



П. В. Сердюк, І. В. Бойчук

Національний університет "Львівська політехніка", м. Львів, Україна

## НОРМАЛІЗАЦІЯ ТРИВИМІРНИХ КООРДИНАТ У БІБЛІОТЕЦІ MEDIAPIPE ДЛЯ ВИМІРЮВАННЯ КУТІВ СУГЛОБІВ

Визначення та коректне використання тривимірних координат людини та глибини зображення досі залишається важливим завданням оцінювання пози людини, що впливає на точність визначення її рухів і розрахунку кутів суглобів. Запропоновано новий підхід до нормалізації тривимірних координат, визначених за допомогою моделі оцінювання пози людини MediaPipe Pose Landmarker. Цей підхід поєднує набори ключових точок людини Landmarks (координати зображення) та WorldLandmarks (тривимірні координати), доповнюючи їх спеціальними перетвореннями, зокрема уніфікацією масштабу та приведенням координат до спільного початку відліку. Розроблено математичну модель суглоба людини. Для оцінювання ефективності підходу та точності розрахованих на підставі нормалізованих координат значень кутів суглобів розроблено програмне забезпечення, що дає змогу провести тестування запропонованого підходу на відео та зображеннях людини, які містять згинання суглобів у складних позах з різних ракурсів та у різних площинах. Набір тестових даних сформовано за допомогою програм для моделювання тривимірних поз людини Magic Poser та Ultimate Poser. На підставі результатів тестування програмного забезпечення встановлено покращення точності визначення кутів суглобів. Проведено порівняння з наявними способами нормалізації координат та підтверджено ефективність запропонованого підходу. Для аналізу точності використано метрики середньої абсолютної похибки (MAE) та середньої абсолютної відсоткової похибки (MAPE), що дало змогу оцінити відхилення розрахованих кутів від еталонних значень. Серед п'яти методів нормалізації тривимірних координат людини, розглянутих у дослідженні, запропонований підхід продемонстрував найвищу точність для двох наборів тестових даних ( $MAE = [12,6^\circ; 7,2^\circ]$   $MAPE = [14,0\%; 8,0\%]$ ). Незважаючи на деяке покращення точності визначення кутів суглобів спостерігається висока похибка у складних випадках, де кінцівку видно погано, вона знаходиться у тіні чи ключові точки накладаються. Подальші дослідження можуть полягати у: поєднанні бібліотеки MediaPipe з іншими інструментами визначення глибини зображення; аналізу координат попередніх кадрів (у випадку відео); використанні декількох камер; використанні технології LiDAR та лазерних давачів.

**Ключові слова:** глибина зображення; оцінювання пози людини; ключові точки; кінематика суглоба; реконструкція 3D-пози.

### Вступ / Introduction

Зі зростанням потужності обчислювальних машин і розвитком сучасних технологій штучного інтелекту засоби комп'ютерного зору набули стрімкого розвитку. Це дало змогу розв'язати та застосовувати задачу визначення пози людини (англ. *Human Pose Estimation*, HPE) у найрізноманітніших сферах людського життя, таких як медицина та реабілітація, розваги та спорт [17], безпека підприємств та військова справа, у режимі реального часу та доповненої реальності [16].

Бібліотека MediaPipe стала однією із найпоширеніших інструментів для HPE та має значні переваги серед конкурентів [1]. Попри високу ефективність та гнучкість, коректне визначення координати Z та глибини зображення досі залишається викликом та є актуальним завданням, позаяк це значення не відповідає глибині об'єкта, а є відносним залежно від його розміру та масштабу [10]. Це ускладнює процес реконструкції 3D-по-

зи людини та знижує точність вимірювання кутів суглобів. Для вирішення цього завдання застосовували геометричні методи [11], спеціальні камери RealSense [19], системи захоплення рухів MoCap [2]. Також потрібно вдосконалити процес нормалізації координат ключових точок, що дасть можливість усунути варіативність значення координати Z.

**Об'єкт дослідження** – визначення тривимірних координат суглобів людини.

**Предмет дослідження** – методи і засоби нормалізації тривимірних координат у бібліотеці MediaPipe, що дасть змогу встановити їхній вплив на точність визначення кутів суглобів тіла людини.

**Мета роботи** – розробити програмне забезпечення для реалізації запропонованого методу нормалізації тривимірних координат, визначених за допомогою бібліотеки MediaPipe, що дасть можливість точніше визначати положення тіла людини у складних позах та вимі-

### Інформація про авторів:

**Сердюк Павло Віталійович**, канд. техн. наук, доцент, кафедра програмного забезпечення.

Email: [pavlo.v.serdiuk@lpnu.ua](mailto:pavlo.v.serdiuk@lpnu.ua); <https://orcid.org/0000-0002-2677-3170>

**Бойчук Іван Віталійович**, магістр, кафедра програмного забезпечення.

Email: [ivan.boichuk.mnpzm.2022@lpnu.ua](mailto:ivan.boichuk.mnpzm.2022@lpnu.ua); <https://orcid.org/0009-0005-6141-3161>

**Цитування за ДСТУ:** Сердюк П. В., Бойчук І. В. Нормалізація тривимірних координат у бібліотеці mediapipe для вимірювання кутів суглобів. Науковий вісник НЛТУ України. 2025, т. 35, № 2. С. 163–169.

