інших алгоритмах.

За допомогою згорткових нейронних мереж і методів машинного навчання можна забезпечити вірогідний захват зображення обличчя людини.

Література

1. Active Appearance Model (AAM) // [Електронний ресурс]. // Режим доступу: <http://ice.dlut.edu.cn/lu/AAM.html> , 15 квітня 2021.
2. 81 Facial Landmarks Shape Predictor . // [Електронний ресурс]. // Режим доступу: // [Електронний ресурс]. // Режим доступу: https://github.com/codeniko/shape_predictor_81_face_landmarks, 3 травня 2021.
3. Algorithmic Principles of Remote PPG," W. Wang, A. C. den Brinker, S. Stuijk and G. de Haan. // [Електронний ресурс]. // Режим доступу: https://pure.tue.nl/ws/portalfiles/portal/31563684/TBME_00467_2016_R1_preprint.pdf.
4. Baker S/ Lucas-Kanade 20 years on: A unifying framework: Part 3 / S. Baker, R. Gross, and I. Matthews // Technical Report CMU-RI-TR-03-35, Carnegie Mellon University Robotics Institute, 2003.
5. Cootes T. Active appearance models. / T. Cootes, G. Edwards, and C. Taylor//In Proceedings of the European Conference on Computer Vision volume 2, pages 484–498, 1998.
6. Cootes T. F. Statistical models of appearance for medical image analysis and computer vision / T. F. Cootes and C. J. Taylor // In Proc. SPIE Medical Imaging 2001, volume 1, pages 236–248. SPIE, 2001.
7. De Haan G. Robust pulse rate from chrominance-based rppg / G. de Haan and V. Jeanne // Biomedical Engineering, IEEE Transactions on, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
8. Guazzi A. R. et al., “Non-contact measurement of oxygen saturation with an RGB camera,” Biomed. Opt. Exp., vol. 6, no. 9, pp. 3320–3338, Sept. 2015.
9. Gudi, A.; Bittner, M.; Lochmans, R. & Van Gemert, J. (2019) Efficient Real-Time Camera Based Estimation of Heart Rate and Its Variability. International Conference on Computer Vision (ICCV) Workshop.

10. Hoffman, W. F. C., & Lakens, D. (2021, March 23). Addressing Reproducibility Issues in Remote Photoplethysmography (rPPG) Research: An Investigation of Current Challenges and Release of a Public Algorithm Benchmarking Dataset. // [Электронный ресурс]. // Режим доступа: <https://doi.org/10.17605/OSF.IO/XJF7U>.
11. Lewandowska M. et al., “Measuring pulse rate with a webcam - a noncontact method for evaluating cardiac activity,” in Proc. Federated Conf. Comput. Sci. Inform. Syst. (FedCSIS), Szczecin, Poland, Sept. 2011, pp. 405–410.
12. Liu, C. Motion magnification / Liu, C., Torralba, A., Freeman, W. T., Durand, F., And Adelson, E. H.// ACM Trans. Graph. 24. – 2005. – P. 519–526.
13. Macwan, Richard & Benezeth, Yannick & Mansouri, Alamin. (2018). Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints. BioMedical Engineering OnLine. 17. 10.1186/s12938-018-0450-3.
14. Poh M.-Z. et al., “Advancements in noncontact, multiparameter physiological measurements using a webcam,” IEEE Trans. Biomed. Eng., vol. 58, no. 1, pp. 7–11, Jan. 2011.
15. Santemiz P. Automatic landmark detection and face recognition for side-view face images / P. Santemiz, L. J. Spreeuwers and R. N. J. Veldhuis // *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, 2013, pp. 1-4.
16. Stegmann M. B. Analysis and Segmentation of Face Images using Point Annotations and Linear Subspace Techniques. Technical report IMM-REP-2002-22, Informatics and Mathematical Modelling, Technical University of Denmark, 2002
17. Tejaswini Mishra. Early Detection Of COVID-19 Using A Smartwatch /, Meng Wang, View ORCID ProfileAhmed A. Metwally, Gireesh K Bogu, View ORCID ProfileAndrew W Brooks, Amir Bahmani, Arash Alavi, Alessandra Celli, Emily Higgs, Orit Dagan-Rosenfeld, Bethany Fay, Susan

Kirkpatrick, View ORCID ProfileRyan Kellogg, Michelle Gibson, Tao Wang, Benjamin Rolnik, Ariel B Ganz, Xiao Li, View ORCID ProfileMichael P Snyder [Електронний ресурс]. // Режим доступу: <https://www.medrxiv.org/content/10.1101/2020.07.06.20147512v1.full-text>.

