Міністерство освіти і науки України

Київський національний університет імені Тараса Шевченка

Кафедра теоретичної кібернетики факультету комп’ютерних наук та кібернетики

ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

за спеціальністю «інформатика» 6.040302

на тему:

**Система відслідковування рухів рук для реабілітації та навчання**

Керівник випускної кваліфікаційної

роботи магістра

доктор фіз.-мат.наук, професор Крак Юрій Василійович

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

“\_\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2018 р.

Виконав студент 2 курсу магістратури

Єфремов Микола Сергійович

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

“\_\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2018 р.

Робота заслухана на засіданні кафедри теоретичної кібернетики та рекомендована до захисту в ДЕК, протокол № 11 від 24.05.2018 р.

Завідувач кафедри математичної інформатики професор Терещенко В.М.

Київ 2018

# 

**ЗМІСТ**

Contents

[ПЕРЕЛІК ВИКОРИСТАНИХ СКОРОЧЕНЬ 3](#_Toc516154765)

[ВСТУП 4](#_Toc516154766)

[РОЗДІЛ 1. Рогзляд методів та комп’ютерних систем для спрощеня процесу реабілітації 8](#_Toc516154767)

[1.1 Комп'ютеризований моніторинг 8](#_Toc516154768)

[1.1.1 Комп’ютеризовані системи реабілітації 8](#_Toc516154769)

[1.1.2 мобілізація систем 8](#_Toc516154770)

[1.1.3 Постановка задачі 9](#_Toc516154771)

[1.2. Огляд методів та технологій потрібних для виконання роботи 9](#_Toc516154772)

[1.2.1 Основи комп’ютерного зору 9](#_Toc516154773)

[1.2.2 Проективна геометрія 10](#_Toc516154774)

[1.2.3 Модель проективної камери 11](#_Toc516154775)

[1.2.4 Стереопара 13](#_Toc516154776)

[1.2.5 Побудова карти глибин 14](#_Toc516154777)

[1.2.6 Моделі відображення кольорів 15](#_Toc516154778)

[1.3 Огляд існуючих методів 16](#_Toc516154779)

[1.3.1 Системи захвату рухів, потребуючі носильну апаратуру 16](#_Toc516154780)

[1.3.2 Системи що використовують відеопотік 18](#_Toc516154781)

[1.3.3 системи з RGB-D та 3D камерами 19](#_Toc516154782)

[РОЗДІЛ 2. Огляд етапів розпізнавання траекторій руху рук 21](#_Toc516154783)

[2.1 Поетапне знаходження траекторії руху 21](#_Toc516154784)

[2.1.1 Силует людини 21](#_Toc516154785)

[2.1.2.Розпізнавання окремих частин тіла 23](#_Toc516154786)

[2.1.2 малювання траєкторії 24](#_Toc516154787)

[2.1.3. Зрівняння траєкторій. 24](#_Toc516154788)

[2.1.4 Виведення результатів 25](#_Toc516154789)

[РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ ВІДСЛІДКОВУВАННЯ ТА СПРОЩЕННЯ РЕАБІЛІТАЦІЇ РУХІВ ЛЮДИНИ 27](#_Toc516154790)

[3.1 Графічний інтерфейс 28](#_Toc516154791)

[СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ ТА ПОСИЛАНЬ 28](#_Toc516154792)

[1. *Наукова праця:* Pattern recognition and modern computers: O.G Selfridge – Massachusetts Institute of Technology, Lexington, Mass -1955 28](#_Toc516154793)

# ПЕРЕЛІК ВИКОРИСТАНИХ СКОРОЧЕНЬ

**2D, 3D** (2-dimentional, 3-d dimetional)

**AR** (Augmented Reality)

**VR**

**API(**

**SDK**

**CCD**(Charged Coupled Device) –тип матриці цифрової камери, яка переобразує заряди пікселей в аналоговий сигнал.

**Стереопара**

RGB-D

Інфорчервоне випромінювання

**Joint** – частина трьохвимірного об’єкта яка розмежує класи частин об’єкта

**офсет**

# ВСТУП

В сучасному світі, з розвитком технологій комп’ютерного зору поширюється його популярність. Комп’ютерний зір та алгоритми розпізнавання широко використовуються у багатьох сферах сучасності. На сьогодні алгоритми комп’ютерного зору використовують для вирішення багатьох задач, таких як задачі розпізнавання, задачі спрощення людям з обмеженими можливостями управління різними інтерфейсами, задачі створення ефективних алгоритмів керування безпілотними авто. Останні роки, різкий стрибок прогресу на ринку мобільних пристроїв сприяв поширенню попиту на додатки які використовують доповнену реальність(AR) у галузях пов’язаних з медициною, як для аналізу конкретних даних, так і для допомоги хворим у реальному часі.

Для використання технологій та алгоритмів комп’ютерного зору, як правило використовують камери, з роками з’являються все більш й більш “розумні камери”, які зазвичай складаються з кількох пар лінз на пару з іншими сенсорами. Так наприклад компанія FLiR Systems для розпізнавання живих об’єктів на зображенні у реальному часі використовує датчики тепла. Перші згадки комп’ютерного зору було опубліковано у 1950х роках професором Массачусецького технологічного Інституту (MIT) Олівером Селфрилджем («Розпізнавання образів і сучасні комп'ютери (1955)»)[1]. В своїй роботі він розглядав такі методи для отримання додаткової інформації про навколишні об’єкти:

* Використання рухомої навколо деякої осі камери
* Використання системи з багатьох камер
* Використання стереопари(система з двох камер)
* Різні алгоритми для постобробки отриманих даних про об’єкти

У 1970х роках досягнення у галузях математики дозволили розпізнавати ще точніше рухомі об’єкти використовуючи теорію аналіза динамічних сцен[2].

На кінець ХХ-століття на світовому ринку стало простіше знайти потрібні сенсори та датчики по доступній ціні, що сприяло зацікавленню багатьма сферами комп’ютерним зором. Алгоритми комп’ютерного зору допомогли зробити великий прорив у сферах медицини, та у сферах, які дозволяють покращити життя людям з різними обмеженнями. Якщо казати про медицину – комп’ютерний зір широко використовується у сучасних хірургічних відділеннях, з його допомогою лікарі більш точно знаходять ураження в тих місцях й тканинах, до яких вважко дістатись звичайним людьским зором. Також комп’ютерний зір використовують для аналізу поведінки хворих на шляху до одуження після таких вважких хвороб як інсульт, й т.д Використовуючи технічні рішення спрощують процес реабілітації хворих, використовуючи камери моніторять за поведінкою пацієнта на відстанні.

