

Система управления курсором «мыши» при помощи движений головы пользователя

Владимир Вежневц, Ростислав Шоргин, Александр Вежневц
Лаборатория Компьютерной Графики и Мультимедиа факультета ВМиК,
Московский Государственный Университет им. М.В.Ломоносова, Москва, Россия
{vvp, shorgin, avezhnevets}@graphics.cs.msu.ru

Аннотация

В настоящее время с ростом уровня использования компьютеров во всех сферах человеческой деятельности все более остро встает задача построения как можно более удобных интерфейсов человек-компьютер. Одним из основных направлений в этой области является создание естественного человеко-машинного интерфейса, позволяющего пользователю вводить информацию в компьютер наиболее привычными и естественными для него способами: с помощью речи, жестов, мимики и т.п.

В данной работе рассматривается разрабатываемая в настоящее время система, позволяющая управлять курсором «мыши» с помощью движений головы и мимики. В статье приводится описание подхода, применяемого для решения поставленной задачи, и результаты, достигнутые на текущий момент.

Ключевые слова: естественный интерфейс человек-компьютер, отслеживание лица, управление курсором «мыши»

1. ВВЕДЕНИЕ

Одним из ключевых требований, предъявляемых к современным вычислительным системам, является возможность наиболее удобного и эффективного взаимодействия человека с компьютером. В настоящее время создаются новые системы, направленные на облегчение передачи информации от пользователя к компьютеру, в частности, различные программы для альтернативного управления курсором «мыши». Такие системы могут найти применение в случаях, когда использование классического устройства типа «мышь» невозможно либо сопряжено с определенными трудностями.

Важной областью применения этих систем является их использование людьми с нарушением опорно-двигательного аппарата, не способных справиться с клавиатурой и «мышью». Другим примером возможного применения на практике таких приложений является их использование на портативных компьютерах, которые зачастую страдают от недостаточного удобных встроенных возможностей для позиционирования курсора «мыши». Наконец, подобные системы могут с успехом применяться для повышения удобства работы с компьютером обычных пользователей.

Существуют различные стратегии построения альтернативных систем позиционирования курсора «мыши». Одним из подходов является распознавание направления взгляда пользователя для указания положения курсора на экране. Для реализации этого метода необходимо высокоточное определение положения зрачков глаз, чего

можно добиться за счет подсветки лица в инфракрасном диапазоне и использования дорогостоящих камер. Однако, даже не принимая в расчет относительно высокую стоимость конечной системы, такой подход сам по себе является не очень удачным, так как согласно ряду исследований существенная часть движений глаз совершается человеком неосознанно, что делает привязку курсора к направлению взгляда нецелесообразным. Другой подход основан на том, что, в отличие от глаз, движения головы лучше поддаются сознательному контролю со стороны пользователя, что позволяет опираться на них при управлении компьютером. Немаловажной особенностью данного метода является использование в качестве дополнительного оборудования недорогой бытовой видеокамеры, что определяет высокую доступность данной системы.

2. СУЩЕСТВУЮЩИЕ СИСТЕМЫ

В течение последних лет проводились исследования, связанные с различными аспектами построения естественных человеко-машинных интерфейсов. Одной из первых систем для альтернативного управления курсором «мыши» стало приложение Nouse [1]. Получая изображение с обыкновенной бытовой видеокамеры, оно отслеживало положение носа в кадре и согласно его перемещениям позиционировало курсор. В качестве управляющего события («клика») выступало моргание пользователя. Эта программа приобрела некоторую известность, однако на практике она не нашла широкого применения, так как использовавшаяся в ней система отслеживания носа была крайне неустойчива а произвольное моргание пользователя зачастую инициировало ложные управляющие события.

На базе проводившихся исследований было создано еще несколько аналогичных коммерческих продуктов: FaceMouse [2], Cam2Pan [3], CameraMouse [4, 5], HeadMouse [9]. Параллельно разрабатывались системы, ориентированные на использование камер, работающих в инфракрасном диапазоне: QuickGlance [6], EyeGaze [7], SmartNav [8].

