

УДК 004.5

СИСТЕМА ВИЗУАЛЬНОГО ЧЕЛОВЕКО-КОМПЬЮТЕРНОГО ВЗАИМОДЕЙСТВИЯ

Петросюк Д.В.

к.т.н., ст. преп. каф. ИС Коваленко Н.В.

Одесский Национальный Политехнический Университет, УКРАИНА

АННОТАЦИЯ. В статье рассматривается предложенная система визуального человеко-компьютерного взаимодействия на основе распознавания движений головы и мимики лица пользователя с использованием обычной веб-камеры. Приводится описание решения поставленной задачи, и результаты, достигнутые на текущий момент.

Введение. Одним из существенных ограничений существующих человеко-компьютерных интерфейсов является необходимость прикосновения к сенсорам или физическим устройствам для управления компьютером. Это создает сложности для людей, которые не могут использовать для управления компьютером привычные клавиатуру и мышь, по различным причинам, в первую очередь по состоянию здоровья. Для преодоления этого ограничения предлагается метод бесконтактного человеко-компьютерного взаимодействия основанный на визуальном управлении.

Цель работы. Целью данной работы является разработка системы отслеживания движений головы и распознавания мимики пользователя, на основе анализа изображений, полученных с помощью веб-камеры. Целевое приложение предназначено для управления курсором с помощью движений головы и осуществления щелчков мыши посредством мимики.

Основная часть работы. В настоящее время проводится довольно много исследований по созданию методов распознавания образов, позволяющих бесконтактно взаимодействовать с компьютером посредством жестов. Тем не менее, коммерческие продукты, которые способны осуществлять полное трёхмерное распознавание движений тела, мимики и направления взгляда, для отслеживания необходимых параметров используют специализированные устройства, которые характеризуются высокой ценой. Наиболее значимыми из подобных практических решений являются: *Intel RealSense, Leap Motion, Microsoft Kinect, Tobii EyeX*.

Для визуального управления курсором мыши с использованием обычной веб-камеры существуют следующие бесплатные приложения: *Camera Mouse, Head Mouse, Enable Viacam*. Их функционал и устойчивость слежения ограничены по отношению к предыдущим решениям, но они не требуют дорогостоящих дополнительных устройств (таблица 1). Движение курсора в них осуществляется движением головы, а в качестве клика мыши выступает задержка головы, на определенное время, в неподвижном положении.

Таблица 1 – Сравнительные характеристики систем визуального управления компьютером

Продукт/ Критерий	<i>Intel RealSense</i>	<i>Leap Motion</i>	<i>Microsoft Kinect</i>	<i>Tobii EyeX</i>	<i>Camera Mouse</i>	<i>Head Mouse</i>	<i>Enable Viacam</i>
Устройство слежения	спец. сенсор	спец. сенсор	спец. сенсор	спец. сенсор	обычная веб-камера	обычная веб-камера	обычная веб-камера
Управление	жесты, мимика	жесты	жесты	направление взгляда	движение головы	движение головы, моргание, открытие рта	движение головы
Трекинг объектов	+	+	+	-	+	-	-
Точность	+	++	+	+/-	+/-	+/-	+

Описанные выше бесплатные приложения при позиционировании курсора опираются только на смещение лица относительно веб-камеры. Соответственно, случайные движения головой, не приводящие к изменению ее ориентации в пространстве, также вызывают перемещение курсора, что приводит к его излишнему дрожанию. Предлагается использовать иной подход, при котором учитывается исключительно поворот головы. Таким образом система не будет реагировать на большую часть случайных движений.

Предложенная система визуального управления работает следующим образом:

- на вход системы поступает последовательность кадров, получаемых с веб-камеры.
- после запуска системы происходит этап инициализации, при котором пользователю требуется расположить параллельно веб-камере лицо с нейтральным выражением.
- в процессе работы пользователь изменяет выражение лица, система распознает его конфигурацию и отправляет соответствующую команду в компьютер.
- пользователь имеет возможность задать в соответствие к каждому управляющему событию мыши одно из доступных для распознавания системой выражений лица.
- повороты головы приводят к соответствующему движению курсора мыши.

Базовой операцией предложенной системы визуального управления является поиск лица и определение точного положение его ключевых точек на текущем кадре. Для решения задачи поиска лица применяется алгоритм *Viola-Jones* из библиотеки компьютерного зрения *OpenCV*. Данная библиотека также используется для захвата изображения с веб-камеры.

Для выделения ключевых точек лица на изображении был взят алгоритм, основанный на ограниченных локальных моделях (*CLM - Constrained Local Models*), подробно описанный в [1].

Оценка позы головы осуществляется на основании ключевых точек с использованием метода прямого линейного преобразования (*DLT - Direct Linear Transform*).

Для распознавания поднятых бровей и улыбки на лице пользователя предлагается использовать значения нормированных расстояний от ключевых точек зоны интереса лица до прямой, проходящей через внешние уголки глаз или нормали к этой прямой, для распознавания улыбки. Нормирование данных показателей осуществляется по межзрачковому расстоянию, также учитываются углы поворота головы. Найденные значения сравниваются с пороговыми, которые были получены путем анализа выборки лиц и в последующем скорректированы на этапе инициализации, а также с помощью доступных настроек. Для распознавания закрытого глаза необходимо найти соотношение расстояния между серединами верхнего и нижнего века к расстоянию между уголками глаза [2]. Для распознавания открытого рта применяется соотношение расстояний между серединами губ и уголками рта. Для обнаружения изменения выражения лица необходимо, чтобы найденные расстояния или соотношения были больше (меньше) пороговых некоторое время. В таблице 2 приведены результаты распознавания.

Таблица 2 – Результат тестирования распознавания мимики без предварительной инициализации

Ошибка	Подняты брови	Улыбка	Открыт рот	Закрыт глаз левый/правый	Головой «Да»/«Нет»
I рода	5,88%	1,16%	0,01%	5,33%	0,01%
II рода	0,70%	0,01%	0,01%	3,67%	0,01%
Распознавание	94,12%	98,84	99,99	94,67%	99,99

Выводы. Предложенная система визуального управления компьютером имеет ряд преимуществ перед аналогами. Важной особенностью данного решения является использование обычной веб-камеры, что определяет высокую доступность данной системы. На данный момент разработанный прототип системы, кроме движений головы, также способен распознать семь выражений лица пользователя, точность распознавания приведена в таблице 2. Управление курсором мыши осуществляется с помощью поворотов головы, а не перемещения, что уменьшило случайное дрожание курсора более чем в два раза по сравнению с аналогами. Возможность настроить вызов управляющего события мыши на определенный жест лица обеспечивает удобство использования данной системы. В процессе тестирования выбранный алгоритм поиска ключевых точек лица показал стабильную работу при повороте головы до 35°.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. A. Asthana, S. Zafeiriou, S. Cheng, M. Pantic. Incremental Face Alignment in the Wild. / A. Asthana, S. Zafeiriou, S. Cheng, M. Pantic // IEEE Computer vision and Pattern Recognition. – 2014.
2. Tereza Soukupová and Jan Čech. Real-Time Eye Blink Detection using Facial Landmarks. / T. Soukupová, J. Čech / In Proc. Computer Vision Winter Workshop. Rimske Toplice, – Slovenia. – February 2016.