В ХХІ столітті існує багато доступних апаратних рішень, які значно спрощують розв’язання проблем, які можна вирішити використовуючи алгоритми комп’ютерного зору. В роботі береться особлива увага на проблеми зв’язані з спрощенням проходження реабілітації пацієнтів, які пережили різні форми інсульта. Також використовуючи теорію комп’ютерного зору можна створити системи, які несуть у собі задачі спрощення вивчення людьми з обмеженими вадами таких предметів як «даткильна абетка», алфавіт, та інш. використовуючи різні інтерактивні методи. Як було зазначено раніше, ці рішення містять у собі комбінацію стереопари та інших допоміжних сенсорів. Стереопара потрібна для перенесення трьовимірних координат у пам’ять машини, зазвичай вони зберігаються у вигляді списків з векторів. Для точності до стереопари популярно додавати інфрачервоні камери Якщо раніше такі рішення потребували значних затрат, то на сьогодні є багато бюджетних камер з датчиками глибини, та їх кількість продовжує зростати, так нещодавно було представлено компанією Qualcomm апаратне рішення, яке з легкістю можна використовувати на мобільних пристроях, що працюють на системі Android[3]. Але відкритого API не було надано. З рішень, до яких існує відкритий API та є доступними пересічному громадянину можна віднести Leapmotion camera від однойменної компанії та Kinect від Microsoft.

**Об’єкт дослідження -** процес розпізнавання рухів людського скелету, та в частності кисті руки.

**Предмет дослідження –** підходи й засоби реалізації програмної системи розпізнавання людського скелету у реальному часі, встановлення траєкторії руху людської кісті, та розрахунок точності промальованої траєкторії відносно заданого набору траєкторій.

**Методи дослідження –** у процессі було використано алгоритми передискретизації траєкторій, засоби реалізації захвату скелету людини та інші методи постобробки мапи глибини.

**Метою кваліфікаційної роботи** є розробка демонстративного програмного рішення, яке знаходить в реальному часі на зображенні людську руку та вказує наскільки точно людина відтворює рухи які їй пропонує програма.

Мета роботи породила такі підзадачі:

* Змоделювати поведінку скелета людини
* Огляд існуючих методів розпізнавання рухів
* Побудова алгоритму, що вказує точність рухів
* Розробка програмного рішення, використовуючи таке апаратне рішення, як камера Microsoft Kinect v1
* Аналіз отриманих результатів на точність та наскільки дане нам апартне рішення є корисним для використання для допомоги ускорення реабілітації хворих

Робота складаєтья з Н розділів.

Перший розділ роглядає основи комп’ютерного зору та існуючі методи розпізнавання людини використовуючи апаратні рішення, акцентуючи увагу на пристрій від Microsoft

# РОЗДІЛ 1. Рогзляд методів та комп’ютерних систем для спрощеня процесу реабілітації

## 1.1 Комп'ютеризований моніторинг

### 1.1.1 Комп’ютеризовані системи реабілітації

Комп'ютеризований моніторинг домашніх реабілітаційних вправ має багато переваг, і це викликало значний інтерес серед спеціалстів з комп'ютерного зору. На сьогодні є існує значна кількість систем реабілітації, більшість цільових пацієнтів - це пацієнти, потребуючі допомги в реабілітації після інсульту. Задача такої реабілітації є виправлення рухового відхилення, яке присутнє у більшості пацієентів. Під час реабілітації пацієнт потребує стороньої допомоги, та змушений знаходитись у лікарні, навіть якщо має незначні відхилення, наприклад у руховому апарті руки. Лікарі прописують таким пацієнтам спеціальні вправи, які зменшують біль та сприяють підвищенню рухливості. Ця вправа сприяє підвищенню рухливості руки і зменшенню болю. Вищезгадані системи допомагають пацієнтам у виконанні цих справ, але є дорогими системами, які потребують пацієнта знаходитись у лікарні під час проведення сеансів реабілітації. З розвитком теоріїї комп’ютерного зору постає можливість використовувати апаратні рішення з записом відеопотоків, для спрощення виконання вищезгаданих вправ.

### 1.1.2 мобілізація систем

В країнах пострадянського союзу, наявність вищезгаданих систем мінімальна, найчастіше це внаслідок недостатнього фінансування медичних закладів та реабілаційних центрів. Але можна знайти й алтернативний вихід.

Оскільки з ростом популярності комп’ютерного зору, та значним спадом цін на потрібну апаратуру, можна збирати такі системи власноруч, та надавати пацієнтам в «аренду», для користування на дому. В роботі розглядається використання апаратного рішення від компаніїї Microsoft, та проводиться дослідження, чи є дійсно воно корисним для використання у вищезгаданних цілях .

### 1.1.3 Постановка задачі

Відповідно до дослідження 2014 року, проведеного Всесвітньою федерацією сердечно-сосудистих захворювань, приблизно п'ять мільйонів людей в світі страждають від інсульту щорічно. Зважаючи на це, програмні рішення для прискорення реабілаціїї хворих є досить актуальною проблемою. Задача полягає у створенні програмного рішення, яке може бути доступним для звичайних обивателів, мала задачу казати які рухові вправи потрібно робити пацієнту, та відслідковуючи їхні руки, видавати інформацію наскільки точними відносно заданих траекторій є рухи користувача данної системи. Програмне забезпечення повинно знаходити у реальному часі силует людини, оскільки більшість пацієнтів які можуть проходити реабілітацію на дому мають вади з руховим апаратом руки чи кісті, то система повинна розпізнавати рухи людьскої руки по заданій траекторії та видавати на виході, наскільки вірно користувач зробив ту чи іншу вправу.

## 1.2. Огляд методів та технологій потрібних для виконання роботи

Проаналізувавши існуючи системи та підходи до реалізації спрощення методів прискорення процесу реабілітації пацієнтів було поставлено задачу розробити систему, що розпізнає людину на зображені, та в реалньому часі відслідоквуючи її рухи, а саме рухи кисті, надає інформацію, чи є покращення у пацієнтів, які проходять курс реабілітації, для цього було вирішено розробити систему, якак вказує точність відтворення рухів, та використовує методи комп’ютерного зору. Отже розглянемо яку базу та математичний апарат треба знати для реалізації нашої задачі.

### 1.2.1 Основи комп’ютерного зору

Комп'ютерний зір — це теорія та технологія створення програмних та апаратних рішень, метою яких виявлення, стеження та класифікація об'єктів[6]. З наукової точки зору, комп’ютерни зір належить до теорії та технологїї створення штучних систем, які отримують інформацію у вигляді зображень. Дані можуть бути представлені у вигляді зображень, у вигляді відеопотоку, чи навіть як 3D данні з медичних пристроїв Комп'ютерний зір — теорія та технологія створення машин, які можуть проводити виявлення, стеження та класифікацію об'єктів [1].

Як наукова дисципліна, комп'ютерний зір належить до теорії та технології створення штучних систем, які отримують інформацію у вигляді зображень. Відеодані можуть бути представлені у вигляді багатьох форм, таких як відеопослідовність, зображення з різних камер або тривимірними даними з медичного сканера [2].