**Citation APA:** Serdyuk, P. V., & Boichuk, I. V. (2025). Normalisation of three-dimensional coordinates in the mediapipe library for measuring joint angles. *Scientific Bulletin of UNFU*, 35(2), 163–169. <https://doi.org/10.36930/40350219>

рювати кути згинання/розгинання суглобів для розв'язання прикладних задач.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати наявні літературні джерела та інтернет-ресурси, на підставі чого буде розглянуто наявні методи вирішення завдання визначення глибини зображення та координати  $Z$ ;
- дослідити та обґрунтувати доцільності розроблення нового алгоритму нормалізації тривимірних координат, що дасть змогу якомога точніше вимірювати кути згинання/розгинання суглобів;
- верифікувати та протестувати новий метод нормалізації координат, на підставі чого буде визначено його ефективність та точність порівняно з іншими методами.

**Аналіз останніх досліджень та публікацій.** Автори роботи [4] визначають переваги та недоліки бібліотеки MediaPipe та довели, що точність оцінювання її роботи залежить від кута огляду камери та типу виконуваних вправ. У наступному дослідженні [3] вони пропонують новий метод реконструкції 3D-поз, заснований на використанні 2D-поз бібліотеки MediaPipe для стереокамер і алгоритмів злиття без попереднього стерео-калібрування координат, поєднуючи переваги високої точності 2D HPE із підвищеною зручністю використання 3D-координат. Отримані ними результати на підставі власного набору даних показали, що їхній метод забезпечує значно покращену продуктивність порівняно з монокулярним 3D HPE (медіана RMSE 30,1 порівняно з 56,3,  $p$ -значення нижче 10<sup>-6</sup>) і може показати, що продуктивність її роботи достатня для розпізнавання можливих помилок.

Автори публікації [10] стверджують, що бібліотека MediaPipe є найбільш зрілим представником 3D-систем для розпізнавання людського тіла, проте вони довели, що монокулярна камера не може точно визначити значення глибини  $Z$  у даних пози людини. Вони пропонують власний метод для корекції (нормалізації) значення  $Z$  на підставі статистичного методу нахилу тіла та нормалізацію пропорцій кінцівок. Для вирішення проблем тремтіння, затримки та періодичного шуму в кількох кадрах, автори виконують фільтрацію "One Euro" та середнього значення спільних даних.

У роботі [12] автори пропонують нове рішення для завдання оцінювання глибини монокулярного зображення з удосконаленнями, яке покращує кількісне та якісне розуміння карт глибини порівняно з наявними методами. Їхня модель складається з двох паралельних мереж автоматичного кодування: автоматичний кодер, на підставі ResNet-50, для видобування функції із вхідного зображення та згорткова мережа графів GCN (англ. *Graph Convolutional Networks*) для оцінювання карти глибини розташування. Дослідження доказало, що запропонований метод дав порівняльні та багатообіцяючі результати з високою точністю прогнозу 89 % на загальнодоступному наборі даних KITTI та Make3D.

Автори роботи [8] представили нову модель DLNet для піксельного оцінювання глибини, яка одночасно виділяє глобальні та локальні характеристики за допомогою блоку глибини Linformer, що складається з інноваційних м'яких розділених багатошарових блоків перцептронів (математична або комп'ютерна модель сприйняття інформації мозком). Під час експериментів на тестах KITTI та Make3D запропонований DLNet досягає продуктивності, зменшуючи часові та просторові складності більш ніж на 62 і 56 % відповідно, тестуван-

ня в ситуаціях реального світу додатково демонструє сильну практичність і здатність до узагальнення запропонованої моделі.

Аналіз останніх досліджень свідчить про те, що наявні методи оцінювання пози людини на підставі MediaPipe мають певні обмеження, зокрема – залежність від кута огляду камери, неточність визначення координати глибини ( $Z$ ) та потребу у додаткових методах нормалізації. Хоча запропоновані рішення, такі як використання стереокамер та глибокі нейронні мережі, демонструють покращення точності визначення кутів пози людини, вони потребують спеціального обладнання або значних обчислювальних ресурсів. Отже, залишається актуальною проблема нормалізації тривимірних координат для підвищення точності розрахунку кутів суглобів без використання додаткового апаратного забезпечення.

**Матеріали та методи дослідження.** Під час проведення теоретико-експериментальних досліджень застосовано метод оцінювання пози людини MediaPipe Pose. Для валідації та порівняння методів нормалізації координат використано 3D модель людини, створеної за допомогою програм для IOS Magic Poser та Ultimate Poser.