18. Tsouri G. R. et al., “Constrained independent component analysis approach to nonobtrusive pulse rate measurements,” *J. Biomed. Opt.*, vol. 17, no. 7, p. 077011, July 2012.
19. Verkruysse, W. Remote plethysmographic imaging using ambient light / Verkruysse, W., Svaasand, L. O., And Nelson, J. S. // *Opt. Express* v16, № 26. – 2008. – P. 21434–21445.
20. Wang W. et al., “A novel algorithm for remote photoplethysmography: Spatial subspace rotation,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1974–1984, Sept. 2016.
21. Yongzhe Yan, Xavier Naturel, Thierry Chateau, Stefan Duffner, Christophe Garcia, et al.. A survey of deep facial landmark detection. RFIAP, Jun 2018, Paris, France. fihal-02892002.
22. Безпека життєдіяльності [текст] : підручник. / [О. І. Запорожець, Б. Д. Халмурадов, В. І. Применко та ін.] – К. : «Центр учбової літератури», 2013. – 448 с.
23. Гальперин С. И. Физиология человека и животных (рус.). — Издание 5-е. — М.: Высшая школа, 1977. — С. 104—105. — 653 с.
24. Голубовська О. А. Інфекційні хвороби (підручник) / Голубовська О. А., М. А. Андрейчин, А. В. Шкурба та ін. (за ред. О. А. Голубовської). — Київ: ВСВ «Медицина» (2 видання, доповнене і перероблене). — 2018. — 688 С.
25. Заневський І. Градація за віком норми частоти серцевих скорочень у стані спокою дітей молодшого шкільного віку при застосуванні проби Руфф'є / І. Заневський // *Фізична активність, здоров'я і спорт*. – 2014. – №2(16). – С. 3-14.

26. Информационно-измерительный комплекс пульс-кардиологической диагностики // Трансфер інноваційних технологій УКРІНТЕІ. // [Електронний ресурс]. // Режим доступу: <http://www.uinteі.kiev.ua/transfer/offer.php?slang=ua&offid=4794>.
27. Іонов І.А. Фізіологія сенсорних систем: методичні рекомендації (видання друге – доповнено та перероблено) / І.А. Іонов, Т.Є. Комісова. – Х. : ФОП Петров В.В., 2018. – С. 6-26.
28. Козловська А. Лихоманка неясного генезу: діагностичний алгоритм від А до Я // Український медичний часопис – 4(2) (132) – VII/VIII 2019 // [Електронний ресурс]. // Режим доступу: <https://www.umj.com.ua/article/159782/lihomanka-neyasnogo-genezu-diagnostichnij-algoritm-vid-a-do-ya/>
29. Контраст объекта съёмки / Фотокинотехника: Энциклопедия // Гл. ред. Е. А. Иофис. — М.: "Советская энциклопедия", — 1981. — 447с.
30. Мельник І.В. Фотодіагностичні медичні технології та пристрої / І.В. Мельник // Прогресивні технології та прилади, 2011. – С.120-127.
31. Методи обстеження пульсу та артеріального тиску в терапевтичній практиці // Кафедра пропедевтики внутрішньої медицини з доглядом за хворими, загальної практики (сімейної медицини). // [Електронний ресурс]. // Режим доступу: <https://prop-intmed.umsa.edu.ua/storage/common/docs/IJF1s4P4Ad0pBLtGd0duhR3j3Eу1Gbk3uo02RaEu.pdf>.
32. Нікітчук Т. М. Система експрес-діагностики стану студентів та моніторингу стану повітря в навчальних приміщеннях у період епідемії коронавірусу COVID-19 / Нікітчук Т. М., Коренівська О. Л., Вакалюк Т. А., Морозов А. В., Морозов Д. С., Фриз С. П // Матеріали конференції. – Державний університет «Житомирська політехніка», 2021. – С.109. // [Електронний ресурс]. // Режим доступу: <https://conf.ztu.edu.ua/wp-content/uploads/2021/01/109.pdf>.

