

УДК 004.93

Коломієць Олег Олексійович

**МЕТОДИ ІНТЕГРАЦІЇ ЕЛЕМЕНТІВ ДОПОВНЕНОЇ
РЕАЛЬНОСТІ З СИСТЕМОЮ РОЗПІЗНАВАННЯ ОБРАЗІВ****METHODS FOR INTEGRATION OF ELAMENTS
OF AGUMENTED REALITY WITH IMAGE RECOGNITIONSYSTEMS**

Дана стаття присвячена методам інтеграції елементів доповненої реальності з системою розпізнавання образів. Мета дослідження – опис алгоритмів позиціонування елементів доповненої реальності на зображенні з використанням даних отриманих від системи розпізнавання образів та обробки відео-потoku для створення доповненої реальності з використанням лінійної регресії. Для розпізнавання обличчя використано сторонню бібліотеку Vision.

Ключові слова: доповнена реальність, опорні точки, проекція, лінійна регресія, апроксимація

Бібл.: 6.

This article is focused on methods for integration elements of augmented reality with pattern recognition system. The research aims to describe the algorithms that adds elements of augmented reality to the image based on results from pattern recognition system and process video for creation of augmented reality using linear regression. Third party library Vision was used for face detection.

Keywords: augmented reality, landmarks, projection, linear regression, approximation

Bibl.: 6.

Актуальність теми дослідження. У зв'язку із значним поширенням та зростом обчислювальних потужності мобільних пристроїв для них відкриваються нові області застосування. Однією з них є створення доповненої реальності на основі зображень та відео потоку, отриманих з камери мобільного пристрою. Тому протягом останніх декількох років сфера створення елементів доповненої реальності набула значної актуальності.

Постановка проблеми. Створення елементів доповненої реальності складається із наступних основних етапів: аналіз оригінального кадру, позиціонування об'єкту, який потрібно додати та генерація вихідного кадру, шляхом накладання необхідного об'єкту на оригінальний кадр. Вважатимемо, що процес розпізнавання об'єктів відбувається з використанням сторонньої бібліотеки. Розглянемо проблему позиціонування та генерації вихідного кадру.

Постановка задачі. Мета даного дослідження – розробка алгоритмів додавання елементів доповненої реальності на зображення (на основі даних, отриманих від системи розпізнавання образів) та відео.

Поняття доповненої реальності. Доповнена реальність (розширена реальність) - результат введення у поле сприйняття будь-яких сенсорних даних (зображення, звук тощо) з метою доповнення інформації про навколишнє середовище та покращення сприйняття інформації. Доповнену реальність, іноді, називають змішаною, що створюється поєднанням реальних об'єктів сприйняття із накладеними за допомогою комп'ютера елементами.

Доповнена реальність пов'язана із розпізнавання образів, адже вона, переважно, базується на накладанні ефектів на існуюче зображення. Процес накладання потребує визначення позиції об'єкта на зображенні, для цього потрібне розпізнавання та визначення координат певних елементів на ньому.

Розпізнавання образів - це підвид машинного навчання що відображає завдання ідентифікації та розпізнавання об'єктів на зображенні або потоці відео.

В межах даного дослідження розглядатиметься накладання ефектів на обличчя людини.

Аналіз вхідних даних та розробка алгоритмів. Для розпізнавання обличчя доцільно використати одне з готових рішень OpenCV, Google Play Services Vision тощо. В контексті застосування на мобільних пристроях під управлінням ОС Android кращим варіантом є Google Play Services Vision, адже ця бібліотека є частиною пакету Google Play Services, який розроблявся спеціально для використання у Android, отже враховує особливості даної платформи. Бібліотека дозволяє розпізнати обличчя, визначити його ключові точки (англ. Landmarks): очі, основа носа, губи тощо, нахил та поворот голови.[1]

Ключовим моментом у розробці системи для додавання елементів доповненої реальності є розміщення ефекту на зображенні. Для двовимірних ефектів можна виділити такі основні параметри для позиціонування:

- Координати опорної точки ефекту на зображенні;
- Масштаб ефекту відносно зображення, що редагується;
- Поворот ефекту відносно вертикалі.

Ефекти потрібно поділити на типи, які мають однакову логіку визначення позиції, наприклад, “окуляри”, “вуса” тощо.