## Результати дослідження та їх обговорення / Research results and their discussion

**Нормалізація координат та математична модель суглоба людини.** Відповідно до офіційної документації [13], модель MediaPipe Pose Landmarker визначає ключові точки (орієнтири) пози тіла у двох форматах: нормалізованих координатах зображення (Landmarks) та координатах тривимірного світу (World Landmarks). У Landmarks (далі L) координати  $x$  та  $y$  нормалізовані між 0,0 і 1,0 за шириною і висотою зображення відповідно, на підставі ширини торсу. Координата  $z$  відповідає глибині всередині стегон як початковою точкою, чим менше значення, тим ближче орієнтир до камери; для величини  $z$  використовують приблизно той самий масштаб, що й для  $x$ . У World Landmarks (далі WL) координати  $x$ ,  $y$  та  $z$  відповідають реальним тривимірним координатам у метрах у проміжку -1,5 та 1,5, із серединою стегон як початком координат, координати отримані не за допомогою людської анотації, а шляхом додавання синтетичних даних (модель GHUM) до 2D-анотації, маски сегментації переднього/фоновому плану людини та внутрішніх параметрів камери. Модель pose landmarker повертає множину з 33 ключових точок (рис. 1).

Для покращення точності визначення кутів суглобів, запропоновано удосконалений метод нормалізації координат відносно зображення, суть якого полягає у поєднанні згаданих вище (L та WL) наборів нормалізованих координат з додатковими змінами, а саме приведення їх до спільного масштабу та початку координат, оскільки у L, точка з координатами (0; 0; 0) знаходиться у лівому верхньому куті зображення, а у WL це середина стегон.

Позначимо координати  $L$  як  $K_L(x_L; y_L; z_L)$ ; WL як  $K_{WL}(x_{WL}; y_{WL}; z_{WL})$ , а їхню множину як (формула (1)); нормалізовані 2D координати  $L$  відносно пікселів зображення  $K_{Ln}(x_{Ln}; y_{Ln})$ , де  $x_{Ln} = x_L \cdot w$ ,  $y_{Ln} = y_L \cdot h$ , множину як (формула (2)); нормалізовані координати удосконаленим методом  $K_{new}(x_{new}; y_{new}; z_{new})$ :

$$PL_{WL} = \{K_{WL0}(x_{WL0}; y_{WL0}; z_{WL0}), K_{WL1}(x_{WL1}; y_{WL1}; z_{WL1}), \dots, K_{WL32}(x_{WL32}; y_{WL32}; z_{WL32})\}; \quad (1)$$

$$PL_{Ln} = \{K_{Ln0}(x_{Ln0}; y_{Ln0}), K_{Ln1}(x_{Ln1}; y_{Ln1}), \dots, K_{Ln32}(x_{Ln32}; y_{Ln32})\} \cdot (2)$$

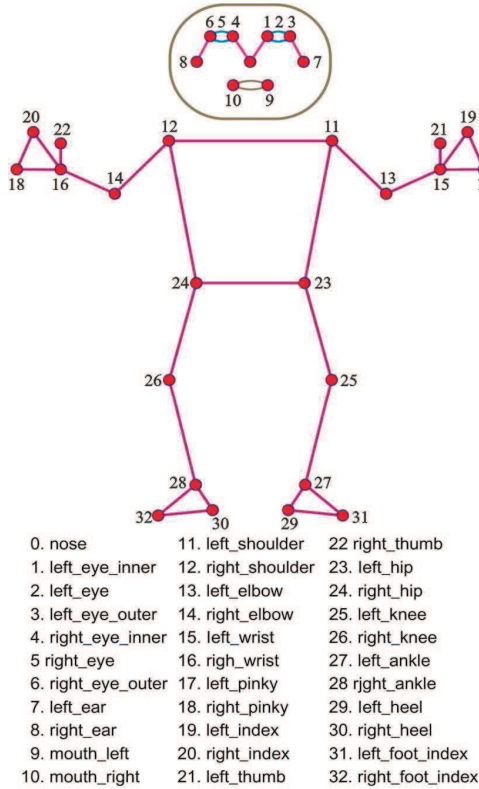


Рис. 1. Ключові точки моделі MediaPipe Pose / Key points of the MediaPipe Pose model [14]

Для початку, потрібно перевести координати  $K_{Ln}$  з пікселів до метрів, тобто координат WL. Для цього введемо коефіцієнт  $k_{p2m}$  (м/пкс), який вираховують на підставі відношення довжин відповідних сегментів тіла людини (див. рис. 1) (для нашого підходу візьмемо тільки нижні кінцівки людини), за такою формулою:

$$k_{p2m} = \sum l_{WL} / \sum l_{Ln}, \quad (3)$$

де:  $l_{Ln}$  та  $l_{WL}$  – довжини нижніх кінцівок людини, розраховані за формулою (4), де відстань між двома точками визначають на підставі  $PL_{Ln}$  та  $PL_{WL}$  відповідно:

$$l_{Ln} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

$$l_{WL} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}. \quad (4)$$

Розрахуємо нові значення координат  $K_{new}(x_{new}; y_{new}; z_{new})$ :  $x_{new} = x_{Ln} \cdot k_{p2m}$ ,  $y_{new} = y_{Ln} \cdot k_{p2m}$ ,  $z_{new} = z_{WL}$ . Тобто для координат  $x$  та  $y$  використовуємо  $K_{Ln}$ , а для координати  $z$  беремо значення з  $K_{WL}$ .