33. Остафійчук Д.І. Клінічні методи дослідження серцево-судинної діяльності (огляд літератури) / Д.І. Остафійчук, А.О. Караняга, І.С. Полікарпова // Клінічна та експериментальна патологія 2020. – Т.19, №2(72). – С.85-90. DOI:10.24061/1727-4338.
34. Павлов С.В. Біомедичні оптико-електронні системи і апарати / Павлов С.В., Кожем'яко В.П., Петрук В.Г., Колісник П.Ф., Марков С.М. // Неінвазивні методи діагностики серцево-судинної системи. – ВДТУ, Вінниця. – 2003. Ч.1. – 142 с.
35. Рогаткин Д.А. Перспективы развития неинвазивной спектрофотометрической диагностики в медицине / Рогаткин Д.А., Лапаева Л.Г. // Медицинская техника, 2003. – №4. – с.31-36.
36. Сторчун Є. Системи пульсової діагностики / Є. Сторчун // Lviv Polytechnic National University Institutional Repository <http://ena.lp.edu>. // [Електронний ресурс]. // Режим доступу до журн. http://ena.lp.edu.ua/bitstream/ntb/39116/1/41_245-248.pdf.

Додатки

1. Лістинг програми

pos_face_seg.py

```
# © Пашенко Р.Б. КНУ 2021
# Програмний код для отримання пульсу з відео

"""
Pulse extraction using POS algorithm (%(version)s)
"""

from math import sqrt
import matplotlib
from matplotlib import pyplot as plt
from matplotlib import path as mplPath
import subprocess
import threading

import os
from scipy.signal import find_peaks
from scipy.signal import peak_prominences
import sys
sys.path.insert(0, './SkinDetector')
import pkg_resources

import numpy as np
import cv2
import dlib

from imutils.video import VideoStream
from imutils import face_utils
import imutils

import argparse
import skin_detector

from image_processing.find_points import get_points;

from shapely.geometry import Point
from shapely.geometry.polygon import Polygon

import copy

from scipy import signal
from scipy.fft import fft, ifft

from sklearn.decomposition import FastICA, PCA
from numpy import linalg as LA

max_frame = 400
max_pulse = 250

def get_length(filename):
    result = subprocess.run(["ffprobe", "-v", "error", "-show_entries",
                            "format=duration", "-of",
```

```

                                "default=noprint_wrappers=1:nokey=1",
filename],
    stdout=subprocess.PIPE,
    stderr=subprocess.STDOUT)
    return float(result.stdout)

def erase_polygon(img, path, points, left, top):
    path_points = []
    for j in range(0, len(path)):
        point = copy.deepcopy(points[path[j]])
        point[0] -= left
        point[1] -= top
        path_points.append(point)
    ppath = Polygon(path_points)

    h,w, _1 = img.shape

    a3 = np.array(path_points, dtype=np.int32 )
    im = np.zeros([h, w], dtype=np.uint8)
    cv2.fillPoly(im, [a3], (255,))

    return im

def apply_fft(input1, l=5, r=45):
    S_F = [fft(x) for x in input1]

    S_F = [np.array([ x if pos > l and pos < r else 0 for pos, x in enumerate(y)
]) for y in S_F]
    return np.array([abs(iff(x)) for x in S_F], dtype = float)

def apply_ica(input, components = 2):
    ica = FastICA(n_components=components, max_iter=2000)
    tmp = ica.fit(input)
    return ica.transform(input)

def apply_pca(input, components = 2):
    pca = PCA(n_components=components)
    tmp = pca.fit(input)
    return pca.transform(input)

rotations = [[]] * 1000
answers = [0] * 1000
my_answers = [0] * 1000

def get_heartrate(num, frames, answer_peaks, framerate, start_point,
finish_point, interesting_segment_length):
    end = finish_point * framerate
    start = start_point * framerate
    frames_in_interesting_segment = interesting_segment_length * framerate
    if len(frames) == 0:
        return -1

    # loop on video frames
    frame_counter = 0

    Us = []
    i = 0
    while (i < frames_in_interesting_segment):
        frame = frames[i]
        h,w,_ = frame.shape

        gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