Описанные выше существующие программы при позиционировании курсора опираются только на положение лица относительно камеры. В результате, движения головы, сохраняющей неизменную ориентацию, в плоскости камеры вызывают перемещение курсора. Нами был выбран принципиально иной подход: учитывается исключительно ориентация (поворот) головы. Подобная стратегия применялась в наших предыдущих разработках [10], однако, в отличие от них, в новой системе удалось отказаться от использования каких-либо дополнительных приспособлений (специального шлема, маркеров и др.), прикрепляемых к пользователю.

3. ПРЕДЛАГАЕМЫЙ МЕТОД

Для создания отвечающей определенным выше требованиям системы требуется решить две подзадачи:

- отслеживание положения головы и определение ее ориентации в трехмерном пространстве;
- преобразование имеющейся информации о положении головы в управляющие сигналы «мышь».

3.1 Определение ориентации головы

Известен ряд методов определения приблизительного направления взгляда человека, основанных на определении положения антропометрических точек лица [15], [16]. Отслеживание положения антропометрических точек лица задача более сложная, нежели поиск лица на изображении, что отражается на более высоких требованиях к качеству и разрешению получаемого с камеры видео и более низкой надежности методов отслеживания.

Основная идея предлагаемого алгоритма заключается в использовании двух детекторов лица, показания одного из которых инвариантны к повороту головы вне плоскости изображения, а другого, напротив, к нему чувствительны. Таким образом, сравнивая их показания, можно делать выводы об ориентации головы.

Для определения ориентации головы в данной системе используются два детектора, хорошо зарекомендовавшие себя как по устойчивости распознавания, так и по скорости работы: Color-Based Face Tracker [11] и Viola-Jones Face Detector [13]. Оба этих алгоритма весьма нетребовательны к качеству и разрешению видео.

3.2 Отслеживание по цветовой статистике

Для отслеживания положения головы применяется алгоритм, использующий цвет кожи как признак присутствия лица. Его основные преимущества: малая вычислительная сложность (как следствие, высокая скорость работы), устойчивость к изменению масштаба и повороту лица (как в плоскости, так и вне плоскости изображения), устойчивость к изменению освещения (за исключением цветного), устойчивость к изменению выражения лица и частичному перекрытию лица другим объектом сцены.

Для обнаружения пикселей цвета кожи был выбран метод Байесовской карты вероятностей [12]. Данный метод обладает высокими показателями вероятности успешного распознавания, малой вероятностью ложного обнаружения, и одинаково успешно работает в различных цветовых пространствах. Результатом применения этого метода является *карта кожи* – полутоновое изображение, в котором интенсивность каждого пикселя пропорциональна мере близости цвета пикселя к цвету кожи (см. рис. 1).



Рис. 1. Результаты цветовой сегментации кожи. Слева – исходное изображение, справа – карта кожи

После цветовой сегментации кожи для определения положения лица применяется ранее разработанный алгоритм, использующий деформируемую эллиптическую модель [11]. Данный алгоритм позволяет значительно повысить надежность отслеживания лица по цвету кожи и в то же время не требует больших вычислительных затрат. Деформируемая модель инициализируется положением лица на предыдущем кадре и, итеративно адаптируясь к карте кожи нового кадра, определяет вектор параметров эллипса, соответствующего наиболее вероятному положению лица (результат см. рис. 2-а).

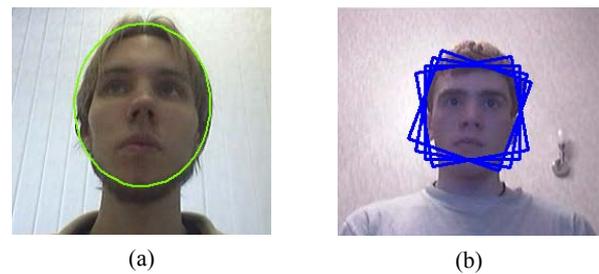


Рис. 2. Обнаружение лица с помощью различных детекторов. (а) – по цветовой статистике. (б) – детектор Viola-Jones.