Якщо відомі реальні розміри людини, чіку позу будемо визначати, можна нормалізувати координати відносно цих значень. Для цього введемо коефіцієнт  $k_{real}$ , що буде означати "масштаб" моделі, тобто скільки в одному метрі визначеному бібліотекою MediaPipe реальних метрів. Цей коефіцієнт визначено на підставі фактичних розмірів людини, як відношення розміру кінцівки до її значення у пікселях. У цьому випадку нові значення координати матимуть такий вигляд:

$$K_{new}(x_{Ln} \cdot k_{p2m} \cdot k; y_{Ln} \cdot k_{p2m} \cdot k; z_{WL} \cdot k).$$

Суглоб людини змодельовано у вигляді точки  $O(x_o; y_o; z_o)$ , у якій з'єднані дві або більше прямих, що його формують (залежно від анатомії). Для визначення кута згинання/розгинання суглоба, потрібні координати кінців цих прямих, тобто додатково дві точки, позначимо  $A(x_A; y_A; z_A)$  та  $B(x_B; y_B; z_B)$ . Тоді для розрахунку зна-

чення кута згинання/розгинання суглоба використовуємо таку формулу:

$$\theta = \arccos \left( \frac{BO^2 + AO^2 - AB^2}{2 \cdot AO \cdot BO} \right). \quad (5)$$

**Програмне забезпечення для визначення та аналізу кінематики суглобів людини.** Для валідації та тестування запропонованого підходу нормалізації тривимірних координат розроблено програмне забезпечення, що дає можливість провести порівняння точності визначення значення кута колінного суглоба на підставі нормалізованих координат. Використано дві 3D-моделі людини у різних позах (рис. 2): у першій (1-10.png) з кутом згинання лівого коліна  $90^\circ$ , у десяти різних ракурсах; у другій (11-20.png) з кутом згинання правого коліна  $60^\circ$ .

Для нормалізації тривимірних координат виділено такі методи:

- 1) 2D: за розмірами зображення  $h$  (висота) і  $w$  (ширина). Координати точки:  $K_{2D}(x_{Ln}; y_{Ln})$ .
- 2) 3D  $w$ : за розмірами зображення  $h$  (висота) і  $w$  (ширина), з  $z$  відносно  $w$ . Координати ключової точки:  $K_{3Dw}(x_{Ln}; y_{Ln}; z_L \cdot w)$ .
- 3) 3D  $h$ : за розмірами зображення  $h$ , з  $z$  відносно  $h$ . Координати ключової точки:  $K_{3Dh}(x_{Ln}; y_{Ln}; z_L \cdot h)$ .
- 4) 3D: без нормалізації. Координати ключової точки:  $K_{3D}(x_L; y_L; z_L)$ .
- 5) WL 3D: використання WL (pose\_world\_landmarks). Координати ключової точки:  $K_{WL3D}(x_{WL}; y_{WL}; z_{WL})$ .
- 6) NEW 3D: запропонований нами удосконалений метод. Координати ключової точки:  $K_{new}(x_{new}; y_{new}; z_{new})$ .

Результати визначення та порівняння значень кута згинання суглоба наведено у табл. 1. Для кожного розрахованого значення кута суглоба  $\alpha$  вираховано абсолютну похибку  $\Delta$ , тобто відхилення значення від реального ( $\alpha_0 \in \{60^\circ, 90^\circ\}$ ,  $\Delta = |\alpha - \alpha_0|$ ). Зеленим кольором позначено клітинки, де  $\Delta \leq 10$ , а червоним – клітинки, де  $\Delta > 10$ . Рядок  $\sum(\Delta)$  підсумовує значення  $\Delta$ . Останній рядок "N" показує кількість зелених клітинок для методів нормалізації.

Аналізуючи результати тестування, можна дійти висновку, що серед шести методів нормалізації, розглянутих у дослідженні, запропонований нами новий підхід продемонстрував найвищу точність. Середня абсолютна похибка (MAE) становить  $12,6^\circ$  для першої пози, та  $7,2^\circ$  для другої, середня абсолютна відсоткова похибка (MAPE) становить 14,0 та 8 % для першої та другої поз відповідно. Незважаючи на покращення точності визначення кутів суглобів спостерігається істотна похибка у складних випадках (5, 7 та 10), де кінцівку видно погано, вона знаходиться у тіні чи ключові точки накладаються.

**Обговорення результатів дослідження.** Автори публікації [13] мали на меті підтвердити й оцінити точність BlazePose (модель машинного навчання, яку використовують у MediaPipe) для вимірювання кутів суглобів нижніх кінцівок під час ключових подій ходи та бігу на вулиці. Унаслідок виконаної роботи модель BlazePose продемонструвала сильну або дуже сильну кореляцію з ручним вимірюванням за допомогою Kinovea ( $r$ : 0,70-0,97). Діаграми Бланда-Альтмана показали хорошу точність із мінімальним зміщенням (середній MAE:  $2,11^\circ$ , MAPE: 1,57 %, RMSE:  $2,63^\circ$ ). Як висновок: BlazePose є вдалим інструментом для вимірювання кутів суглобів під час бігу, що є доступною альтернативою для біомеханічного аналізу.