### 1.2.2 Проективна геометрія

В геометрії комп’ютерного зору значну роль відіграє проективна геометрія. Проективну геометрію можна розглядати з координатної точки зору(вводяться поннятя гемоетричних об’єктів та аксіом й з них виводять властивості проективного простору), так і з аналітичної (розглядати все в координатах) Надалі будемо розглядати аналітичний підхід до проективної гемоетрії.

**Точки проекційної площини**. Розглянемо 2D простір, який надалі будем називати «проективною площиною».У той час, коли у евклідовому просторі точки описуються парою координат , на проективній площині точки описуються трьохкомпонентним вектором . При цьому для будь-якого не нульового числа , вектори відповідають одній й тій самій точці. А нульовий вектор не відповідає ніякій точці та просто не використовується. Такий опис точок на площині називається однорідними координатами.(homogeneous coordinates).

Точкам проективної площини можна зіставити точки звичайної Евклідової площини. Координатному вектору при зіставимо точку Евклідової площини з координатами . Якщо ж = 0, то будемо казати, що ця точка знаходиться у нескінченності. Таким чином, проекційну площину можна розглядати як Евклідову площину доповнену точками з нескінченності.

Перехід від однорідних координат до евклідових можна шляхом ділення координатного вектора на останню його компоненту й подальше її відкидання →. А від Евклідових координат перейти до однорідних можна за рахунок доповнення координатного вектора одиницею:

→.

### 1.2.3 Модель проективної камери

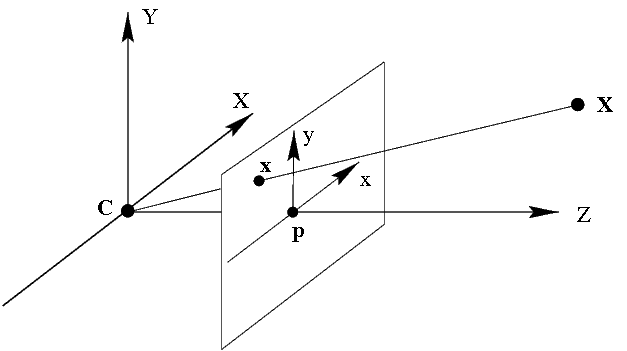


Рис 1. Модель камери. C - центр камери, Cp - головна ось камери. Точка X трьохвимірного простору проеціюється в точку x - на плоскості зображення.[4]

Для розв’язання задач комп’ютерного зору використовують різні камери. Сучасні CCD-камери добре описуються за допомогою наступної моделі, званої проективної камерою (projective camera, pinhole camera). Проективна камера визначається центром камери, головною віссю - променем починається в центрі камери і спрямованим туди, куди камера дивиться, площиною зображення - площиною на яку виконується проектування точок, і системою координат на цій площині. У такій моделі, довільна точка простору X проектується на площину зображення в точку x лежить на відрізку CX, який з'єднує центр камери C з вихідною точкою X (див. Рис. 1).

Формула проектування(проекціювання?) має простий математичний запис в однорідних координатах:

де X - однорідні координати точки простору, - однорідні координати точки площини, P - матриця камери розміру 3 × 4. Матриця P виражається наступ ним чином P = KR [I | -c] = K [R | t], де K - верхня трикутна матриця внутрішніх параметрів камери розміру 3 × 3 (конкретний вид наведено нижче), R - ортогональна матриця розміру 3 × 3, яка визначає поворот камери щодо глобальної системи координат, I - одинична матриця розміру 3 × 3, вектор c - координати центру камери, а t = -Rc.

Варто зазначити, що матриця камери визначена з точністю до постійного ненульового множника, який не змінить результатів проектування точок по формулі x = P X. Однак цей постійний множник зазвичай вибирається так, що б матриця камери мала вищеописаний вид. У самому простому випадку, коли центр камери лежить на початку координат, головна вісь камери сонаправлени осі Cz, осі координат на площині камери мають однаковий масштаб (що еквівалентно квадратним пікселям), а центр зображення має нульові координати, матриця камери буде позначатись як P.

, P = K [ I | 0] (1)

У реальних CCD камер пікселі зазвичай незначно відрізняються від квадратних, а центр зображення має ненульові координати. В такому випадку матриця внутрішніх параметрів набуде вигляду:

(2)

Коефіцієнти f, αx, αy - називаються фокусною відстанню камери (відповідно до загальних і вздовж осей x і y).

Крім цього, в силу неідеальності оптики, на зображеннях, отриманих з камер, присутні спотворення-дисторсії (distortion). Дані дисторсіїї мають нелінійну математичну запис :

Де - коефіцієнти дисторсії, є параметрами оптичної системи; ; (X ', y') - координати проекції точки відносно центру зображення при квадратних пікселях і відсутності дисорцій; (X ", y") - спотворені координати точки відносно центру зображення при квадратних пікселях.

Дисорсія не залежить від відстані до об'єкта, а залежить тільки від координат точок, в які проектуються пікселі об'єкта. Відповідно для компенсації дисторсії зазвичай виконується перетворення вихідного зображення отриманого з камери. Це перетворення буде однаковим для всіх зображень, отриманих з камери, за умови сталості фокусної відстані (тобто, матриці внутрішніх параметрів будуть співпадати).

У ситуації, коли відомі внутрішні параметри камери і коефіцієнти дисторсії кажуть, що камера відкалібрована.

### 1.2.4 Стереопара

Про знаходження тривимірних координат об’єктів спостереження (точок), можна казати, якщо присутня у системі хоча б одна стереопара.

**калібрування.** Нехай є дві камери, задані своїми матрицями P і P' (1) в деякій системі координат. У такому випадку говорять, що є пара відкаліброваних камер. Якщо центри камер не збігаються, то цю пару камер можна використовувати для визначення тривимірних координат об'єктів які спостережуються.

Найчастіше, система координат вибирається так, що матриці камер мають вигляд P = K [I | 0], P '= K' [R '| t']. Це завжди можна зробити, якщо вибрати початок координат збігається з центром першої камери, і направити вісь Z уздовж її оптичної осі.

Калібрування камер зазвичай виконується, за рахунок багаторазової зйомки деякого калібрувального шаблону, на зображенні можна легко виділити ключові точки, для яких відомі їх відносні положення в просторі. Далі складаються і вирішуються (груба результати) системи рівнянь, що зв'язують координати проекцій, матриці камер і положення точок шаблону в просторі.

Існують загальнодоступні реалізації алгоритмів калібрування, наприклад, Matlab Calibration toolbox. Так, наприклад у бібліотеці OpenCV реалізовано алгоритми калібрування камер і пошуку каліброваного шаблону на зображенні.

### 1.2.5 Побудова карти глибин

Карта глибини (depth map) - це зображення, на якому для кожного пікселя, замість кольору, зберігатися його відстань до камери. Карта глибини може бути отримана за допомогою спеціальної камери глибини (наприклад, сенсор Kinect є свого роду такою камерою), а так само може бути побудована по стереопарі зображень.