3.3 Отслеживание детектором Viola-Jones

В качестве второго детектора в нашей системе применяется алгоритм обнаружения лица на черно-белом изображении Viola-Jones [13]. Нами была использована его реализация в открытой библиотеке компьютерного зрения Intel OpenCV [14]. Данный детектор лиц является наилучшим из существующих по соотношению показателей эффективности распознавания / скорость работы. Для случая, когда лицо отклонено от положения «строго ан фас» не более чем на 15° как в плоскости изображения, так и вне ее, в общем случае этот детектор обладает намного более высокой надежностью обнаружения, нежели детектор, использующий цвет кожи. Также этот метод обладает крайне низкой вероятностью ложного обнаружения лица.

Недостатком метода является его неинвариантность к повороту головы. Вероятность верного обнаружения лица по мере увеличения его отклонения от положения «строго ан фас» падает, и для отклонений более чем 15° близка к нулю. Другим недостатком детектора является относительно высокая вычислительная сложность - скорость существующих реализаций недостаточна для работы в real-time режиме на «среднем» персональном компьютере так, чтобы оставлять достаточно вычислительной мощности основным задачам пользователя.

3.4 Комбинация показаний детекторов

Для того чтобы получить устойчивую и одновременно нетребовательную к ресурсам систему отслеживания лица необходимо использовать в полной мере сильные стороны обоих детекторов и компенсировать их недостатки за счет друг друга.

3.4.1 Инициализация системы

Перед началом работы системы пользователю предлагается посмотреть прямо в камеру. На полученном изображении производится поиск лица с помощью детектора Viola-Jones. Так как поворот головы относительно положения «ан фас» на этом кадре невелик, проблем с обнаружением лица не возникает. Найденное положение лица используется для инициализации цветовой модели кожи [12].

3.4.2 Отслеживание лица на новом кадре

На каждом новом кадре производится поиск вероятного положения лица с помощью детектора на основе цветовой статистики. Это дает примерное положение лица на текущем кадре. Для уточнения положения лица используется детектор Viola-Jones.

Для стабильной работы системы требовалось добиться частичной инвариантности детектора Viola-Jones к повороту головы в плоскости изображения. Для решения этой проблемы область кадра, в которой производится поиск лица, последовательно поворачивается вокруг своего центра на 10 градусов в интервале от -30 до 30 градусов. В результате трансформации алгоритм сохраняет семь изображений, в каждом из которых производится поиск лиц с помощью детектора Viola-Jones. В сочетании с диапазоном в +/- 15° самого детектора это дает устойчивую к повороту в интервале +/- 45°. Полученные положения лица затем робастно усредняются, после чего используются в качестве информации о положении лица (см. рис. 2-b).

Для сохранения высокой скорости работы системы поиск лица на этом шаге производится не во всем кадре, а в пределах полученной от первого детектора области, взятой с некоторым запасом. Такой подход позволяет также установить и обратную связь между двумя детекторами: удачное обнаружение лица с помощью Viola-Jones сигнализирует системе об устойчивой работе первого детектора.

3.4.3 Диагностика сбоев

Детекторы при определении положения лица в кадре опираются на различные, независимые друг от друга признаки. Это позволяет надежно комбинировать результаты обнаружения лица, что существенно повышает стабильность функционирования системы. Простейший пример взаимного контроля детекторов - сильное расхождение в определении положения лица сигнализирует о некорректной работе системы.