Рис. 2. Тестові пози / Test poses

Удосконалити модель BlazePose намагаються автори публікації [7]. Для цього вони впроваджують моделі глибокого навчання автокодувальника Seq2seq та синхронізоване захоплення руху за допомогою RGB-камери, а також десяти камер Vicon, що дає змогу прогнозувати кінематику суглобів та отримати спільні нормалізовані координати.

Автори публікації [20] пропонують нову та ефективну архітектуру для пошуку та реконструкції 3D-пози людини на підставі двовимірних координат. Для цього вони нормалізують 3D-позу за допомогою ортографічних проєкцій та набору даних Motion Capture (MoCap),

усунувши трансляцію, орієнтацію та розбіжності розміру скелета в позах. На підставі створеної бази знань, проєктують підмножину з'єднань нормалізованих 3D-поз на площини 2D-зображень. Задачу реконструкції 3D-пози також розглядають автори публікації [5], де вони детально описали методи нормалізації 3D-пози через використання лінійних шарів, нормалізації та зважених втрат, які вони використали для покращення базової моделі оцінювання 3D-пози людини, на підставі процесу переходу від 2D до 3D з використанням двоетапного підходу. Свої зміни автори перевірили за допомогою тесту Human3.6M, та отримали задовільні результати.

Табл. 1. Порівняння методів нормалізації (поза № 1) / Comparison of normalization methods (pose No. 1)

| Ракурс           | 2D n | $\Delta$ | 3D w | $\Delta$ | 3D h | $\Delta$ | 3D   | $\Delta$ | WC 3D | $\Delta$ | NEW 3D | $\Delta$ |
|------------------|------|----------|------|----------|------|----------|------|----------|-------|----------|--------|----------|
| 1.png            | 90   | 0        | 93   | 3        | 101  | 11       | 94   | 4        | 89    | 1        | 91     | 1        |
| 2.png            | 96   | 6        | 83   | 7        | 67   | 23       | 65   | 25       | 115   | 25       | 93     | 3        |
| 3.png            | 158  | 68       | 50   | 40       | 28   | 62       | 28   | 62       | 96    | 6        | 94     | 4        |
| 4.png            | 35   | 55       | 49   | 41       | 30   | 60       | 30   | 60       | 80    | 10       | 87     | 3        |
| 5.png            | 86   | 4        | 71   | 19       | 53   | 37       | 57   | 33       | 95    | 5        | 67     | 23       |
| 6.png            | 129  | 39       | 160  | 70       | 170  | 80       | 159  | 69       | 157   | 67       | 85     | 5        |
| 7.png            | 152  | 62       | 77   | 13       | 53   | 37       | 54   | 36       | 125   | 35       | 144    | 54       |
| 8.png            | 89   | 1        | 113  | 23       | 139  | 49       | 121  | 31       | 80    | 10       | 98     | 8        |
| 9.png            | 104  | 14       | 122  | 32       | 143  | 53       | 120  | 30       | 99    | 9        | 93     | 3        |
| 10.png           | 131  | 41       | 59   | 31       | 32   | 58       | 43   | 47       | 115   | 25       | 112    | 22       |
| $\Sigma(\Delta)$ | 290  |          | 279  |          | 470  |          | 397  |          | 193   |          | 126    |          |
| N                | 4    |          | 2    |          | 0    |          | 1    |          | 6     |          | 7      |          |
| MAE              | 29,0 |          | 27,9 |          | 47,0 |          | 39,7 |          | 19,3  |          | 12,6   |          |
| MAPE, %          | 32,2 |          | 31,0 |          | 52,2 |          | 44,1 |          | 21,4  |          | 14,0   |          |

Табл. 2. Порівняння методів нормалізації (поза № 2) / Comparison of normalization methods (pose No. 2)