```

```

points_ = get_points(frame)
if len(points_) == 0:
    i+=1
    print("Cannot find key points! Skip frame.")
    print(i)
    continue
points = points_[0]

left = max(min(points, key = lambda x: x[0])[0], 0)
right = min(max(points, key = lambda x: x[0])[0], w)
top = max(min(points, key = lambda x: x[1])[1], 0)
bottom = min(max(points, key = lambda x: x[1])[1], h)

width = abs(right - left)
height = abs(bottom - top)
show_frame = frame
face = frame[top:bottom,left:right]

if(face.size==0):
    print("Any face has not been found, skip frame!")
    i += 1
    continue
face_mask = face

path = [1,2,3, 30,13,14,15,72, 80, 71, 70, 69, 76]
path2 = [77,17,18,19,20,21,22, 23, 24, 25, 26, 74, 72, 80, 71, 70,
69, 76, 77]
path3 = [1,28,15,14,13,12,29,4,3,2]
face_mas = erase_polygon(img = face_mask, path = path, points =
points, left = left, top = top)
forehead_mask = erase_polygon(img = face_mask, path = path2, points
= points, left = left, top = top)
cheeks_mask = erase_polygon(img = face_mask, path = path3, points =
points, left = left, top = top)

mask = cv2.bitwise_or(forehead_mask, cheeks_mask)
mask = cv2.bitwise_and(mask, face_mas)
masked_face = cv2.bitwise_and(face_mask, face_mask, mask=mask)

mask = skin_detector.process(masked_face)
masked_face = cv2.bitwise_and(masked_face, masked_face, mask = mask)

cropped = get_cropped_frame(masked_face)

# cv2.imshow("image", cropped)
# cv2.waitKey(0)

pixels = [x for x in cropped.transpose(2,0,1).reshape(3,-1).T if x[0]
!= 0 or x[1] != 0 or x[2] != 0]
number_of_skin_pixels = len(pixels)

V = np.array(pixels, dtype=np.uint64).T

C = V.dot(V.T) / number_of_skin_pixels
#ax = plt.axes(projection='3d')
#ax.scatter3D(np.array(pixels).T[0], np.array(pixels).T[1],
np.array(pixels).T[2], cmap='Greens');

values, vectors = LA.eig(C)
Us.append(vectors)
#plt.pause(0.05)

```

```

        frame_counter +=1
        i += 1
    # plt.show()

    #base U
    Rs = []
    for u in Us:
        Rs.append([np.array([u[0])).dot(np.array(Us[0][0])),
np.array([u[0])).dot(np.array(Us[0][1])),
np.array([u[0])).dot(np.array(Us[0][2]))])
        Rot_v = []
        for row in Rs:
            Rot_v.append([sqrt(abs(1 - row[0][0]*row[0][0])), row[1], row[2]])
        rotation = np.array(Rot_v)
        f = np.arange(start_point, finish_point, 1/framerate)
        rotation_ica = apply_pca(np.array([rotation.T[0][0], rotation.T[0][1],
rotation.T[0][2]]).T, components=3)
        rotation_ica_fft = apply_fft(input1 = [rotation_ica.T[0],
rotation_ica.T[1], rotation_ica.T[2]], l=4, r=45)

        peaks = find_peaks(rotation_ica_fft[1], distance = 60 * framerate /
max_pulse)

    peaks_cnt = len(peaks[0])
    peaks_len = peaks[0][-1] - peaks[0][0]
    k = (end - start) / peaks_len

    global my_answers, answers, rotations
    my_answers[num] = peaks_cnt
    rotations[num] = rotation_ica_fft[1]
    answers[num] = len(answer_peaks)

    return peaks_cnt

def getAbsError(p):
    er = 0
    for i, rotation in enumerate(rotations):
        peaks = find_peaks(rotation, distance = 9, prominence=p)
        peaks_cnt = len(peaks[0])
        er += abs(peaks_cnt - answers[i])
    return er

def get_framerate(video):
    camera = cv2.VideoCapture(video)
    return int(camera.get(cv2.CAP_PROP_FPS))

def get_cropped_frame(frame):
    grayscale = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    mask = grayscale>0
    return frame[np.ix_(mask.any(1),mask.any(0))]

def get_frames(video, s, f):
    camera = cv2.VideoCapture(video)

    framerate = get_framerate(video)
    i = 0
    frames = []
    lenvid = get_length(video) * framerate
    if lenvid < f + 10:
        return []
    while (i < f):
        (grabbed, frame) = camera.read()

