

УДК 621.307

В.Г. Абакумов, д-р техн. наук, В.В. Васильева, Е.Ю. Ломакина

Современные средства взаимодействия человека с компьютером

В работе рассматривается возможность использования жестов рук в качестве нового способа ввода информации для управления различными автоматизированными устройствами. Целью данной статьи является анализ, оценка и систематизация методов, которые используются в системах с распознаванием жестов, а также возможность дальнейшего применения.

Hand gestures are considered in this paper as a new input mean to control various automated devices. The main aim of this article is an analysis and systematization of existing solutions, estimation of gesture-recognition methods which are used in systems, as well as the possibility of further application.

Введение

В настоящее время основными стандартными устройствами ввода информации и обеспечения интерфейса человек-компьютер являются клавиатура и мышь. В других сферах, таких как робототехника или игровые приложения, где требуется трехмерная информация, используют джойстики или информационные перчатки. Все эти устройства накладывают некоторые ограничения на взаимодействие пользователя с компьютером. Для коммуникации между собой люди в основном используют зрение и речь, следовательно, взаимодействие человека с автоматизированными устройствами было бы намного удобнее с использованием распознавания видео- или аудиоинформации. Преимущество данного интерфейса состоит в том, что с его помощью можно обеспечить дистанционное управление различными устройствами. Учитывая все возможные помехи и наличие шумов в окружающей среде, предпочтение отдается системам на основе компьютерного зрения.

Одно из перспективных направлений развития информационных технологий связано с разработкой и исследованием новых способов обеспечения интерфейса человек-машина, основанных на распознавании образов. Перед разработчиками подобных интерфейсов ставится задача использования естественных для человека способов общения с компьютерами с помощью жестов, мимики. Жесты являются особенно перспективными для построения интерфейсов управления программным и аппа-

ратным обеспечением компьютеров, роботов, позволяют расширить возможности интерфейса для людей с дефектами слуха и речи.

Основная цель исследований в области распознавания жестов состоит в том, чтобы создать систему, которая может идентифицировать определенные человеческие жесты и использовать их для передачи информации или для управления различными устройствами. Результаты соответствующих фундаментальных и прикладных исследований непосредственно используются в планировании и автоматизации научных исследований, неразрушающем контроле, технической диагностике устройств, дистанционном зондировании, прогнозировании и диагностике в медицине, поиске в геологии, прогнозировании в химии, экологическом мониторинге.

Области применения

Рассмотрим конкретные примеры использования систем, где жесты служат в качестве устройств ввода информации и управления.

Разработанные и публично представленные М. Зеллером системы [1] управляют с помощью жестов программой молекулярно-динамического моделирования и осуществляют интерактивное моделирование биополимеров.

В медицине авторы [2] нашли успешное применение устройству ввода информации, разработанному на основе анализа жестов, для дистанционного просмотра изображений в радиологии. Бесконтактное управление медицинскими приборами, обеспечивает стерильность, снижает опасность проникновения инфекции, что играет немаловажную роль при лечении пациента. Системы с распознаванием жестов способны заменить сенсорные мониторы, которые используются в операционных на данный момент.

Интерактивные системы также применяются для улучшения коммуникации с глухонемыми, они интерпретируют жесты и присваивают их определенным фрагментам речи [3].

Другой важной областью применения является управление транспортными средствами [4,5]. Основной мотивацией исследований в этой области является усовершенствование системы управления дорожным движением.

Системы на основе распознавания и идентификации жестов находят свое применение и в сфере искусства. Такая система, согласно ра-

боте [6], применяется для распознавания жестов дирижера с целью понимания музыкального темпа. Она отслеживает область руки и кодирует информацию о ее движениях.

Внедрение средств робототехники в повседневную жизнь позволило реализовать управление бытовыми устройствами обычным пользователем с помощью жестов и речи [7].

Как дополнение, следует упомянуть, что не так давно компания «Toshiba» представила ноутбуки (laptops) модели Qosmio, которые используют жесты как альтернативный способ ввода информации. Такие средства посредством определенных наборов жестов перед экраном монитора позволяют работать с программными приложениями [8].

Анализ существующих тенденций в этой области [9] показывает, что в будущем управление техническими средствами с использованием распознавания жестов будет использоваться во многих приложениях. Решающим фактором, который делает более практичным и эффективным внедрение систем на основе распознавания жестов является снижение вычислительных затрат.

Технология распознавания жестов и положения руки

Человеческая рука имеет сложную анатомическую структуру, состоящую из соединенных между собой частей и суставов, которые тесно взаимосвязаны. Для того чтобы ее математически описать требуется как минимум 27 степеней свободы [10]. Разработка интерфейса для пользователя требует глубокого понимания анатомической структуры человеческой руки для того, чтобы определять какие положения и жесты руки следует выбирать для обучения системы. Существует различие между положением руки и жестами. Положение руки является статическим в отличие от жестов, где присутствует движение. Например, выставленный вперед указательный палец является статическим положением, в то время как жесты руки определяются динамическим перемещением и соответствуют последовательностям положений руки в течение некоторого промежутка времени. Исходя из этого, процесс распознавания можно разделить на два уровня: низкий - определение положения руки и высокий - распознавание жестов.