3.4.4 Определение вектора направления головы

Алгоритм обнаружения лица, основанный на цветовой сегментации кожи, остается инвариантным относительно вращения головы во всех плоскостях. Вместе с тем, будучи (после описанных в разделе 3.4.2 модификаций) инвариантным относительно поворота головы в плоскости изображения, детектор Viola-Jones является неинвариантным относительно ее поворота вне этой плоскости. В результате, показания этих детекторов примерно согласуются при фронтальном положении лица, однако начинают расходиться при небольших поворотах головы влево-вправо и ее наклонах вверх-вниз. Это расхождение позволяет получить приблизительный вектор поворота головы (см. рис. 3).

3.5 Перемещение курсора

На этом этапе в системе уже имеется информация о векторе направления головы. Остается задача максимально эффективно преобразовать ее в управляющие сигналы «мышью».



Рис. 3. Приблизительный вектор поворота головы.

В системе предусмотрено два различных режима трансляции поворота головы в движения курсора – по скорости (аналогично манипулятору типа «джойстик») и по положению (аналогично манипулятору типа «мышь»). Первый тип управления означает, что поворот головы на определенный угол влечет перемещение курсора в соответствующем направлении со скоростью, пропорциональной углу поворота. Во втором типе управления положение курсора напрямую зависит от угла поворота головы.

Система позволяет комбинировать два вида управления в зависимости от предпочтений пользователя. На основе вычисленных координат u и v вектора поворота головы перемещение курсора мыши вычисляется по следующим формулам:

$$x_t = x_{t-1} + \alpha_x \cdot (f_d(u_t - u_0) + \int_{u_{t-1}-u_0}^{u_t-u_0} f_m(u) du)$$
$$y_t = y_{t-1} + \alpha_y \cdot (f_d(v_t - v_0) + \int_{v_{t-1}-v_0}^{v_t-v_0} f_m(v) dv)$$

В приведенных формулах (x_{t-1}, y_{t-1}) и (x_t, y_t) есть координаты курсора «мышью» соответственно в предыдущий и текущий моменты времени, (u_{t-1}, v_{t-1}) и (u_t, v_t) - координаты вектора поворота головы на предыдущем и текущем кадрах. Вектор с координатами (u_0, v_0) соответствует «нейтральному» положению головы, когда перемещение курсора не происходит. Кусочно-линейные функции $f_d()$ и $f_m()$ регулируют скорость и характер перемещения курсора в зависимости от положения головы. Коэффициенты α_x и α_y регулируют интенсивность движения по каждой из осей.

Чтобы избежать ненужного дрожания курсора, возникающего при произвольных малых движениях головы, в системе используется понятие настраиваемой «нулевой зоны» - области игнорируемых программой значений вектора поворота.

3.6 Результаты и будущие исследования

По ходу проведения исследований нами был разработан опытный прототип будущего приложения. Эксперименты по работе с ним показывают, что построить устойчивую систему на основе описанных выше идей более чем реально.

В дальнейшем, планируется развивать приложение в трех основных направлениях. Во-первых, это создание более развитой системы обнаружения ошибок и восстановления после сбоя. Второе направление - это проведение экспериментов с различными алгоритмами отслеживания лица. В том числе исследуется возможность добавления в программу модуля, отвечающего за динамическое подключение и отключение различных детекторов, исходя из условий, в которых работает система (например, освещения). Наконец, планируется дальнейшее развитие аппарата настроек приложения.

4. ЗАКЛЮЧЕНИЕ

В данной статье приведено описание разрабатываемой в настоящий момент системы, позволяющей управлять курсором «мыши» с помощью движений головы и мимики пользователя. Предложены алгоритмы для устойчивого определения ориентации лица (см. рис. 4) и ее трансляции в управляющие сигналы «мыши». Описанные методы решают поставленные задачи, однако, безусловно, требуется проведение дальнейших исследований в этой области.



Рис. 4. Определение вектора ориентации головы.