```

```

        if not grabbed:
            print("Frame hasn't been grabbed, skip frame!")
            continue
        if i > s and i < f:
            frames.append(get_cropped_frame(frame))
        i += 1
    camera.release()
    cv2.destroyAllWindows()
    return frames

def get_answer_peaks(s, f, time_beats):
    res = []
    for time in time_beats:
        if time > s and time < f:
            res.append(time)
    return res

answer_peaks = [[]]*1000

def run_test(video, s, f, num, framerate, length, time_beats):
    frames = get_frames(video = video, s = s * framerate, f = (s + length +
1) * framerate)
    answer_peaks[num]=get_answer_peaks(s=s, f=s+length,
time_beats=time_beats)
    return get_heartrate(num, frames=frames, answer_peaks=answer_peaks[num],
framerate=framerate, start_point=s, finish_point=s + length,
interesting_segment_length=length)

def apply_test(subdir, file):

    video = os.path.join(subdir, file)

    time_beats = []
    for s,d,fs in os.walk(subdir):
        for filename in fs:
            if filename.endswith(".rr"):
                f = open(os.path.join(subdir, filename))
                a = f.read().split('\n')
                for timebeat in a:
                    if timebeat != '':
                        time_beats.append(float(timebeat.split(' ')[0]))

    s = 0
    length = 6
    start_time = []
    framerate = get_framerate(video)
    video_len = get_length(video)

    i=0
    # while s + length < video_len - 10:
    start_time.append(s)
    run_test(video, s, s+length, i, framerate, length, time_beats)
    s += 30
    i +=1

    for i,time in enumerate(start_time):
        beats0 = next((x[0] for x in enumerate(time_beats) if x[1] > time),
-1)
        beats1 = next(( x[0] for x in enumerate(time_beats) if x[1] > time +
length), -1)
        if beats0 == -1 or beats0 == -1:
            print("Don't know an answer")

```

```

        continue
        answer = beats1 - beats0
        my_ans = my_answers[i]
        print(file + "_" + str(i) + "\t:\t" + str(answer) + "\t:\t" +
str(my_ans))

def main(user_input=None):
    # EXTRACT PULSE
    ap = argparse.ArgumentParser()
    ap.add_argument("-v", "--video", help = "path to the (optional) video
file")

    sumerr = 0
    orsumerr=0
    for subdir, dirs, files in os.walk("./data"):
        for file in files:
            if file.endswith(".avi"):
                apply_test(subdir=subdir, file=file)

    return

if __name__ == "__main__":
    main()

```

find_points.py

```

from imutils import face_utils
import dlib
import cv2

p = "shape_predictor_81_face_landmarks.dat"
detector = dlib.get_frontal_face_detector()
predictor = dlib.shape_predictor(p)

def get_points(image):
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

    # Get faces into webcam's image
    rects = detector(gray, 0)
    res = [];
    # For each detected face, find the landmark.
    for (i, rect) in enumerate(rects):
        # Make the prediction and transform it to numpy array
        shape = predictor(gray, rect)
        res_p = face_utils.shape_to_np(shape)
        res.append(res_p)

    return res;

```