В системах распознавания на основе компьютерного зрения, движение руки фиксируется с помощью видеокамер. Полученные таким образом видеоданные анализируются, после чего осуществляется выделение необходимых компонентов (признаков). Выбор признака для последующей сегментации является важной ча-

стью процесса распознавания, так как движения рук отличаются разнообразием форм, движений и текстур. Для распознавания статического положения руки применяют геометрические признаки, такие как положения кончиков пальцев, направление пальцев, контуры рук. Такие признаки не всегда доступны и не всегда надежны из-за самозатенения и влияния условий освещения. Распознавание изменения положения руки во времени требует не наличия, как пространственных признаков, так и временных признаков. Для этого используют двумерную (2D) локализацию реперной области – области пикселей на изображении, сегментированных по определенному признаку.

Для устранения избыточной информации и выделения необходимых компонентов применяются различные типы фильтрации. Например, геометрическая фильтрация, которая используется для отделения области рук от остальных областей, соответствующих частям тела и фону. Таким образом, распознаются разные положения руки. Поскольку жесты это композиция из последовательности положений рук, то систему распознавания обучают так называемому «алфавиту» - набору положений. Распознанные жесты и соответствующие им сигналы – «решения» применяются для управления различными автоматическими устройствами.

Методы распознавания жестов и положений рук

Методы распознавания в системах компьютерного зрения можно разделить на два вида – методы на основе создания трехмерной модели руки и методы на основе выделения признаков [11].

Методы на основе создания трехмерной модели руки базируются на построении кинематической модели, учитывающей все возможные степени свободы. Для этого требуется оценить жесты руки, посредством сравнения положения руки на входном изображении и двумерной проекции модели жеста из базы данных.

Известный отслеживающий алгоритм Монте-Карло основан на использовании выборки по значимости. Идея выборки по значимости базируется на том, что некоторые значения случайных данных о движении руки в процессе моделирования имеют большую значимость для оцениваемого положения руки. Если более значимые величины будут появляться в процессе выбора случайной величины чаще, дисперсия оцениваемой функции уменьшится. Базовая методология заключается в построении распределения, способствующего выбору более значимых случайных величин. Такое распределение изменяет оцениваемую функцию и приме-

няется непосредственно в процессе расчетов. Метод эффективен при удачном выборе и построении такого распределения, так как даёт существенное сокращение времени вычислений. Недостатком данного метода является слабая помехоустойчивость и неэффективность при значительных изменениях положения рук.

Метод Стренгера [12] применяется на практике для создания трехмерной модели руки. Анатомически точная модель строится из малых квадратов (рисунок 1). Поза руки оценивается при помощи фильтра Калмана [13], который минимизирует геометрические ошибки между профилями и краями, найденными на изображении. Использование данного фильтра позволяет достигнуть большей разрешающей способности на изображении, чем, например, при фильтровании по частицам, при котором можно достигнуть большей точности. Недостатком метода является линейный рост вычислительных затрат с увеличением числа камер в системе, что значительно замедляет работу системы распознавания.

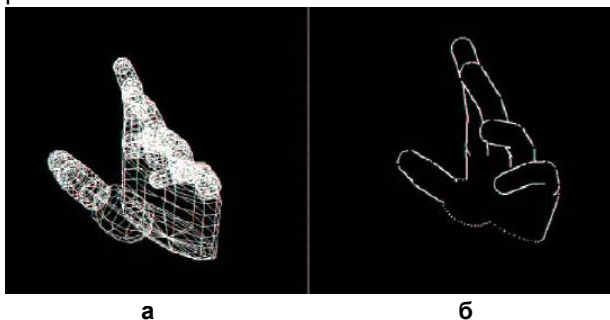


Рис. 1. Трехмерная модель (а) и ее контур (б)

Методы на основе создания трехмерной модели руки потенциально позволяют распознавать значительное количество жестов. Однако, с целью его реализации, требуется создание большой базы данных изображений для сравнения с построенной моделью и преодоление сложностей при выделении признаков с учетом анатомических особенностей.

Методы на основе выделения признаков основаны на учете особенностей изображений, которые используются для определения положения руки. Одним из подходов является нахождение участков кожи на изображении [14] с использованием цветных признаков. Способ довольно прост, но очень чувствителен к условиям освещения. В настоящее время используются методы для обнаружения участков кожи по цвету с применением управляемого освещения. Точность метода ухудшается, если на изображении присутствуют другие органы человека или объекты, имеющие подобный цвет.

Другой подход базируется на использовании подпространства, представляемого набором базисных векторов. Этот подход предполагает ортогональное построение системы векторов, соответствующей большинству признаков изображений. Метод инвариантен к входным данным, однако сопровождается сложными преобразованиями входных изображений. Для маленького набора жестов этот подход достаточен. Для большего набора жестов компактность подпространства, необходимая для эффективной обработки, не будет обеспечиваться [15].