| Ракурс           | 2D n | $\Delta$ | 3D w | $\Delta$ | 3D h | $\Delta$ | 3D   | $\Delta$ | WC 3D | $\Delta$ | NEW 3D | $\Delta$ |
|------------------|------|----------|------|----------|------|----------|------|----------|-------|----------|--------|----------|
| 11.png           | 57   | 3        | 66   | 6        | 63   | 3        | 55   | 5        | 77    | 17       | 63     | 3        |
| 12.png           | 52   | 8        | 61   | 1        | 52   | 8        | 56   | 4        | 65    | 5        | 58     | 2        |
| 13.png           | 53   | 7        | 55   | 5        | 46   | 14       | 42   | 18       | 70    | 10       | 55     | 5        |
| 14.png           | 62   | 2        | 68   | 8        | 63   | 3        | 56   | 4        | 83    | 23       | 66     | 6        |
| 15.png           | 112  | 52       | 43   | 17       | 23   | 37       | 23   | 37       | 95    | 35       | 74     | 14       |
| 16.png           | 53   | 7        | 52   | 8        | 46   | 14       | 31   | 29       | 73    | 13       | 56     | 4        |
| 17.PNG           | 82   | 22       | 91   | 31       | 101  | 41       | 79   | 19       | 107   | 47       | 91     | 31       |
| 18.PNG           | 54   | 6        | 52   | 8        | 41   | 19       | 38   | 22       | 67    | 7        | 58     | 2        |
| 19.PNG           | 0    | 60       | 84   | 24       | 94   | 34       | 87   | 27       | 65    | 5        | 62     | 2        |
| 20.png           | 1    | 59       | 57   | 3        | 42   | 18       | 43   | 17       | 52    | 8        | 57     | 3        |
| $\Sigma(\Delta)$ | 226  |          | 111  |          | 191  |          | 182  |          | 170   |          | 72     |          |
| N                | 6    |          | 6    |          | 3    |          | 3    |          | 4     |          | 7      |          |
| MAE, °           | 22,6 |          | 11,1 |          | 19,1 |          | 18,2 |          | 17,0  |          | 7,2    |          |
| MAPE, %          | 25,1 |          | 12,3 |          | 21,2 |          | 20,2 |          | 18,9  |          | 8,0    |          |

Автори публікації [9] пропонують вирішити проблему неоднозначності глибини оцінювання 3D-пози шляхом додавання функції втрат відхилення центру маси від центру опорних ніг і штрафних функцій щодо відповідного діапазону повороту кута суглоба. Вони також розробили складнішу гуманоїдну модель та поєднали її з методом швидкої оптимізації для оцінювання кутів з'єднань для оцінювання 3D-пози.

У своїй публікації [6] автори реалізували новий підхід нормалізації даних для вирівнювання кореневого суглоба та уніфікації біомеханічних характеристик навчальних зразків, що полягає у розробленні нової моделі: біомеханічна модель пози рук (НРВМ). Запропонований нормалізований метод демонструє чудову продуктивність порівняно з моделями, навченими без нормалізованих наборів даних, з використанням ключових точок від бібліотеки MediaPipe спостерігається зменшення похибок на 11,3 і 14,3 % для наборів даних STB і RHD.

У роботі [15] розглянуто алгоритм фітнесу для виявлення та корекції постави людини. Автори навели комплексний аналіз останніх досягнень у виявленні та корекції постави людини за допомогою методів машинного навчання. Погана постава є поширеною проблемою, що сприяє різним опорно-руховим розладам та загальному дискомфорту. Традиційні методи корекції постави часто спираються на суб'єктивні оцінки або дороге спеціалізоване обладнання. Алгоритми машинного навчання стали перспективними інструментами для вирішення цих проблем шляхом автоматичного аналізу даних, пов'язаних з поставою, та надання персоналізованого зворотного зв'язку користувачам. Гібридний підхід, що поєднує згорткові нейронні мережі 3НМ (англ. *Convolutional Neural Networks*, CNNs) та рекурентні нейронні

мережі РНМ (англ. *Recurrent Neural Networks*, RNNs), використовують для аналізу як статичних, так і динамічних особливостей постави людини. ЗНМ витягують просторові ознаки з глибинних зображень, тоді як РНМ фіксують часові залежності з послідовностей даних інерційних вимірювальних одиниць ІМО (англ. *Inertial Measurement Units*, IMUs). Після виявлення неправильної постави система надає користувачеві зворотний зв'язок у режимі реального часу через візуальні, слухові або тактильні інтерфейси, спрямовуючи його до правильного вирівнювання постави. Механізми зворотного зв'язку можуть містити візуальні накладання, звукові підказки або вібраційні шаблони, адаптовані до уподобань користувача.

Автори публікації [21] вирішують проблему точності абсолютної 3D-пози людини на підставі монокулярного введення 2D-пози шляхом перетворення вхідних даних з простору пікселів у 3D-нормалізовані промені. Запропонована нова модель НРЕ Ray3D явно набуває зовнішніх параметрів камери як вхідних даних та спільно моделює розподіл між 3D-променями пози та зовнішніми параметрами камери.

Виконане обговорення результатів дослідження вказує на те, що різні науковці намагаються зробити процес реконструкції 3D-пози людини та здійснити нормалізацію координат її ключових точок з плином часу значно точнішою, проте це досі залишається для них викликом. Застосування нових архітектурних рішень та додаткового апаратного забезпечення допомагають дещо краще вирішити це актуальне завдання. Запропонований нами підхід нормалізації тривимірних координат на підставі синтезу наборів координат бібліотеки Medi-

aPіpe продемонстрував хороші результати у визначенні кутів суглобів тіла людини та може бути перспективним напрямом вирішення завдання реконструкції 3D-позы людини.

Отже, за результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

*Наукова новизна отриманих результатів дослідження* – удосконалено метод визначення кутів згинання/розгинання суглобів під час оцінювання позы людини за допомогою нормалізації тривимірних координат, що містить поєднання, перетворення та масштабування двох наборів нормалізованих координат: відносно розмірів зображення у пікселях (Landmarks) та відносно простору у метрах (WorldLandmarks).

