

ЕФЕКТИВНЕ РОЗПІЗНАВАННЯ КООРДИНАТНИХ МІТОК З ВИКОРИСТАННЯМ НАДШИРОКОКУТНОЇ КАМЕРИ ДЛЯ ІНТРАОПЕРАЦІЙНОЇ НАВІГАЦІЇ

Давидович Ілля Вікторович

магістр кафедри біомедичної кібернетики
НТУУ «КПІ ім. Ігоря Сікорського», м. Київ, Україна

Алхімова Світлана Миколаївна

к.т.н., доцент кафедри біомедичної кібернетики
НТУУ «КПІ ім. Ігоря Сікорського», м. Київ, Україна

Анотація. Розпізнавання координатних міток (маркерів) у полі зору камери надає можливість коректного розміщення віртуальних об'єктів під час використання головних дисплеїв у якості пристроїв доповненої реальності для систем інтраопераційної навігації. Поле зору камери, що використовується у головному дисплеї, обмежує робочий простір можливої взаємодії з маркерами. Додаткове обмеження на робочий простір взаємодії із маркерами накладають алгоритми розпізнавання останніх, які не спрацьовують у випадках часткового виходу маркерів із поля зору камери. Об'єктом дослідження у даній роботі є оцінка підходу до покращення розпізнавання маркерів $ARUCO$, коли положення такого маркера призводить до його часткового виходу із поля зору камери.

Ключові слова: доповнена реальність, розпізнавання маркерів, маркери $ARUCO$, точність, інтраопераційна навігація, камера.

ВСТУП

Наразі пристрої з технологією доповненої реальності почали широко використовуватися в різних професійних сферах людської діяльності, в тому числі і в медицині. Спеціальні головні дисплеї набувають все більшого попиту для різних задач у системах інтраопераційної навігації. Це відбувається за рахунок того, що їх використання не займає руки оператора і не потребує переводити зір на окремий дисплей [1]. Саме ці дві особливості є вагомими перевагами використання головних дисплеїв доповненої реальності по відношенню до традиційних систем інтраопераційної навігації або таких пристроїв для візуалізації сцен доповненої реальності як звичайні смартфони або планшети.

Головні дисплеї зазвичай використовують метод позиційного відстеження за принципом *inside-out*, тобто для орієнтування у просторі використовуються лише ті камери або датчики, що вбудовані у самий пристрій. Пристрій може відстежувати, як змінюється його положення по відношенню до навколишнього середовища. Як результат, коли оператор рухається, він змінює своє місце в просторі, а віртуальне середовище відповідно реагує в режимі реального часу. Перевагами такого методу позиційного відстеження є його простота і те, що він не потребує використання додаткового зовнішнього обладнання. Однак

недоліком є те, що не можна відстежувати взаємозв'язки між об'єктами у навколишньому середовищі. Не зважаючи на те, що можна легко розміщувати віртуальні об'єкти відносно тривимірного простору, не можливо здійснювати їх коректне позиціонування відносно об'єктів реального світу без використання спеціальних координатних міток (маркерів) [2].

Робочий простір можливої взаємодії з маркерами обмежений полем зору камери, яка використовується на пристрої доповненої реальності. Такий популярний головний дисплеї як Microsoft HoloLens 2 за інформацією від виробника має моноскопічну камеру, що надає поле зору лише у 52 градуси [3], в той час як поле зору людини без урахування повороту очей і периферійного зору становить 114 градусів [4]. У підсумку робочий простір Microsoft HoloLens 2 – це невелика зона, за межами якої віртуальні об'єкти не будуть мати коректне позиціонування по відношенню до об'єктів реального оточення. Незважаючи на наявність двох пар стереоскопічних камер по боках зазначеного пристрою, що могли б вирішити проблему невеликого поля зору моноскопічної камери, вони мають меншу роздільну здатність світлочутливої матриці у порівнянні із моноскопічною. Дана пара камер націлена на використання їх заради розрахунку відстаней до об'єктів та оцінки їх об'ємів, але погано розрізняють текстури поверхонь, а отже для задачі розпізнавання маркерів не підходять. Таким чином можна зробити висновок, що пристрій доповненої реальності має бути обладнаний камерою з великою роздільною здатністю матриці, а також камера має забезпечувати великий робочий простір для потенційної взаємодії з маркерами.

Поле зору, що відповідатиме потребам оператора, може забезпечити використання на головних дисплеях надширококутних камер.

У багатьох програмах на основі комп'ютерного зору одним із найпопулярніших підходів є використання систем бінарних маркерів ArUco [5]. Це обумовлене високою швидкістю розпізнавання та стійкістю до хибних спрацювань, до того ж бібліотека для роботи з ними [6] безкоштовна й дуже функціональна. Такі маркери є квадратними чорно-білими (бінарними) матрицями, що для досягнення контрастності додатково розміщують всередині білої зони.