Для реалізації розширюваної системи у моделі збереження даних доцільно використати шаблон проектування композит (англ. Composite), який відображатиме структуру типів ефектів [2]

Позиціонування ефектів на зображенні. Розглянемо приклад визначення позиції ефекту “окуляри”. За опорну точку для даного ефекту доцільно обрати центр зображення(для фронтального вигляду окулярів). Для позиціонування використаємо такі опорні точки обличчя: ліве і праве око та основу носа.

1. Визначимо точку проекції основи носа на пряму, що проходить через точки очей, позначимо її як Projection (таким чином враховуватиметься поправка на нахил обличчя);

2. Абсцису точки розташування ефекту знайдемо за формулою

$$x = \frac{\left(\frac{leftEye.x + rightEye.x}{2} + Projection.x\right)}{2}, \quad (1)$$

де leftEye.x - абсциса точки лівого ока, rightEye.x - правого, projection.x - точки проекції;

3. Ординату обчислимо за формулою :

$$y = \frac{(leftEye.y + rightEye.y + Projection.y)}{3}, \quad (2)$$

де leftEye.y - ордината точки лівого ока, rightEye.y - правого, projection.y - точки проекції;

4. Масштаб визначимо як відношення відстані між очима до половини ширини зображення(виходячи з того, що око повинне знаходитись посередині скельця, відстанню між скельцями знехтуємо)

5. Значення нахилу голови відносно нормалі отримаємо з результатів розпізнавання обличчя.

За переліченими вище даними можна намалювати ефект поверх вхідного зображення. Також слід врахувати масштаб вхідного зображення, адже розпізнавання обличчя відбувається для його абсолютних розмірів.

Розробка алгоритму для обробки відео-потоків.

Відео-потік - це набір кадрів, які змінюються із певною частотою. Для людського ока комфортна частота зміни кадрів - 24 FPS (framepersecond), мінімальна прийнятна частота кадрів 12-4 [3]. Потреба у забезпеченні плавності відео накладає жорсткі обмеження на алгоритм додавання елементів доповненої реальності.

Основним завданням алгоритму є мінімізація часу, необхідного для обробки фреймів вхідного потоку. Зменшення часових затрат на розпізнавання можна досягнути з використанням комбінованого методу обробки вхідних даних: проводити розпізнавання лише деяких кадрів із заданою періодичністю, а для кадрів, які не обробляються системою розпізнавання образів прогнозувати розташування необхідних опорних точок за допомогою регресійного аналізу.

Поняття і обґрунтування доцільності застосування регресійного аналізу для прогнозування опорних точок об'єктів.

Регресійний аналіз - це статична техніка для дослідження і моделювання взаємозв'язків між змінними. Лінійна регресія є підвидом регресійного аналізу - це лінійний підхід до моделювання зв'язку між скалярним результатом (залежною змінною) і однією або декількома описовими(незалежними) змінними. З загальному випадку розглядається вектор регресорів (незалежних змінних). Випадок коли досліджується взаємозв'язок між результуючою змінною(y) та однією незалежною змінною(x) називається простою лінійною регресією.[4]

Об'єкти на відео рухаються за певною криволінійною траєкторією. Оскільки регресійний аналіз для невідомої криволінійної функції є складною та ресурсоємкою задачею, її не доцільно використовувати в умовах обмежених ресурсів та часу для обчислень.

Криволінійну функцію можна апроксимувати до ламаної лінії(англ. polygonal chain) [5]. Аналіз та моделювання ламаної лінії є значно простішою, з точки зору необхідних ресурсів, задачею. Її можна розбити на лінійні ділянки та досліджувати кожну із них окрему, як двовимірну лінію, задану рівнянням

$$y = \hat{a} + \hat{b}x, \quad (3)$$

Для визначення параметрів рівняння($\hat{a}, \hat{b}$) застосуємо просту лінійну регресію.[6]

$$\hat{b} = \frac{\sum_{i=1}^n x_i * y_i - n * \bar{x} * \bar{y}}{\sum_{i=1}^n x_i^2 - \bar{x}^2}, \quad (4)$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}, \quad (5)$$

де $\bar{y}, \bar{x}$ – середні вхідних даних(відомих значень координат опорних точок)

Опис алгоритму

1. Ініціалізувати список попередніх опорних точок recognizedLandmarks, граничний кут зміни вектору руху опорних точок – $d = \text{const}$, період розпізнавання кадрів – $n = \text{const}$, лічильник розпізнаних кадрів – $\text{currentIteration} = 0$