*Практична значущість результатів дослідження* – результати дослідження реалізовано у вигляді програмного забезпечення, яке можна застосувати для вимірювання та аналізу кутів згинання/розгинання суглобів під час різної активної діяльності, наприклад ходьба, біг, їзда на велосипеді, виконання фізичних вправ, що стане корисним у сфері реабілітації після травм чи поранень, для спортивної аналітики та вирішення інших прикладних задач, що потребують точного визначення кутів суглобів людини.

## Висновки / Conclusions

Розроблено програмне забезпечення для аналізу кутів згинання/розгинання суглобів людини, розрахованих на підставі нового підходу до нормалізації тривимірних координат, визначених за допомогою бібліотеки оцінювання позы людини MediaPipe. За результатами дослідження можна зробити такі основні висновки.

1. Проаналізовано проблематику задачі визначення глибини зображення та тривимірних координат під час оцінювання позы людини на фото та відео. Встановлено, що це завдання є актуальним, тому потребує проведення нових досліджень, запропоновані рішення яких дадуть хороші результати. Розглянуто наявні рішення покращення точності реконструкції 3D-позы людини.
2. Розроблено новий підхід для нормалізації 3D-координат людини, визначених за допомогою бібліотеки HPE MediaPipe, який полягає у комбінації двох наборів ключових точок Landmarks та WorldLandmarks з додатковими перетвореннями та масштабуванням. Запропонований підхід описано у вигляді математичної моделі суглоба людини та реалізовано у вигляді програмного забезпечення, яка дає можливість визначати та аналізувати кінематику суглобів людини.
3. Розроблено модуль програмного забезпечення, що дає змогу верифікувати та протестувати новий метод нормалізації координат. Результати дослідження довели, що запропонований підхід має високу точність та ефективність у визначенні кута суглобів та реконструкції 3D-позы людини ( $MAE = [12,6^\circ; 7,2^\circ]$   $MAPE = [14,0\%; 8,0\%]$ ), проте у випадках складних поз та ракурсів похибка зростає. Майбутні дослідження можуть полягати у комбінації MediaPipe та інших інструментів визначення глибини, аналізу координат попередніх кадрів відео, використання декількох камер та ракурсів, LiDAR та лазерних давачів.

## References

1. Chung, J.-L., Ong, L.-Y., & Leow, M.-C. (2022). Comparative Analysis of Skeleton-Based Human Pose Estimation. *Future Internet*, 14(12), article ID 380. <https://doi.org/10.3390/fi14120380>

2. Clemente, C., Chambel, G., Silva, D. C. F., Montes, A. M., Pinto, J. F., & Silva, H. P. D. (2023). Feasibility of 3D Body Tracking from Monocular 2D Video Feeds in Musculoskeletal Telerehabilitation. *Sensors*, 24(1), article ID 206. <https://doi.org/10.3390/s24010206>
3. Dill, S., Ahmadi, A., Grimmer, M., Haufe, D., Rohr, M., Zhao, Y., Sharbafi, M., & Hoog Antink, C. (2024). Accuracy Evaluation of 3D Pose Reconstruction Algorithms Through Stereo Camera Information Fusion for Physical Exercises with MediaPipe Pose. *Sensors*, 24(23), article ID 7772. <https://doi.org/10.3390/s24237772>
4. Dill, S., Rösch, A., Rohr, M., Güney, G., De Witte, L., Schwartz, E., & Hoog Antink, C. (2023). Accuracy Evaluation of 3D Pose Estimation with MediaPipe Pose for Physical Exercises. *Current Directions in Biomedical Engineering*, 9(1), 563–566. <https://doi.org/10.1515/cdbme-2023-1141>
5. Gui, Z., & Luo, J. (2022). *Advanced Baseline for 3D Human Pose Estimation: A Two-Stage Approach* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2212.11344>
6. Guo, Z., Li, J., & Tan, J. (2024). A Normalization Strategy for Weakly Supervised 3D Hand Pose Estimation. *Applied Sciences*, 14(9), article ID 3578. <https://doi.org/10.3390/app14093578>
7. Hulleck, A. A., AlShehhi, A., El Rich, M., Khan, R., Katmah, R., Mohseni, M., Arjmand, N., & Khalaf, K. (2024). BlazePose-Seq2Seq: Leveraging Regular RGB Cameras for Robust Gait Assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32, 1715–1724. <https://doi.org/10.1109/TNSRE.2024.3391908>
8. Jia, S., Pei, X., Yao, W., & Wong, S. C. (2021). *Self-supervised Depth Estimation Leveraging Global Perception and Geometric Smoothness Using On-board Videos* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2106.03505>
9. Kim, J.-W., Choi, J.-Y., Ha, E.-J., & Choi, J.-H. (2023). Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model. *Applied Sciences*, 13(4), article ID 2700. <https://doi.org/10.3390/app13042700>
10. Lin, Y., Jiao, X., & Zhao, L. (2023). Detection of 3D Human Posture Based on Improved Mediapipe. *Journal of Computer and Communications*, 11(02), 102–121. <https://doi.org/10.4236/jcc.2023.112008>
11. Liqiang, Z., Qingrou, W., Xinrui, Z., & Qiming, N. (2023). *A geometric method for generating Euler angles from 3D landmark key-points*. <https://doi.org/10.22541/au.169288926.64567564/v1>
12. Masoumian, A., Rashwan, H. A., Abdulwahab, S., Cristiano, J., Asif, M. S., & Puig, D. (2023). GCNDepth: Self-supervised monocular depth estimation based on graph convolutional network. *Neurocomputing*, 517, 81–92. <https://doi.org/10.1016/j.neucom.2022.10.073>
13. MediaPipe Solutions guide. *Google for Developers*. (n.d.). (2023). URL: <https://developers.google.com/mediapipe/solutions/guide>
14. On-device, Real-time Body Pose Tracking with MediaPipe BlazePose (n.d.). (2025). URL: <https://developers.google.com/mediapipe/solutions/guide>
15. Panwar, G., Bhargava, D., & Khatter, H. (2024). Fitness Algorithm for Human Posture Detection and Correction. In: *2024 1st International Conference on Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI)*, Greater Noida, India, pp. 1–6. <https://doi.org/10.1109/ICSCAI61790.2024.10866663>
16. Rosique, F., Losilla, F., & Navarro, P. J. (2021). Deep Learning Computer Vision Algorithms for Real-time UAVs On-board Camera Image Processing. *Applied Sciences*, 11(19), article ID 9132. <https://doi.org/10.3390/app11199132>
17. Serdyuk, P. V., & Boichuk, I. V. B. (2024). Determining the correct bicycle fit using computer vision. *Scientific Bulletin of UNFU*, 34(3), 87–95. <https://doi.org/10.36930/40340311>
18. Srinivasan, H. K., Mathunny, J. J., Devaraj, A., Gogoi, H., Govindasamy, K., & Karthik, V. (2025). Can BlazePose accurately assess joint angles in outdoor running environments? *Biomedical Human Kinetics*, 17(1), 13–25. <https://doi.org/10.2478/bhk-2025-0002>