Нажаль розпізнавання маркерів ArUco за допомогою вбудованих засобів із бібліотеки OpenCV можливе тільки за умови, що тіло маркера буде повністю розміщене у полі зору камери. Отже, частковий вихід маркера із поля зору камери призводить до проблем із розпізнаванням останнього на зображенні. Як результат, робочий простір для можливої взаємодії з маркерами зменшується відповідно до розмірів використовуваних маркерів.

У попередньому дослідженні [7] був запропонований підхід для добудови фрагмента маркера ArUco, який частково вийшов із поля зору камери. У результаті застосування цього підходу виконується добудова частини тіла маркера, що вийшла із поля зору камери, із його подальшим розпізнаванням вбудованими засобами із бібліотеки OpenCV. Це дозволило розширити робочий простір для потенційної взаємодії з маркерами для випадків відсутності менше ніж 43% площі маркера (Рис. 1).

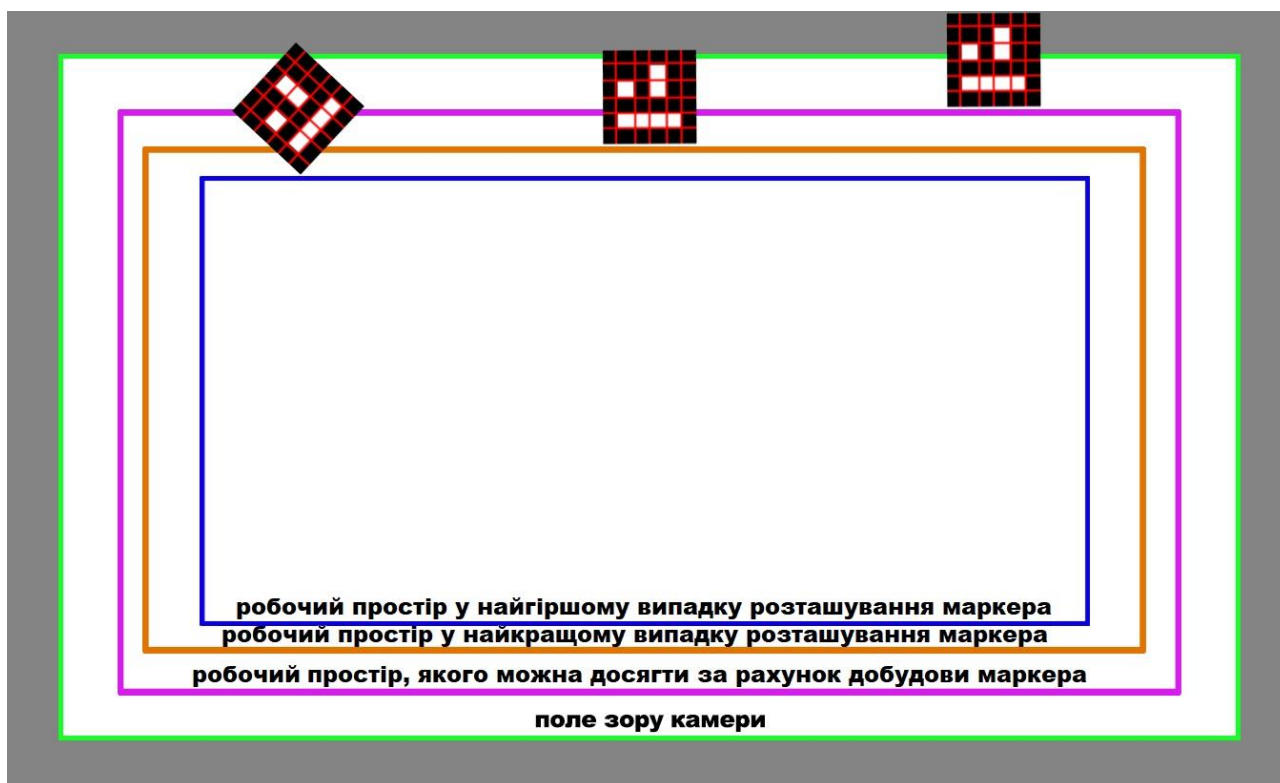


Рис. 1. Зменшення робочого простору для потенційної взаємодії з маркером ArUco під час виходу тіла маркера із поля зору камери.

Метою даної роботи є оцінка підходу до покращення розпізнавання маркерів ArUco, що був запропонований у дослідженні [7], з урахуванням різних розмірів маркерів.