5. БИБЛИОГРАФИЯ

- [1] D.O. Gorodnichy, S. Malik and G. Roth. *Nouse 'Use Your Nose as a Mouse' - a New Technology for Hands-free Games and Interfaces*. Proceedings of International Conference on Vision Interface, pp. 354-361, Calgary, 2002.
- [2] FaceMouse. <http://www.aidalabs.com/FaceMOUSE/>.
- [3] MouseVision, Inc. <http://www.mousevision.com/>.
- [4] B. M., G. J., and F. P. *The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities*. IEEE Trans Neural Syst Rehabil Eng., 10(1):1-10, 2002.
- [5] CameraMouse, Inc. <http://www.cameramouse.com/>.
- [6] EyeTech Digital Systems. <http://www.eyetechds.com/>.
- [7] LC Technologies, Inc. <http://www.eyegaze.com/>.
- [8] NaturalPoint, Inc. <http://www.naturalpoint.com/>.
- [9] S.L. Kalman, T. Keller-Markus, A. Lorincz, and the NIP Group. *Use of Headmouse for Persons with Special Communication Needs: Preliminary Results*. Alternative and Augmented Communication, Conference, Zolyom, Sk, 2003.
- [10] Вежнев В.П. *Алгоритмы анализа изображения лица человека для построения интерфейса человек-компьютер*. Автореферат диссертации на соискание ученой степени кандидата физико-математических наук, Москва, 2004
- [11] V. Vezhnevets. *Method For Localization Of Human Faces In Color-Based Face Detectors And Trackers*. In Proc. Third

International Conference on Digital Information Processing And Control In Extreme Situations, pp. 51-56, Minsk, Belarus, May 2002.

[12] Vezhnevets V., Sazonov V., Andreeva A. *A Survey on Pixel-Based Skin Color Detection Techniques*. Proc. Graphicon-2003, pp. 85-92, Moscow, Russia, September 2003.

[13] P. Viola and M. Jones. *Robust Real-time Object Detection*. In Proc. 2nd Int'l Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing and Sampling, Vancouver, Canada, July 2001.

[14] Intel Open Source Computer Vision Library. <http://www.intel.com/technology/computing/opencv/index.htm>.

[15] P. Ballard, G.C. Stockmann: "Controlling a Computer via Facial Aspect", IEEE Trans. Sys, Man and Cybernetics, 25(4):669-677, 1995.

[16] Jilin Tu, Thomas Huang, Hai Tao, "Face as Mouse Through Visual Face Tracking" In Proceedings of CRV , pp. 339-346, 2005.

Об авторах

Владимир Вежнев - к.ф.-м.н., старший научный сотрудник ВМК МГУ. Его электронный адрес: vvp@graphics.cs.msu.ru

Ростислав Шоргин - студент ВМК МГУ. Его электронный адрес: shorgin@graphics.cs.msu.ru

Александр Вежнев - студент ВМК МГУ. Его электронный адрес: avezhnevets@graphics.cs.msu.ru

Mouse Control via User's Head Movements

The paper is devoted to the recent researches in the field of camera based human-computer interface (HCI) system developing. This software is designed primarily for the persons who are unable to use conventional means of HCI (keyboard, mouse), for example for children suffering from cerebral palsy. Moreover, while camera can be found in PC systems in increasing frequency, possibility of interaction between human and computer via visual face tracking seems to be very enticing.

We introduce a novel hands free HCI system that enables to navigate mouse cursor through head movements. The algorithm stages are described and the first results are discussed.

Keywords: *Natural Human-Computer Interface, Face Tracking, Hands Free Mouse Control.*

About the author(s)

Vladimir Veznevets, Ph.D., is a research fellow at Graphics and Media Laboratory of Moscow State Lomonosov University. His email address is vvp@graphics.cs.msu.ru.

Rostislav Shorgin is a third year student of Moscow State Lomonosov University, Faculty of Computational Mathematics and Cybernetics. His email address is shorgin@graphics.cs.msu.ru.

Alexander Vezhnevets is a fifth year student of Moscow State Lomonosov University, Faculty of Computational Mathematics and Cybernetics. His email address is avezhnevets@graphics.cs.msu.ru.