

Список литературы

1. ПИД-регуляторы [Электронный ресурс] // энциклопедия АСУТП URL: http://www.bookasutp.ru/Chapter5_1.aspx.
2. Подогрев нефти. Печи и установки подогрева нефти [Электронный ресурс] // URL: http://www.intech-gmbh.ru/oil_preheaters.php.
3. Денисенко В. ПИД-регуляторы: вопросы реализации // В записную книжку инженера.

УДК 004

РЕАЛИЗАЦИЯ РАСПОЗНАВАНИЯ ЖЕСТОВ В РЕАЛЬНОМ ВРЕМЕНИ

А.В. Байэр

*Научный руководитель: Ю.А. Болотова, к.т.н., доцент каф. ВТ ИК ТПУ
Национальный исследовательский Томский политехнический университет
634000 г. Томск, пр. Ленина, 30
E-mail: anton.bauer@inbox.ru*

Abstracts. *This paper describes a realization of method of Human-Computer interaction based on using video-stream. Described method applies Viola-Jones algorithm and Hidden Markov models for gestures capturing and recognition.*

Keywords: HCI, gestures, Viola-Jones, Hog-descriptors, HMM, gestures database.

Ключевые слова: Человеко-машинное взаимодействие, жесты, алгоритм Виолы-Джонса, HOG-дескрипторы, база жестов.

Ведется разработка прототипа программы, которая позволяет управлять компьютером с помощью жестов.

Проект состоит из 3-х модулей:

- Захват жеста
- Преобразование последовательности изображений
- Обработка жеста (обучение либо классификация).

Для захвата жеста используется алгоритм Виолы-Джонса [1]. Началом жеста считается появление в кадре раскрытой ладони.



Рис. 1. Ладонь в кадре

Жест считается законченным в тот момент, когда на текущем кадре не будет обнаружено раскрытой ладони.

Для представления изображения в виде последовательности чисел используются HOG-дескрипторы [2].

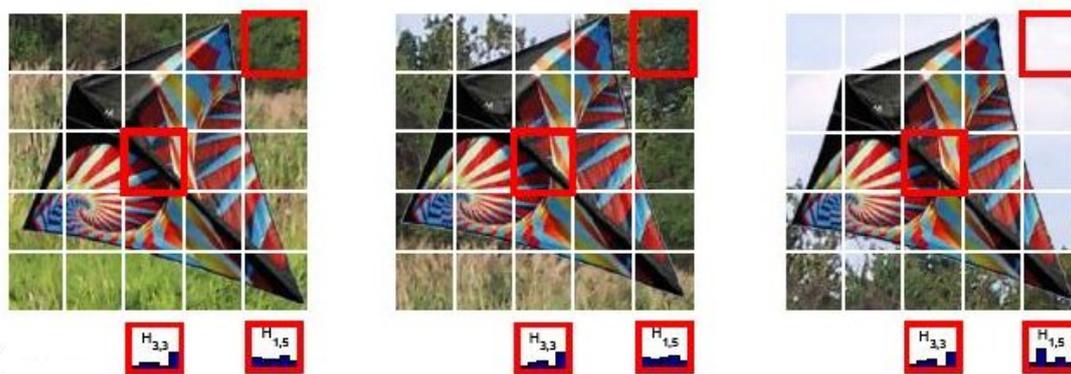


Рис. 2. HOG-дескрипторы различных изображений

В качестве классификатора используются скрытые Марковские модели. Для каждого жеста создается Марковская модель, которая обучается при помощи алгоритма Баума-Велша [3]. Классификация жеста основана на алгоритме прямого-обратного хода [3].

В табл. 1 представлены входные и выходные данные для модулей.

Таблица 1

Входные и выходные данные модулей приложения

Модуль	Входные данные	Выходные данные
Захват жеста	Видеопоток с камеры устройства	Последовательность ключевых кадров жеста
Преобразование жеста	Последовательность кадров	Последовательность HOG-дескрипторов каждого кадра
Классификация жеста	Числовая последовательность	Класс, к которому относится жест

Рассмотрим детали реализации каждого из модулей.