Ідея, що лежить в основі побудови карти глибини по стереопарі дуже проста. Для кожної точки на одному зображенні виконується пошук парної їй точки на іншому зображенні. А по парі відповідних точок можна виконати тріангуляцію і визначити координати їх прообразу в тривимірному просторі. Знаючи тривимірні координати прообразу, глибина обчислюється, як відстань до площини камери.

Парну точку потрібно шукати на епіполярной лінії. Відповідно, для спрощення пошуку, зображення вирівнюють так, що б все епіполярние лінії були паралельні сторонам зображення (зазвичай горизонтальні). Більш того, зображення вирівнюють так, що б у точці з координатами (x0, y0) відповідна їй епіполярная лінія задавалася рівнянням x = x0, тоді для кожної точки відповідну їй парну точку потрібно шукати в тій-же сходинці на зображенні з другої камери. Такий процес вирівнювання зображень називають ректифікацією (rectification). Зазвичай ректифікацію здійснюють шляхом ремеппінга зображення і її поєднують з позбавленням від дисторсій.

Після того як зображення ректифікованого, виконують пошук відповідних пар точок. Найпростіший спосіб проілюстрований на зображенні 3 і полягає в наступному. Для кожного пікселя лівої картинки з координатами (x0, y0) виконується пошук пікселя на правій картинці. При цьому передбачається, що піксель на правій зображенні повинен мати координати (x0 - d, y0), де d - величина звана невідповідність / зміщення (disparity). Пошук відповідного пікселя виконується шляхом обчислення максимуму функції відгуку, в якості якої може виступати, наприклад, кореляція околиць пікселів. В результаті виходить карта зсувів (disparity map), приклад якої наведено на рис. 3.

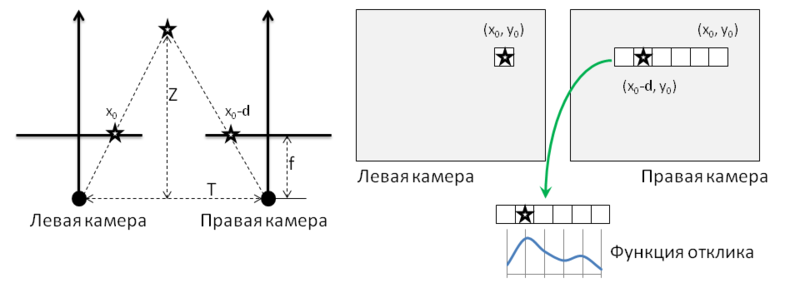


рис 3[5]

Власне значення глибини обернено пропорційні величині зсуву пікселів. Якщо використовувати позначення з лівої половини малюнка 4, то залежність

між disparity і глибиною :  
https://habrastorage.org/storage1/3421e1fa/fff760cd/b7f55435/48a8a99b.gif

Через зворотній залежності глибини і зміщення, роздільна здатність систем стерео зору, які працюють на основі даного методу, краще на близьких відстанях, і гірше на далеких.

### 1.2.6 Моделі відображення кольорів

Окрім використання лише геометричних методів, для розпізнавання об’єктів існують також методи, що використовують можливості сучасних нейронних мереж для ефективного розпізнавання положення у просторі силуету людини, або більше детальне розпізнавання окремих частин тіла, таких як кінцівки [25]. Але результати використаного підходу не були достатньо оптимального для розпізнавання у реальному часі. В деяких роботах [11, 38] для сегментації ділянки з рукою на зображенні ефективно використовувались можливості інших колірних просторів (YCbCr, HSV), відмінних від поширеного RGB. Але результати сегментації не були достатньо задовільними для практичного використання, через не врахуванні випадків з нетиповим освітленням. У ряді робіт [11, 12] крім дослідження оптимальних методів знаходження рук на зображенні, також розглядалося прикладне використання комп’ютерного зору у створенні інтерфейсних пристроїв для використання побутових та комп’ютерних засобів. Були також запропоновані нові алгоритми [14] для спостереження за положенням руки, використовуючи колір шкіри.

## 1.3 Огляд існуючих методів

### 1.3.1 Системи захвату рухів, потребуючі носильну апаратуру

В роботі [7] Хижникова-Клочкова згадується про використання спеціальної перчатки. Первинний широко визнаний пристрій для керування рукою в віртуальному просторі стало перчаткою Digital Data Entry Glove розроблена Г. Граймсом у 1981р. [8]. Вона була призначена для створення буквенно-цифрових символів у комп'ютерному просторі та запропонована як альтернатива клавіатури. Впоследствие перчатку використовував Д. Ланьер як метод маніпулювання предметами у віртуальному просторі, створеному за допомогою віртуального шолома. Така системе не є практичною у використанні, оскільки потребує додатко ще важкий шолом. Але вона може підійти для лікування деяких психічних розладів. У наш час існують методики подолання психічних розладів, які потребують пацієнта знаходитись у закритому приміщенні, так існують так званні кімнати викрикування, ціль яких, заводити людину у кімнату повну LCD дислпеїв та давати їм указання, викричатись, й залежно від крику, міняються кольорові тони LCD дисплеїв. Варто додати, що системи, які потребують VR шолом, можуть добре справлятись з тією задачею яку поставили перед собою. В україні існує стартап який несе у собі мету, відтворити задумку кмнат для викрикування, використовуючи шолом доповненої реальності, для допомоги людям з певними психчіними розладами.

Повертаючись до пацієнтів, що потребують надання допомоги в реабілітаціїї після випадків інсульта, варто зазаначити розробку кампаніїї Neofect. – RAPAEL Smart Glove[9]

Корейська технічна компанія "Neofect" створила розумне рішення для реабілітації хворих на інсульт, які потребують навчитися повторно використовувати свої руки. Рукавичка призначена для індукування нервової пластичності у пацієнта за допомогою спеціальних та індивідуальних вправ з імітацією. Продукт Neofect Rapael Smart Glove - це рукавичка, яку одягають на кисть руки для подальшого фіксування рухів, яка допоможе пацієнтам з інсульту відновити руховий апарт за допомогою спеціальних вправ, які користувач виконує в режимі віртуальної реальності. Рукавичка має вбудовані датчики, які вимірюють рухи пацієнта, коли вони проводять реабілітаційні вправи. Стороньо від розробника перчатки були додані елементи гри до даного апаратного рішення , що дозволяють хворим займатися і зберігають інтерес до повторних вправ реабілітації. Елементи гейміфікації призначені для підвищення у пацієнтів інтересу до виконнання вправ під час реабілітації.  
Реабілітаційні ігри для цієї перчатки на сьогодні оновлюються щомісяця, і кожна гра призначена для реабілітації навколо певних рухів. Для згинання і розтягування, вправи з видавлюванням апельсина, а також для пронації

передпліччя і супінації.

Деякі реабілітаційні центри використовують добре відому відео консоль від японської компаніїї Nintendo – Nintendo Wii. У минулому році в журналі BMC Muscoskeletal Disorders[10] була опублікована стаття німецьких лікарів з клініки ревматології та клінічної імунології в Charite University Medicine в Берліні. Вони розповіли про результати свого одноцентровую дослідження, вивчивши принципову можливість використовувати ігрову консоль Nintendo ™ Wii Fit Plus для реабілітації пацієнтів з ревматоїдним артритом.  
  
