

К.т.н, асистент Морозов К.В., магістрант Лунгов О.В.

**Національний технічний університету України
«Київський політехнічний інститут імені Ігоря Сікорського»**

МЕТОД ВІДСТЕЖЕННЯ ПОГЛЯДУ КОРИСТУВАЧА ДЛЯ ВЗАЄМОДІЇ З КОМП'ЮТЕРОМ

Abstract

Konstantyn Morozov, assistant, PhD; Lunhov Olexandr, student
User's eye tracking method for computer interaction

This work presents a method for real-time eye tracking to enable human-computer interaction using a standard camera. The approach combines computer vision techniques, Haar features, and a convolutional neural network (CNN) trained with deep learning to predict the user's gaze coordinates on the screen. The proposed system was implemented using OpenCV, PyTorch, and PyQt, providing accurate gaze tracking and interactive control without specialized hardware. Experimental results demonstrate the method's effectiveness, accuracy, and stability under varying lighting conditions and user positions.

Вступ

Сучасні комп'ютерні системи пропонують широкий спектр засобів взаємодії з користувачем, серед яких клавіатура, миша, сенсорні екрани та інші пристрої вводу. Проте традиційні методи керування мають низку обмежень, пов'язаних зі швидкістю, ергономікою та потребою фізичного контакту з пристроєм. Це стимулює розвиток нових підходів до організації людино-комп'ютерної взаємодії.

Одним із перспективних напрямів є використання відстеження погляду як альтернативного засобу керування комп'ютером. Такий підхід забезпечує інтуїтивну взаємодію, може застосовуватися у спеціалізованих системах та підвищує доступність технологій для людей з обмеженими можливостями. У більшості сучасних рішень для відстеження погляду використовуються дорогі сенсори, що обмежує їх поширення. Тому актуальним є розроблення методу, який реалізує відстеження погляду користувача з використанням звичайної камери та алгоритмів глибокого навчання, здатних забезпечити високу точність і стабільність роботи в реальному часі.

Постановка задачі

Метою даного дослідження є розроблення методу відстеження погляду користувача для взаємодії з комп'ютером на основі методів глибокого навчання, який забезпечує роботу в реальному часі без застосування спеціалізованих сенсорів. Для досягнення цієї мети необхідно проаналізувати існуючі методи відстеження погляду та визначити їхні переваги й недоліки. Далі слід розробити архітектуру алгоритму, що поєднує засоби комп'ютерного зору та глибокого навчання для визначення положення очей і напрямку погляду користувача. Після цього потрібно реалізувати програмний прототип та провести експериментальне дослідження для оцінки точності, швидкодії й ефективності запропонованого методу в умовах реального використання.

Термінологія

Відстеження погляду (eye tracking) - процес визначення напрямку та положення очей користувача з метою інтерпретації його уваги та намірів. У даній роботі використовується для керування комп'ютером без фізичного контакту.

Комп'ютерний зір (computer vision) - область штучного інтелекту, що займається автоматичною обробкою та аналізом зображень або відеопотоку для отримання корисної інформації. У нашому методі застосовується для виявлення обличчя та очей на зображенні.

Глибоке навчання (deep learning) - підгалузь машинного навчання, що використовує багатoshарові нейронні мережі для навчання моделей на великих обсягах даних. У цій роботі глибоке навчання застосовується для передбачення координат погляду користувача на основі зображення очей.

Згорткова нейронна мережа (CNN, Convolutional Neural Network) - тип нейронної мережі, що використовується для обробки зображень шляхом виділення низько- та високорівневих ознак. CNN забезпечує точне визначення напрямку погляду на дисплеї.

Ознаки Хаара (Haar features) - набір простих прямокутних ознак, що застосовуються для швидкої детекції об'єктів на зображенні, таких як обличчя та очі. У поєднанні з каскадними класифікаторами дозволяють ефективно локалізувати ключові області на зображенні.

Аналіз існуючих підходів та алгоритмів

Існуючі методи відстеження погляду можна поділити на дві великі групи: апаратні та програмні. Апаратні методи використовують спеціалізовані сенсори, інфрачервоні камери та трекари очей для високоточного визначення напрямку погляду, проте вони є дорогими та складними у використанні. Програмні методи базуються на обробці

зображень або відеопотоку з звичайної камери та дозволяють здійснювати відстеження погляду без додаткового обладнання, проте їх точність та швидкодія зазвичай нижчі.

Сучасні дослідження активно застосовують методи глибокого навчання для підвищення точності та надійності програмних рішень. Використання нейронних мереж дозволяє ефективно визначати положення очей і напрямок погляду навіть у складних умовах освітлення та з різними позами користувача. Популярними підходами є застосування згорткових нейронних мереж для аналізу зображень обличчя та очей, а також комбінації CNN із рекурентними мережами для обробки відеопотоку в реальному часі.

Незважаючи на існування численних рішень, більшість програмних методів ще не досягли високої точності та стабільності порівняно з апаратними трекерами. Це зумовлює актуальність розробки нового методу відстеження погляду на основі глибокого навчання, який забезпечує ефективну взаємодію користувача з комп'ютером у реальному часі за допомогою звичайної камери.