Модуль захвата изображения получает данные из видеопотока с веб-камеры. Используя загруженный каскад классификаторов для алгоритма Виолы-Джонса определяем, что в кадре есть ладонь. Следующий пункт – определение, нужно ли добавить текущий кадр в последовательность ключевых кадров. Для этого используется метод `cvCompareHist` из библиотеки `EmguCV`. Указанный метод сравнивает две гистограммы методом, указанным в качестве параметра. Если разница между текущим кадром и последним ключевым кадром больше некоторого порога, то текущий кадр добавляется в качестве очередного ключевого кадра.

После окончания жеста каждый кадр из последовательности ключевых кадров преобразуется в последовательность чисел следующим образом: для кадра рассчитывается HOG-дескриптор, после чего к дескриптору добавляются смещения по вертикали и горизонтали относительно предыдущего кадра. Полученная последовательность чисел нормируется.

Последний этап – классификация жеста – использует скрытые Марковские модели. Для каждого жеста обучаем свою модель. Полученную из ключевых кадров последовательность чисел подаем на все сети, и, с помощью алгоритма прямого-обратного хода, определяем вероятность того, что сеть возвращает переданную последовательность наблюдений. После этого сравниваем вероятности, которые вернули имеющиеся сети. Выбираем класс, с которым связана сеть, вернувшая наибольшую вероятность – это и будет класс, к которому относится захваченный жест.

Проект основан на платформе `.NET Framework`, написан на языке `C#`. Прототип представляет собой десктопное приложение. Для построения графического пользовательского интерфейса используется технология `WPF (Windows Presentation Foundation)`.

Для обработки изображений использовалась библиотека компьютерного зрения EmguCV – кроссплатформенная .NET обертка над OpenCV. Позволяет вызывать методы OpenCV в .NET-совместимых языках. Может использоваться для разработки мобильных приложений, в том числе тех, которые используют Xamarin. Таким образом проект может быть собран для всех популярных на сегодняшний день мобильных платформ.

Список литературы

1. P. Viola and M.J. Jones, «Robust real-time face detection», International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154.
2. Navteem Dalal and Bill Triggs, “Object Detection using Histogram of Oriented Gradients”, Pascal VOC 2006 Workshop ECCV, 2006.
3. Sushnita Mitra and Tinku Acharya, “Gesture Recognition: A Survey”, IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews, Vol. 37, No 3, May 2007.

УДК 004

РАСПОЗНАВАНИЕ ОБРАЗА ЧЕЛОВЕКА НА ВИДЕО

А.А. Анисимова

*Научный руководитель: И.И. Савенко, ассистент каф. АиКС ИК ТПУ
Национальный исследовательский Томский политехнический университет
E-mail: anisimovanastyaa@gmail.com*

Abstracts. *This article rassmotrena problem of pattern recognition. The basic methods of recognition. Also considered the principle of the touch controller Kinect. It is worth noting that the Kinect is Microsoft's breakthrough in the field of pattern recognition.*

Key words: pattern, recognition, Microsoft, Kinect, computer vision.

Ключевые слова: распознавание, образ, Майкрософт, Кинект, компьютерное зрение.

В настоящее время задача распознавания образов становится наиболее актуальной и находит все большее количество применений. Некоторыми из областей применения распознавания образов являются видеонаблюдение на дорогах, охранное видеонаблюдение, системы военного назначения, идентификация и аутентификация. Исследователями используется большое количество ресурсов для решения проблемы автоматического распознавания объектов.

Существует три основных метода распознавания образов. Первый из них, метод перебора. При использовании этого метода происходит сравнение объекта с базой данных, в которой хранятся различные модификации отображения этого объекта. Таким образом, если речь идет об оптическом распознавании, то в базе данных отображения этого объекта могут храниться под разными углами, в разных масштабах и так далее. Второй метод – глубокий анализ характеристик образа (при оптическом распознавании это, например, определение геометрических характеристик). Третий метод – искусственные нейронные сети. Не смотря на то, что метод требует использование большого количества примеров, при обучении, он отличается более высокой производительностью и эффективностью.

Говоря о распознавании образа человека на видео, следует отметить устройство, разработанное компанией *Microsoft*, бесконтактный сенсорный игровой контроллер *Kinect*.

Kinect является внушительным достижением в области распознавания. Его устройство основано на стандартном распознавании образов. Следует отметить большую многоядерную вычислительную мощность, которая позволяет сделать обучающее множество достаточно