Принцип роботи цієї консолі полягає в тому, що людина управляє відбувається на екрані телевізора рухами свого тіла. Для цього в руках у нього спеціальні пристрої (для різних типів завдань та ігор вони можуть бути різними), які вимірюють стан його тіла, а на телевізорі встановлено інфрачервоний датчик, що зчитує показання цих пристроїв. Також, до цього додається балансувальна дошка з чотирма сенсорами (вперед-назад-вліво-вправо), за допомогою якої можна управляеться елементами йоги і тренувати рівновагу. Варто додати, що у 2017 році було випущенно оновлену версію Nintendo Switch, яка на сьогоднішній день має відкрите homebrew, що є великим проривом у мобільності, та несе у собі великі сподівання, що через декілька років, що допоміжні апаратні рішення можна буде навіть носити в кишенях, та виконувати реабілітаційні вправи частіше.

### 1.3.2 Системи що використовують відеопотік

У дослідженні Журнала Nature Communications.[11]розглядається, як оптико-електронна техніка і машинне навчання можуть бути використані для аналізу змін в рухових навичках, допомагаючи реабілітації пацієнтів.  
Новий терапевтичний підхід, над розробкою якого працювали в Швейцарському федеральному технологічному інституті (Swiss Federal Institute of Technology - ETH; Цюріх, Швейцарія), Гейдельберзькому університеті (University of Heidelberg; Німеччина) та інших установах, заснований на використанні оптогенетики для активації кортікостерінальной схеми. Оптогенетічна стимуляція в поєднанні з інтенсивною запланованою реабілітацією може привести до відновлення втрачених моделей руху (а не вимушеними компенсаційними діями), що підтверджується автоматичним аналізом поведінки на основі комп'ютерного зору в дослідженні моделі інсульту у щурів. Рухи щура реєструвалися за допомогою відеокамери і автоматично аналізувалися для моніторингу процесу реабілітації з метою коригування оптогенетіческой стимуляції. Результати показали, що оптогенетіческі активовані кортікоспаніческіе нейрони сприяють проростанню аксонів від інтактного до денервированного шийного відділу спинного мозку.

Варто згадати за проект Google creative lab[14], який довозляє виконувати розпізнавання людьских постатей у реальному часі, використовуючи при цьому лише ресурси інтернет браузера. Проект має назву PoseNet.

PoseNet можна використовувати для розпізнавання поз однієї або декількох людей, тобто є версія алгоритму. Розпізнавання однієї людини здійсньєься за одним, а декількох за іншим алгоритмом слідувати.  
Розглянувши реалізацію PoseNet, можна виділити декілька рівнів за якими відбувається оцінка пози. Вона відбувається у два етапи:

Вхідне зображення RGB подається через згорточну нейронну мережу.  
Для декодування пози використовується або алгоритм декодування з однієї позою, або багатопозиційний декодер, позиціонують оцінки довіри, позиції ключових точок, і показники довіри ключових точок від вихідних даних моделі.

На кафедрі теоретичної кібернетики факультету комп'ютерних наук та кібернетики при університеті імені Т.Г Шевченко було розроблено систему розпізнавання дактильної мови[22]. У роботі розглядаються підходи для вирішення задачі розробки технології, яка дозволяє розпізнавати елементи дактильної мови. Автор використовує методи перетворенн вхідного кольорового зображення у сірі тони, тобто перетворює вхідний потік, що містить у собі RGB модель, у потік зображень представлених моделю HSV. Також було показано ефективність методів контурного аналізу та методів скелетизації використовуючи аналіз кутових параметрів рухів. Робота присвячена розробки системи для ознайомлення користувача з дактильною абеткою, але методи розроблені Шкілюнюком можна використовувати й для розробки систем комп’ютерної реабілітації

### 1.3.3 системи з RGB-D та 3D камерами

Камери що мають можливість знімати крім зображення у RGB, а ще й будувати карту глибини зображення, є на сьогодні рішенням, яке дозволяє фіксувати та відслідковувати трьохвимірні об’єкти та їх рухи у реальному часі та водночас потребують мінімум налаштувань та є досить легкі у налаштуванні їх під проблеми у різних сферах. В роботі [7] було проведено досілдження, яке показало, що рухи руки у віртуальній реальності, що зчитуються за допомогою сенсорів LeapMotion, більш точно відповідають рухам руки у реальному середовищі.

Пошуки програмних рішень, що основані на камері від Microsoft привели до новини 2016 року, де було описано, що при університеті Pompeu Fabra, що у Барселоні було започатковано проект, який створював додатки для камер глибини, які допомогали пацієнтам одужувати у ігровій формі. [12] Автори проекту стверджують, що, використовуючи доповнену реальність, можна не просто відновлювати рухові функціїї організму після інсульту, але й проблеми з ковтанням, та нейро-психологічні розлади.

3D камери відрізняються від RGB-D камер тим, що в них присутня лише стереопара (відустній інфорчервоний проектор) В роботі [23] описується створення моделей трьовимірних моделей використовуючи стереопару за допомогою методу гнучких моделей, та метод наближення полігонами для більш детальної візуалізації ідентифікованого об’єкту[22, 24]

# РОЗДІЛ 2. Огляд етапів розпізнавання траекторій руху рук

Розглянувши існуючі роботи та програмні рішення в області допомоги людям, що перенесли існульт, та проаналізувавши їх на недоліки було розроблено план дій, для реалізації власної системи. Оскільки найчастіше у такій системі, якак може бути використана й у домашніх умовах потребують ті, в кого є розлади під час руху кисті рук чи самих рук. Було вирішено розробити систему, яка отримує у реальному часі на вхід відеопотік даних, та порівнюючи його з заданою інформацію, повідомляє, наскільки точно й правильно було виконано ту чи іншу вправу. В наступних підрозділах описується які алгоритми було використано для реалізаціїї ціїї задумки.

## 2.1 Поетапне знаходження траекторії руху

Для виділення силуету людини на зображені у реальному часі, можна використовувати багато методів, які описані у минулому розділі. В роботі Пустовіта О.В “Розпізнавання траекторіїї руху кисті людини” [15] описуються методи видалення фону, доти доки не не залишиться на вихідонму зображенні тільки силует людини. Такі методи використовують алгоритми для видалення фону, що оперують з кольоровою моделю YCrCb. В даній роботі було показано, що програмні рішення для розпізнавання використовуючи звичайні камери RGB, потребують реалізаціїї процесу видалення багатьох ознак, які можуть некоректно представити вихідні данні. Тому було вирішено використовувати камеру, що має можливість розраховувати карту глибини зображень.