Пропонований метод

Основним завданням роботи є створення методу відстеження погляду користувача, який у реальному часі передбачає напрямок погляду відносно камери та дозволяє визначати положення погляду на дисплеї. Для реалізації методу використовуються сучасні бібліотеки та фреймворки: OpenCV забезпечує зчитування відеопотоку з камери та попередню обробку зображень, включно з масштабуванням та нормалізацією кадрів; PyTorch застосовується для розробки та навчання нейронної мережі; PyQt використовується для створення графічного інтерфейсу користувача та відображення актуального стану програми, включно з позиціюванням погляду на екрані.

Метод складається з кількох основних етапів. На першому етапі здійснюється детекція обличчя та очей за допомогою ознак Хаара, що дозволяє локалізувати ключові області на зображенні і забезпечити ефективну обробку кадрів у реальному часі. Далі відбувається попередня обробка зображення, яка включає перетворення кольорового зображення у чорно-біле та масштабування до розмірів, необхідних для подальшої роботи нейронної мережі.

На наступному етапі використовується власна згорткова нейронна мережа, яка приймає оброблене зображення та передбачає координати погляду користувача. Архітектура мережі побудована на принципах ієрархічного виявлення ознак: перші шари виділяють низькорівневі характеристики, а наступні шари поєднують їх у ознаки вищого рівня, необхідні для визначення положення очей і напрямку погляду.

Процес навчання нейронної мережі реалізований за допомогою алгоритмів оптимізації градієнтного спуску. Використовується метод стохастичного градієнтного спуску та його модифікації з імпульсом, а також метод Адам, що забезпечує адаптивне оновлення ваг і стабільну збіжність моделі. Під час навчання мережа набуває здатності точно визначати положення погляду на екрані незалежно від освітлення та положення користувача відносно камери.

На фінальному етапі отримані координати погляду трансформуються у позицію курсора або вибір елементів інтерфейсу. За допомогою PyQT відбувається відображення поточної позиції погляду на екрані, що забезпечує зворотний зв'язок користувачу. Крім того, система дозволяє відстежувати кліпання очей для симуляції натискання кнопок миші, що відкриває можливості для виконання базових операцій на робочому столі.

Завдяки поєднанню комп'ютерного зору, ознак Хаара та глибокого навчання запропонований метод забезпечує високу точність і швидкодію, дозволяючи реалізувати інтуїтивну та доступну взаємодію користувача з комп'ютером у реальному часі без спеціалізованого обладнання.

Результати експериментальних досліджень

Експериментальна перевірка запропонованого методу відстеження погляду користувача проводилася з метою оцінки точності визначення координат погляду на дисплеї та швидкодії алгоритму у реальному часі. Для цього було створено тестовий прототип системи, що включав камеру, бібліотеки OpenCV, PyTorch та графічний інтерфейс на PyQT.

Для оцінки точності алгоритму було проведено серію тестів, користувач виконував стандартні завдання на екрані, зокрема наведення курсора на задані точки та симуляцію кліків очима. Результати показали, що середнє відхилення координат погляду від реальної позиції на дисплеї становило менше пікселів. Це свідчить про високий рівень точності запропонованого методу навіть без спеціалізованого обладнання.

Додатково було перевірено стабільність роботи методу за різних умов освітлення та положення користувача відносно камери. Алгоритм продемонстрував здатність до адаптації завдяки використанню глибокої нейронної мережі, що дозволяє підтримувати точність відстеження погляду при зміні кута огляду, відстані до камери та інтенсивності світла.

Результати експериментів підтверджують ефективність запропонованого методу, його практичну придатність для взаємодії користувача з комп'ютером та можливість використання без спеціалізованого обладнання.

Висновки

Отже, дана стаття була присвячена розробці методу відстеження погляду користувача для взаємодії з комп'ютером на основі методів глибокого навчання. Запропонований метод поєднує використання комп'ютерного зору, ознак Хаара та згорткових нейронних мереж, що забезпечує точне визначення координат погляду у реальному часі за допомогою звичайної камери.

Експериментальні дослідження показали, що метод дозволяє досягти високої точності визначення положення погляду на дисплеї та стабільної швидкодії, що робить його придатним для інтерактивної взаємодії з комп'ютером. Система демонструє стійкість до змін освітлення, кута огляду та відстані до камери.

Запропонований метод відкриває можливості для широкого застосування в інтерфейсах людина–комп'ютер, зокрема для користувачів з обмеженими фізичними можливостями, а також може бути використаний у навчальних, ігрових та професійних додатках. Подальший розвиток методу може включати оптимізацію архітектури нейронної мережі та розширення функціоналу для керування складнішими елементами інтерфейсу.

Література

1. OpenCV. Бібліотека обробки зображень. [Електронний ресурс] – Режим доступу: <https://opencv.org/>.
2. PyTorch. Офіційна документація. [Електронний ресурс] – Режим доступу: <https://pytorch.org/>.
3. PyQT. Офіційна документація. [Електронний ресурс] – Режим доступу: <https://riverbankcomputing.com/software/pyqt/>.
4. Viola P., Jones M. Robust Real-Time Object Detection. International Journal of Computer Vision, 2004.
5. LeCun Y., Bengio Y., Hinton G. Deep Learning. Nature, 2015, 521:436–444.