МКТОДИКА ПРОВЕДЕННЯ ДОСЛІДЖЕННЯ

Бібліотека для роботи із маркерами ArUco включає в себе декілька так званих словників, що містять набори маркерів з різною кількістю бітів. Наявність меншого бітового розміру допомагає краще ідентифікувати маркери, якщо їх розмір на зображенні замалий, більший бітовий розмір сприяє точнішій ідентифікації положення у просторі.

Для перевірки можливого впливу різних розмірів й словників на стабільність розпізнавання в дослідженні були використані екземпляри маркерів словників DICT_4X4 та DICT_7X7, а саме – DICT_4X4 із розміром 3x3 см, DICT_4X4 із розміром 5x5 см і DICT_7X7 із розміром 5x5 см.

Для запису маркерів була використана надширококутна камера 8 Мп (апертура f/2.2, фокусна відстань 13 мм, поле зору 120 градусів, діагональ світлочутливої матриці 1/4,0"). Відстань від шолому з камерою до маркера складала 70 см, що є середньостатистичним значенням довжини витягнутої руки людини за даними [8].

У дослідженні [7] було показано, що стабільність алгоритмічної реалізації із побудови маркера досягається для випадків відсутності менше ніж 43% площі маркера. Таким чином, у даному дослідженні оцінка підходу до покращення розпізнавання маркерів ArUco виконувалася тільки для зазначених випадків.

Окремо було проаналізовано випадки втрати кутової та бокової частин тіла маркера до зазначеного відсотка втрати площі.

Розпізнавання маркерів ArUco виконувалося для синтетичних зображень, на яких була виконана трансляція сцени з метою зафіксувати частковий вихід тіла маркера із поля зору камери. Отримані синтетичні зображення подавалися на вхід програми, що добудовувала тіло маркера. Далі отримане зображення оброблялося вбудованими засобами із бібліотеки OpenCV. Оригінальне зображення також оброблялося вбудованими засобами із бібліотеки OpenCV і враховувалося під час аналізу разом із коефіцієнтом трансляції, що був використаний для отримання синтетичного зображення.

Оскільки зображення всього тіла маркера було відновлене алгоритмічно, в цьому дослідженні виконувалася оцінка його коректного розташування у просторі після добудови.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для оцінки якості добудови маркерів був використаний індекс структурної подібності SSIM [9]. Це простий і зручний спосіб оцінити міру схожості двох зображень. Для цього із оригінальних і синтетичних зображень були вирізані частини, що відповідали тілу маркера. Вирізані фрагменти були розташовані в центрі зображень однакового розміру із білим кольором в якості фону. В якості результату подібності були отримані значення від 0 до 1, де значення 0 відповідало повній невідповідності, а значення 1 – повній ідентичності двох зображень (Рис. 2).

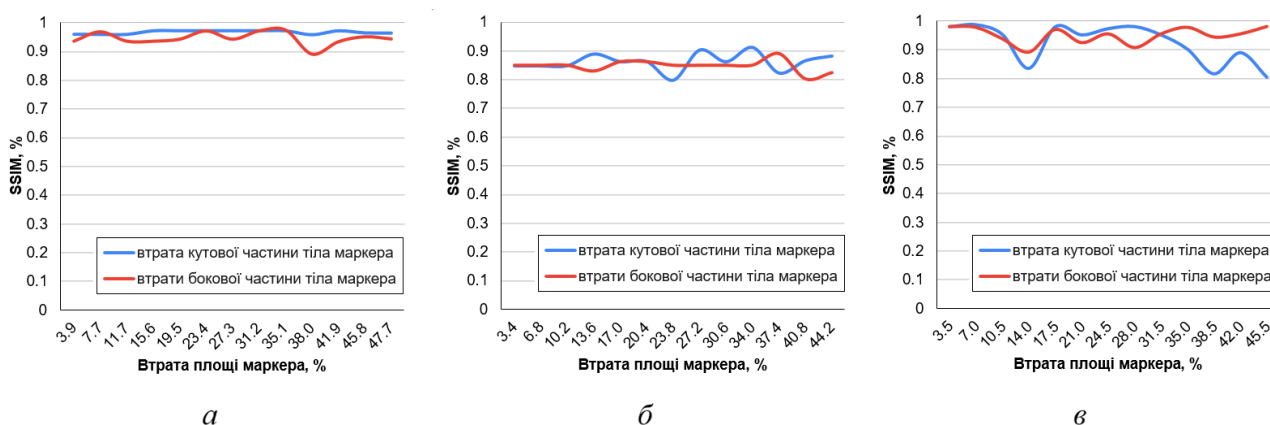


Рис. 2. Структурна подібність добудованого маркера до оригінального: *a* – для маркерів DICT_4X4 із розміром 3x3 см; *б* – для маркерів DICT_4X4 із розміром 5x5 см; *в* – для маркерів DICT_7X7 із розміром 5x5 см.