### 2.1.1 Силует людини

Для того щоб розпізнати на вхідному зображенні певний об’єкт, треба чітко відокремети його від інших присутніх об’єктів. Постає проблема, що під час малого освітлення, або взагалі коли освітлення відсутнє, об’єкти сильно зливаються з фоном, що дає неможливим подальшу обробку та розпізнання об’єктів, які цікавлять для виконання тієї чи іншої мети. Камери, які здатні будувати карту глибини, використовуючи сенсор, що містить у собі інфрачервоний проектор, за допомогою якого виконується побудова матриці зображення, елементи якої на відмінну від звичайних камер, що зберігають кольори, несуть інформацію у кожному пікселі про яка є відстань від камери, до об’єкту на який потрапляє інфрачервоне випромінювання.[17] Використовуючи такий сенсор, ми отримуємо матрицю відстаней, та робимо подальший її аналіз, для відокремлення потрібних нам об’єктів.

Для того, щоб прийняти рішення чи є на зображенні, яке було отримано з сенсора глибини потрібним нам об’єкти робиться локальний аналіз кожного пікселя на певні ознаки. Розпізнавання виконується за допомогою структуру прийняття рішень, яка навчена на безлічі зразків. Для того, щоб структура працювала, треба надати класифікатор з великою кількістю значень ознак, які несуть у собі потрібну інформацію, необхідну для розпізнавання об'єкта.[18] Ознаки для кожного пікселя вираховуються з наступної формули[16, 18]:

**f = d( x + u/d(x) ) — d( x + v/d(x) )**  (3)

де u, v - пара векторів зсуву, а d (x) - глибина пікселя, тобто відстань від Kinect до точки, проецирующейся на x. Це дуже простий ознака, по суті це просто різниця в глибині двох пікселів, зміщених щодо вихідного на u і v. (Варіюючи u і v, отримуємо набір ознак.

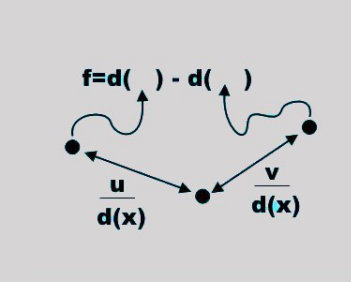


Рис.4 Знаходження ознак для кожного пікселя

Для остаточного результата знаходження одних лише ознак не є достатнім. Для того, щоб система мала можливість розпізнати на зображенні силует людини, її треба прогнати через класифікатор для правильної класифікації для кожної частини тіла на тестовому наборі зображень. Враховуючі всі можливі рухи та пози, навчання такої системи може потребувати базу з близько мільйону зображень. Для вирішення цієї проблеми, було обрано використовувати Kinect v1, оскільки компанія Microsoft, разом з SDK для Kinect, постачає базу даних, яка була натренована на різновидності класифікатора «ліс рішень», тобто на наборі дерев рішень. Кожне дерево навчалося на наборі ознак на глибинних зображеннях, які заздалегідь були прив'язані до відповідних частин тіла. Тобто дерева перебудовувалися до тих пір, поки вони не стали видавати правильну класифікацію для певної частини тіла на тестовому наборі зображень. Ця база має безліч дерев, які були навчені на вибірці з більше ніж 1 мільйону зображень.

### 2.1.2.Розпізнавання окремих частин тіла

Використовуючи метод описаний вище, навчені класифікатори видають ймовірність належності пікселя до тієї чи іншої частини тіла. Після цього області з максимальною вірогідністю назначають собі окрему частину тіла кожного типу. Наприклад, область буде віднесена до категорії 'нога', якщо «ніжний» класифікатор видав максимум ймовірності в цій області. На рисунку 5, можна бачити, як помилково людську руку було класифіковано до класу “нога” у випадку, коли людський скелет було зчитано не повністю.

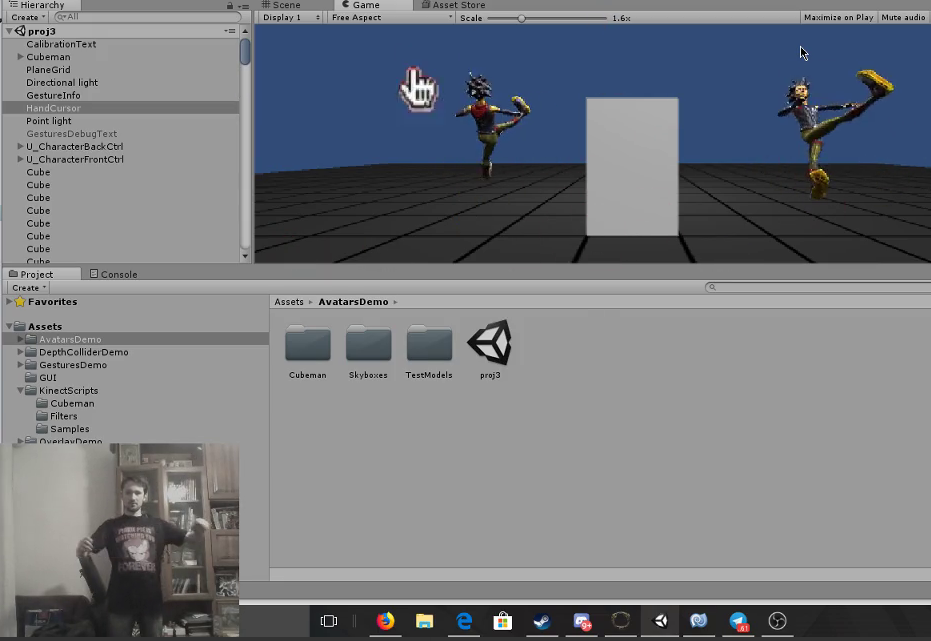


Рис 5. Випадок помилкової класифікації частин тіла

Розпізнавання об’єкта, як було сказано у розділі 2.1.1 використовує інформацію про кожен піксель, В статті П.Мьєра[20] детально описується як саме Kinect використовує Гаусівський розподіл на пару з алгоритмом середнього зсуву. Система використовуючі дані алгоритми робить аналіз кожного пікселя, та перераховує кожен кадр, до якої частини тіла можна віднести кожен з пікселів, який зберігає у собі елемент матриці карти глибини кадрів відеопотоку.

### 2.1.2 малювання траєкторії

Для запису у реальному часі, траєкторій, які у повітрі малює користувач, було вирішено створювати об’єкт, який буде відслідковувати зміну 3D координат об’єкта який належить класу “кисть руки”, та синхронізуючись з ним змінювати власні координати, паралельно записуючи пройдені координати у список. Візуалізація кожної точки, яка буде належати такому списку надасть візуальну картину того, що користувач малює певну траєкторію у повітрі. Варто зазначити, що малюючи траєкторію, треба зменшити частоту кадрів , тобто збільшити різницю між кадрами, на яких створюється новий об’єкт що є елементом траєкторії, оскільки, як було помічено у розділі 2.1.2 – можуть трапитись ситуації, що піксель було класифіковано до некоректного класу. В наступному розділі роглядається знаходження оптимальних частот оновлення кадрів.