19. Van Crombrugge, I., Sels, S., Ribbens, B., Steenackers, G., Penne, R., & Vanlanduit, S. (2022). Accuracy Assessment of Joint Angles Estimated from 2D and 3D Camera Measurements. *Sensors*, 22(5), article ID 1729. <https://doi.org/10.3390/s22051729>
20. Yasin, H., & Krüger, B. (2021). An Efficient 3D Human Pose Retrieval and Reconstruction from 2D Image-Based Landmarks. *Sensors*, 21(7), article ID 2415. <https://doi.org/10.3390/s21072415>
21. Zhan, Y., Li, F., Weng, R., & Choi, W. (2022). *Ray3D: Ray-based 3D human pose estimation for monocular absolute 3D localization* (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2203.11471>

**P. V. Serdyuk, I. V. Boichuk**

*Lviv Polytechnic National University, Lviv, Ukraine*

## **NORMALISATION OF THREE-DIMENSIONAL COORDINATES IN THE MEDIAPIPE LIBRARY FOR MEASURING JOINT ANGLES**

The identification and accurate utilization of human three-dimensional coordinates and image depth remain critical challenges in human pose estimation, directly influencing the precision of movement analysis and joint angle calculation. This study proposes a novel approach for normalizing three-dimensional coordinates obtained using the MediaPipe Pose Landmarker model. The proposed method integrates human keypoint sets, specifically Landmarks (normalized image coordinates) and WorldLandmarks (normalized three-dimensional world coordinates), supplemented by specialized transformations, including scale unification and coordinate alignment to a common reference point. A mathematical model of the human joint has been developed in the course of study. To assess the effectiveness of the proposed approach and the accuracy of joint angle values derived from the normalized coordinates, a software system was designed to enable testing on videos and images containing joint flexion in complex poses from various angles and planes. The test dataset was generated using human three-dimensional pose modeling software Magic Poser and Ultimate Poser. The software-based evaluation results demonstrate an improvement in joint angle estimation accuracy. A comparative analysis with existing coordinate normalization methods confirmed the effectiveness of the proposed approach. Accuracy assessment was conducted using mean absolute error (MAE) and mean absolute percentage error (MAPE) metrics, allowing for the evaluation of deviations between computed angles and reference values. Among the five normalization methods examined in this study, the proposed approach exhibited the highest accuracy across two test datasets (MAE = [12.6°; 7.2°], MAPE = [14.0 %; 8.0 %]). Despite the observed improvement in joint angle estimation accuracy, significant errors persist in challenging scenarios where the limb is poorly visible, shadowed, or where keypoints overlap. The proposed approach can be applied in biomechanics, sports analysis, and rehabilitation. Future research directions may include integrating the MediaPipe library with other depth estimation tools, analyzing keypoint coordinates from previous frames (in the case of video sequences), utilizing multiple cameras, and employing LiDAR (Light Detection and Ranging) technology and laser sensors.

**Keywords:** image depth; human pose estimation; landmarks; joint kinematics; 3D pose reconstruction.