Для маркерів DICT_4X4 із розміром 3x3 см, DICT_4X4 із розміром 5x5 см і DICT_7X7 із розміром 5x5 см середнє значення індексу SSIM для випадків до 43% втрати площі маркера становило 0.956, 0.839 і 0.937, відповідно.

Отримані координати кутів розпізнаних маркерів на оригінальних і синтетичних зображеннях були використані для оцінки коректного розташування маркера у просторі. В якості метрики була розрахована евклідова відстань (Рис. 3).



Рис. 3. Усереднене зміщення координат добудованого маркера по відношенню до оригінального: *a* – для маркерів DICT_4X4 із розміром 3x3 см; *б* – для маркерів DICT_4X4 із розміром 5x5 см; *в* – для маркерів DICT_7X7 із розміром 5x5 см.

Для маркера DICT_4X4 із розміром 3x3 см результати показали, що для випадків до 43% втрати площі маркера отримані координати у середньому зміщуються на 1.14 пікселі у випадку втрати кутової і на 1.09 пікселі у випадку втрати бокової частин тіла маркера. Для маркера DICT_4X4 із розміром 5x5 см ці значення становили 2.44 пікселі і 2.38 пікселі відповідно. Для маркера DICT_7X7 із розміром 5x5 см ці значення становили 1.43 пікселі і 1.21 пікселі відповідно.

Для маркерів із однаковими фізичними розмірами (DICT_4X4 із розміром 5x5 см і DICT_7X7 із розміром 5x5 см) кращі результати були отримані для маркерів DICT_7X7 із розміром 5x5 см. Це пояснюється залежністю метода пошуку ключових точок від складності структури шуканого об'єкта. Як результат, чим більше різноманітних ліній та кутів, тим більше інформації для встановлення відповідності.

Для маркерів із різними фізичними розмірами (DICT_4X4 із розміром 3x3 см і DICT_4X4 із розміром 5x5 см) кращі результати були отримані для маркерів DICT_4X4 із розміром 3x3 см. Це пояснюється більшою щільністю пар ключових точок при менших розмірах маркера, тим самим відповідність встановлюється точніше.

ВИСНОВКИ

За результатами проведеної оцінки підходу до покращення розпізнавання маркерів ArUco можна побачити, що досліджуваний підхід краще добудовує відсутню на зображенні частину маркера для випадків втрати бокової, а не кутової частин тіла маркера. Це пояснюється специфікою візуального подання зазначених маркерів, адже при однакових відсотках втрати площі тіла маркера випадки кутової втрати призводять до більшої втрати бітового подання маркера на зображенні.

Загалом результати оцінки говорять про можливість ефективного використання запропонованого підходу з метою розпізнавання маркерів ArUco

для випадків відсутності менше ніж 43% їх площі. Отже, робочий простір потенційної взаємодії з маркерами може бути наближений до фактичних розмірів поля зору надширококутної камери, що є вкрай важливим для задач інтраопераційної навігації.

Список літератури:

1. Thabit, A., Niessen, W. J., Wolvius, E. B., & van Walsum, T. (2022, April). Evaluation of marker tracking using mono and stereo vision in Microsoft HoloLens for surgical navigation. In *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*, 12034, 253-262.
2. Langley, H. Inside-out v Outside-in: How VR tracking works, and how it's going to change. (2017). <https://www.wearable.com/vr/inside-out-vs-outside-in-vr-tracking-343>. Accessed 30 May 2022.
3. Microsoft: Learn about HoloLens 2 features and review technical specs. <https://www.microsoft.com/en-us/hololens/hardware> (2022). Accessed 30 May 2022.
4. Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford psychology series №29. Oxford University Press, USA, 736.
5. Romero-Ramirez, F. J., Muñoz-Salinas, R., & Medina-Carnicer, R. (2018). Speeded up detection of squared fiducial markers. *Image and vision Computing*, 76, 38-47.
6. OpenCV: Tutorials for contrib modules. Detection of ArUco Markers. https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html (2022). Accessed 30 May 2022.
7. Алхімова, С.М., Давидович, І.В. (2022). Розширення робочого простору до розмірів поля зору камери в системах інтраопераційної навігації. In Proceedings of the XXII International Scientific and Practical Conference «Multidisciplinary academic research, innovation and results». - Prague, Czech Republic, Jun 7-10, 2022. P.729-733.
8. Looker, J., & Garvey, T. (2015). Reaching for Holograms. In *Proceedings from International Design Congress*, 504-511.
9. Bakurov, I., Buzzelli, M., Schettini, R., Castelli, M., & Vanneschi, L. (2022). Structural similarity index (SSIM) revisited: A data-driven approach. *Expert Systems with Applications*, 189, 116087.