### 2.1.3. Зрівняння траєкторій.

У програмне рішення завантажені 10 пар траєкторій, які були заздалегідь намальовані та внесені у програмне рішення як списки, що складаються з тривимірних точко для того, аби користувач міг дізнатись наскільки добре він може повторювати ту чи іншу траєкторію.

Вварто зазначити, що намальована та задана траєкторії мають різну щільність точок та різну їх кількість, що залежить від швидкості, точності рухів користував та багатьох інших факторів, таких як навколишнє середовище, що може мінятись під час виконання вправ. Це сприяє винекненню проблеми того, що точки траєкторії, яку малює користувач та точки заданої траєкторії у програмі не можуть бути однозначно порівняні. Для вирішення данної проблеми потрібно зрівняти кількість точок, з яких складаються траєкторія. що намальована користувачем та траєкторї які використовуються у програмному рішенні. Для цього робиться операція передескритизацїї зображень( в нашому випадку траєкторій). Для вирішення цієї проблеми, робиться операція передискретизації заданої траєкторії яка визначена множиною *М* точок у множину з *N* рівнорозподілених точок, які належать до траєкторії, яка вводиться у реальному часі користувачем програми на рисунку 6 наведено приклад дескритизованої фігури у різну кілкість точок.

w

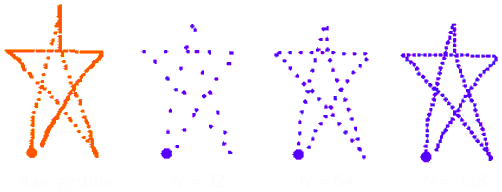


Рис 6.

траєкторія

N = 32

N = 64

N = 128

На практиці велике число точок у множині M значно збільшує процес порівняння фігур. Було визначено, що оптимальний діапазон для множини  M є значення між 60 ≤ *M* ≤ 250. Передисеретизація є не настільки поширеним методом, як наприклад фільтрація, але його також використовують у різних роботах присвячених проблемам комп’ютерного зору.[20, 21]

### 2.1.4 Виведення результатів

Для досягнення поставленої мети, було вирішено знаходити офсет між двома траєкторіями, та розраховувати відсоток, наскільки великим є відхилення між точками остаточної та заданої траєкторій. Оскільки до фігур було задіяно алгоритм дескритизації – то кількість тривимірних точок, з яких вони побудовані є однакова. Це дозволяє нам зробити порівняння двох траєкторій використовуючи прості алгебраїчні формули. Проходячи по двом спискам R та S, між двома найблищими точками та , , Де - S – список точок з яких складається задана траєкторія, а - R – список точок з яких складається намальована користувачем траєкторія. Для знаходження наскільки відмінно від заданої була намальована остаточна траєкторія, рахується значення офсету між двома точками та

-

# РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ ВІДСЛІДКОВУВАННЯ ТА СПРОЩЕННЯ РЕАБІЛІТАЦІЇ РУХІВ ЛЮДИНИ

Була створена програма, що представляє собою демонстраційне програмне забезпечення, у якому користувачеві надається можливість перевірити точність виконання рухів рукою, методом порівняння заданих траєкторій з траєкторією руху руки людини, що користується програмним рішенням. Для реалізації програми було обрано використовувати камеру від компанії Microsoft – Kinect v1, це є камера, що складається з стереопари та датичку глибини, який представленний у вигляді сенсора, що випромінює інфрочервоні промені. У наступних підрозділах описуються технічні особоливості створеного програмного забезпечення та особливості проектування данного програмного забезпечення.

На рисунку 3.1 представлено рекомендовану середу користування програмною системою. Системні вимоги до програми мінімальні, підійде комп’ютер з процесорором споживчого сегменту та інстальована на ньому операційна система Windows, яка підтримує програмне забезпечення для камери глибини Kinect, та звісно сама камера.

****

Рис 3.1  Рекомендована середовище роботи системи(у положені стоя)

# 3.1 Графічний інтерфейс програмного застосунку

Інтерфейс програмного застосунку має одновіконну структру. Програмне вікно, яке відображає синхоризовані відеопотоки з стереокамери та потік відеоданих з камери глибини. У лівому верхньому кутку відібражається точнісь виконання руху. Трохи нище є меню, яке дозволяє користувачеві міняти місцезнаходження фігур на екрані, та вказувати швидкість малювання та кількість точок, з яких повинна складатись кінцева траєкторія. На рисунку 3.2 показаний інтерфейс створеної програми у процесі запису траєкторії руху.

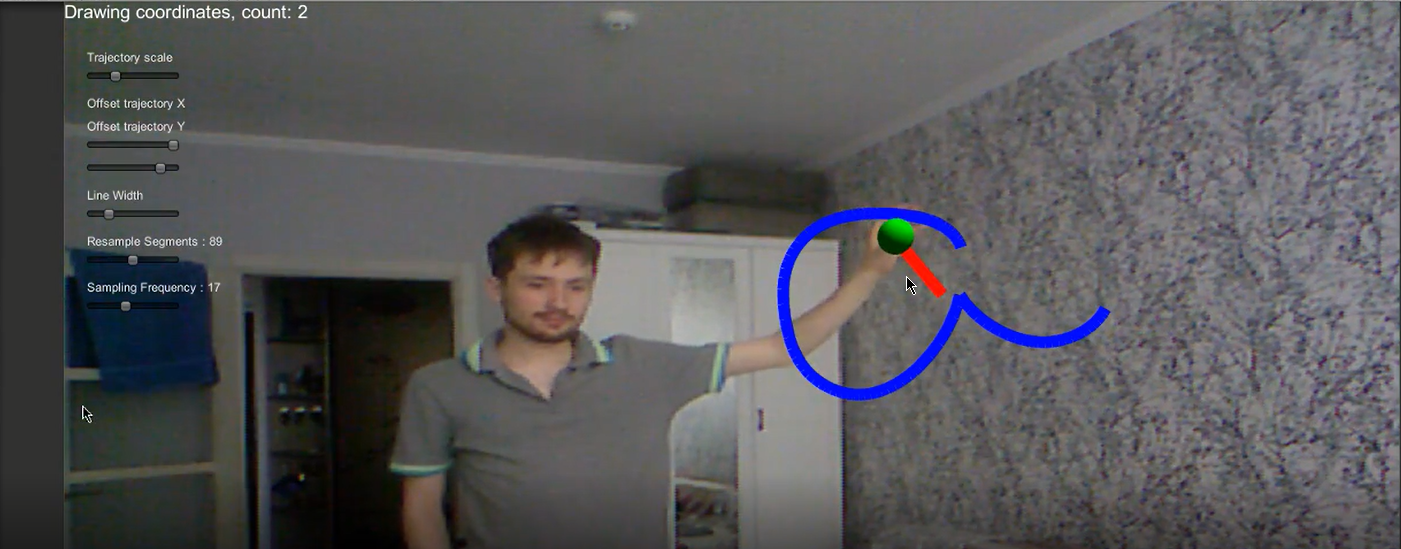


Рис 3.2 Інтерфейс програмного застосунку

Повзункове меню у лівому верхньому кутку інтерефейса було вирішено додати для спрощення налаштування програми під різні середовища. Користувач може самостійно підігнати кількість точок, товщину траєкторії та швидкість відображення, орієнтуючись на такі фактори оточення як світло, наявність великої кількості стороннії об’єктів. Для зміни значення параметрів використовуються графічні слайдери. Для задання позиції заданих фігур, користувачєві пропонується змінювати координати по осі X та Y теж за допомогою графічного слайдеру.