2. Початкове розпізнавання
 - 2.1. Провести розпізнавання 1 кадру, записати отримані опорні точки у масив попередніх точок
 - 2.2. Провести позиціонування ефектів та відобразити оброблений кадр
 - 2.3. Провести розпізнавання наступного, записати отримані опорні точки у масив попередніх точок
 - 2.4. Провести позиціонування ефектів та відобразити оброблений кадр
3. Для кожного наступного кадру
 - 3.1. Провести регресійний аналіз для отриманих з попередніх розпізнавань опорних точок
 - 3.2. На основі функцій руху опорних точок визначити координати опорних точок для поточного кадру
 - 3.3. Провести позиціонування ефектів та відобразити оброблений кадр
4. Паралельно із пунктом 3 для кожного n-го кадру (виконувати у паралельному потоці)
 - 4.1. `currentIteration++`
 - 4.2. Провести розпізнавання кадру, додати розпізнані опорні точки у список `recognizedLandmarks`
 - 4.3. Перевірити напрямок вектору руху опорних точок: для пар опорних точок (`recognizedLandmarks[current]`, `recognizedLandmarks [current-1]`) та (`recognizedLandmarks[current-1]`, `recognizedLandmarks[current-2]`) побудувати прямі та порівняти їх коефіцієнти нахилу – якщо різниця коефіцієнтів нахилу більша ніж d – стався перегин кривої руху об'єкту $\rightarrow$ повернутись до пункту 1; інакше продовжити

Для додавання елементів доповненої реальності на кожен фрейм, використовується описаний вище алгоритм для обробки зображень.

Висновки. Дане дослідження описує алгоритми накладання елементів доповненої реальності на зображення та відео-потік. Основне завдання описаних алгоритмів зменшити затримки при обробці відео-потіку та запобігти значному зменшенню показника плавності(кількості кадрів за секунду) вихідного відео-потіку.

Список використаної літератури

1. MobilevisionAPI [Електронний ресурс] – Режим доступу до ресурсу: <https://developers.google.com/vision/introduction>.
2. Design Patterns: Elements of Reusable Object-Oriented Software / E.Gamma, J. Vlissides, R. Johnson, R. Helm., 1994. – 395 с.
3. LaValle S. Vrtual reality / Steven LaValle. – Cambridge: Cambridge University Press, 2017. – 428 с.
4. Montgomery D. C. Introduction to linear regression analysis / D. C. Montgomery, E. A. Peck, G. G. Vining. – New Jersey: Wiley, 2012. – 642 с. – (1).
5. Hiroshi IMAI. Polygonal Approximations of a Curve — Formulations and Algorithms / Hiroshi IMAI, Masao IRI // Machine Intelligence and Pattern Recognition / Hiroshi IMAI, Masao IRI., 1988.
6. Seber G. A. Linear regression analysis / G. A. Seber, A. J. Lee. – New Jersey: WILEY, 2012. – 549 с. – (2).

ДОВІДКА ПРО АВТОРІВ

Коломієць Олег Олексійович – студент 6-го курсу кафедра обчислювальної техніки, факультету інформатики та обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kolomiets Oleh – 6th year student of computer technics department of National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Kolomiets Oleh

METHODS FOR INTEGRATION OF AUGMENTED REALITY ELEMENTS WITH IMAGE RECOGNITION SYSTEMS

The actuality of research. Augmented reality on mobile devices become very popular during the last couple of years. Mostly it related to increasing the performance of such devices. But still, there are issues with the processing of the video which requires image recognition for adding augmented reality elements. This issue is connected with the high complexity of image recognition that impacts on video FPS (frame per second) rate.

Problem setting. The high complexity of image recognition which leads to low FPS of the output video.

Target setting. Make investigation the problem of adding augmented reality element into photo and video. The aim of the research is to describe the algorithms that add elements of augmented reality to the image based on results from pattern recognition system and process video for the creation of augmented reality using linear regression.

Main material description. During this research was conducted an investigation of possible ways for improving video processing and created an algorithm which combines image recognition for getting coordinates of required landmarks on the image with regression analysis for prediction of these landmarks for frames when we are not running recognition algorithms. Linear regression is used for cases when the object is moving on the linear trajectory in another case we run additional recognition and mark all previous landmark as inappropriate for regression analysis (fallback mechanism)

Conclusions. This research describes algorithms for creation augmented reality based on results from image recognition system. The key purpose of video processing algorithms is reducing of the latency of output video and increase its FPS rate.