Методы на основе выделения признаков применяются при условии, если возможно идентифицировать характерные точки или области на объектах, а сам объект может быть представлен как совокупность этих областей. В данном случае вместо моделирования объекта в целом, моделируется некое множество характерных областей. Преимущество этого подхода заключается в том, что он устойчив к деформациям и другим изменениям входных данных. При определении достаточного количества характерных признаков, объект надежно идентифицируется. Для задач определения и распознавания жестов в реальном времени данные методы являются перспективными.

Выводы

В данной статье рассматриваются методы распознавания жестов рук, которые можно разделить на две группы – методы на основе создания трехмерной модели руки и методы на основе выделения признаков.

Использование отслеживающего алгоритма Монте-Карло, основанного на использовании выборки по значимости, даёт существенное сокращение времени вычислительных операций. Однако метод не является помехоустойчивым и не эффективен при значительных изменениях положения рук.

Метод Стренгера широко используется на практике для создания трехмерной модели руки. Анатомически точная модель строится из малых квадратов. Поза руки оценивается при помощи фильтра Калмана или при помощи фильтрования по частицам. При использовании фильтра Калмана достигается большая разрешающая способность изображения, а при фильтровании по частицам достигается большая точность. Недостатком метода является линейный рост вычислительных затрат с увеличением числа камер в системе.

Метод на основе выделения признаков базируется на нахождении участков кожи на изображении. Метод прост в реализации, но чувствителен к

условиям освещения и работает не стабильно, если на изображении присутствуют другие органы человека или объекты, имеющие подобный цвет.

Подход, базирующийся на использовании подпространства, представленного набором базисных векторов инвариантен к входным данным, однако требует большого объема преобразований входного изображения. Он эффективен для маленького набора жестов.

Методы на основе выделения признаков являются стойкими к деформациям или изменениям во входных данных. Они перспективны для задач выделения рук в реальном времени.

Методы на основе трехмерной модели руки потенциально позволяют распознавать большое количество жестов. Однако для этого требуется создание базы данных большого объема.

Все вышеуказанные методы, с учетом их преимуществ и недостатков, позволяют создавать перспективные интерфейсные системы для разных задач в сфере информационных технологий.

Литература

1. Zeller, M., et al, "A Visual Computing Environment for Very Large Scale Biomolecular Modeling", Proc. IEEE Int. Conf. on Application-specific Systems, Architectures and Processors (ASAP), Zurich, pp. 3-12, 1997
2. Juan P. Wachs, Helman I. Stern, Yael Edan, Michael Gillam, Jo Handler, Craig Feied, Mark Smith, "A Gesture-based Tool for Steril Browsing of Radiology Images", Journal of the American Medical Informatics Association (2008; 15:321-323, DOI 10.1197/jamia.M24)
3. Thad Starner and Alex Pentland, "Real time American Sign Language Recognition from Video using Hidden Markov Models", Technical Report 375, MIT Media Lab, 1995
4. Dong Guo Yonghua, "Vision-Based Hand Gesture Recognition for Human-Vehicle Interaction", International Conference on Control, Automation and Computer Vision, 1998
5. Pickering, Carl A. Burnham, Keith J. Richardson, Michael J. Jaguar, "A research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction", 3rd Conference on Automotive Electronics, 2007
6. Hongmo Je Jiman Kim Daijin, "Hand Gesture Recognition to Understand Musical Conducting Action" 16th IEEE International Symposium on Robot and Human interactive Communication, 2007, RO-MAN, 2007, pg 163-168
7. A Malima, E Ozgur, M Cetin, "A fast algorithm for Vision based hand gesture recognition for robot control", 14th IEEE conference on Signal Processing and Communications Applications, April 2006
8. [Online] Available from: http://www.xbit-labs.com/news/mobile/display/2008062420361_1_Toshiba_Formally_Unveils_Notebooks_with_SpursEngine_Chip.html
9. J. K. Aggarwal, Q. Cai, "Human motion analysis: a review", Nonrigid and Articulated Motion Workshop, 1997. Proceedings., IEEE (1997), pp.90-102.
10. Y. Wu and T.S. Huang, "Hand modeling analysis and recognition for vision-based human computer interaction", IEEE Signal Processing Mag. – Special issue on Immersive Interactive Technology, vol.18, no.3, pp. 51-60, May 2001
11. H. Zhou, T.S. Huang, "Tracking articulated hand motion with Eigen dynamics analysis", In Proc. Of International Conference on Computer Vision, Vol 2, pp. 1102-1109, 2003
12. B. Stenger P. R. S. Mendonc, a R. Cipolla, "Model-Based 3D Tracking of an Articulated Hand", In proc. British Machine Vision Conference, volume I, Pages 63-72, Manchester, UK, September 2001
13. [Online] Available from: http://en.wikipedia.org/wiki/Kalman_filter
14. Elena Sanchez-Nielsen, Luis Antyn-Canalos, Mario Tejera, "Hand Getsure recognition for Human Machine Intercation", In Proc. 12th International Conference on Computer Graphics, Visualization and Computer Vision : WSCG, 2004
15. P. Viola and M. Jones, "Robust real-time object detection", Cambridge Res. Lab., Cambridge, MA, pp.1-24, Tech. Rep. CRL 2001/01, 2001.