## 3.2 Обрана мова програмування реалізації, середовище та бібліотеки

**3.2.1 Рушій гри – Unity**

Поставлена задача у ці роботі потребує не лише відображення відеопотоку даних з камери, але й створення графічних об’єктів, з яким потім потрібно виконувати ряд певних дій. Для спрощення цієї задачі було обрано виконувати розробку програми на мові програмування C#, тому що

# СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ ТА ПОСИЛАНЬ

# 1. *Наукова праця:* Pattern recognition and modern computers: O.G Selfridge – Massachusetts Institute of Technology, Lexington, Mass -1955

посилання на електроне джерело: <https://dl.acm.org/citation.cfm?id=1455310>

2. *Книга*: «Теорія аналізу динамічних сцен»: Ганс-Хельмут Нагель -1979-

3. *Патент*: Depth map generation techniques for conversion of 2d video data to 3d video data: пат. US20110096832A, Qualcomm Inc: Rong Zhang, Ying Chen, Marta Karczewicz,публ:2009.10.23; посилання:https://patents.google.com/patent/US9100640

4. *Книга:* «Multiple View Geometry in Computer Vision»: Richard Hartley, Andrew Zisserman, «Cambridge University Press -2004». -673 стр.

5. *Книга:* «Learning OpenCV»: Gary Bradski, Adrian Kaehler. «O'Reilly Media - 2008». -  580 стр.

*6. Книга*: «Компьютерное зрение»: Л. Шапиро, Дж. Стокман. «Бином. Лаборатория знаний -2006» -752 стр.

7. *Наукова праця:* Виртуальная реальность как метод восстановления двигательной функции руки: А.Е. Хижникова, А.С. Клочков, А.М. Котов-Смоленский, Н.А. Супонева, Л.А. Черникова – ФГБНУ «Научный центр неврологии», Москва -2016

8. *Патент:* Digital data entry glove interface device; пат: US4414537A, Nokia Bell Labs , [Gary J. Grimes](https://patents.google.com/?inventor=Gary+J.+Grimes) -1983

9. *Електроний ресурс:* RAPAEL Smart Glove for Hand Rehab [<http://www.neofect.com/en/product/rapael/>]

10. *Наукова праця:* Isometric hand grip strength measured by the Nintendo Wii Balance Board – a reliable new method: A. W. Blomkvist, S. Andersen, E. D. de Bruin, M. G. Jorgensen -*BMC Musculoskeletal Disorders -2016.* посилання на електроне джерело:

<https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/s12891-016-0907-0>

11. *Наукова праця:* Optogenetically stimulating intact rat corticospinal tract post-stroke restores motor control through regionalized functional circuit formation: A. S. Wahl, U. Büchler, A. Brändli, B. Brattoli, S. Musall, H. Kasper, B. V. Ineichen, F. Helmchen, B. Ommer & M. E. Schwab

посилання: <https://www.nature.com/articles/s41467-017-01090-6>

12.*Електроний ресурс*: Система реабілітації комп’ютерним зором;

посилання: [<http://rgs-project.upf.edu/>]

*13.Частина книги, періодичного видання:* ФИЛЬТРАЦИЯ ДАННЫХ ГЛУБИНЫ С СЕНСОРА KINECT: Дусеев В.Р.  Мальчуков А.Н.  Мыцко Е.А. Журнал

Современные проблемы науки и образования. – 2015. – № 1 (часть 1); Национальный исследовательский Томский политехнический университет»

 Посилання: <https://www.science-education.ru/ru/article/view?id=17554>

14. *Електроний ресурс*:  білбліотека мови програмування JavaScript.

Посилання: [<https://github.com/tensorflow/tfjs-models/tree/master/posenet>]

15.

15«РОЗПІЗНАВАННЯ ТРАЄКТОРІЇ РУХУ КИСТІ ЛЮДИНИ» Пустовіт О В

16 Real-Time Human Pose Recognition in Parts from Single Depth Images Jamie Shotton Andrew Fitzgibbon Mat Cook Toby Sharp Mark Finocchio Richard Moore Alex Kipman Andrew Blake Microsoft Research Cambridge & Xbox Incubation

https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/BodyPartRecognition.pdf

17 ] D. Grest, J. Woetzel, and R. Koch. Nonlinear body pose estimation from depth images. In In Proc. DAGM, 2005. 1, 2

18 S. Knoop, S. Vacek, and R. Dillmann. Sensor fusion for 3D human body tracking with an articulated 3D body model. In Proc. ICRA, 2006

19 D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. IEEE Trans. PAMI, 24(5), 2002.

20. Keneth Dawson-Howe, A Practical Introduction to Computer Vision with OpenCV / Keneth Dawson-Howe // Wiley: 2014. – 235 p.

21. Singh SK, Chauhan DS, Vatsa M, Singh R (2003) A robust skin color based face detection algorithm. Tamkang J Sci Eng 6(4):227-234

[useful link] [All about Kinect](http://www.i-programmer.info/babbages-bag/2003-kinect-the-technology-.html) <http://www.i-programmer.info/babbages-bag/2003-kinect-the-technology-.html>

https://groups.google.com/forum/#!topic/cudacsmsusu/LBa88kygILA

22

<http://www.icyb.kiev.ua/file/disser/dis_Shkilniuk.pdf>

Для того, щоб дізнатись оптимальне число кадрів, було зроблено дослідження: Було порівняно різні швидкості малювання траєкторії “Зірка” використовуючи різні частоти оновлення кадрів за секунду.

23 Местецький Л.М. Восстановленний в реальном времени пространственных характеристик гибкого объекта по стереопаре изображений. Метематические методы разпознавания образзов: сборник докладов 13 всерос конф(30 сентября -6 октября 2007г) г. Зеленогорск. : Макс Пресс, 2007 – стр 359-363

24 Местецкий Л.М Непрерывный скелет бинарного растрового изображения . Труды 8й международной конф. По комп. Графике и анимации МГУ 1998 стр 71-78

25. Комп’ютерне розпізнавання жестів: програмно-алгоритмічний підхід / О. Годич, М. Давидов, Ю. Нікольський та ін. – Л. : ТОВ “Компанія “Манускрипт”, 2011. – 310 с.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1 кадр/c | 10 кадрів/с | 40 кадрів/с | 120 кдр/с | 200 кадрів/c |
| Повільно: |  |  |  |  |  |