

МИНИСТЕРСТВО
ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное
образовательное учреждение высшего
профессионального образования
«Сибирский государственный
аэрокосмический университет
имени академика М. Ф. Решетнева»
(СибГАУ)
просп. им. газеты «Красноярский рабочий», 31
г. Красноярск, 660014
Тел.: (391) 264-00-14. Факс: (391) 264-47-09
E-mail: info@sibsau.ru
http://www.sibsau.ru
ОКПО 02069734, ОГРН 1022402056038
ИНН/КПП 2462003320/246201001

«УТВЕРЖДАЮ»

Проректор по НИД,
д.ф.-м.н., профессор



«28» февраля 2014 г.

04 MAR 2014

№

41/1088

На №

от

ОТЗЫВ

ведущей организации на диссертационную работу Нгуен Тоан Тханг «Алгоритмы распознавания жестов на видеопоследовательностях», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Актуальность темы диссертации

В процессе эволюции пользовательского интерфейса устройства человеко-машинного взаимодействия стали более удобными и естественными для пользователя. С введением новых понятий, таких как «виртуальная среда», «человеко-машинная интеллектуальная интеракция», «перцепционный пользовательский интерфейс» требуется разработка более мощных и удобных способов взаимодействия человека с компьютером. Таким образом, взаимодействие на основе естественных жестов человека доказало свою перспективность и возможность широкого применения. В связи с этим, распознавание жестов стало важным направлением исследования и привлекает множество исследователей.

В настоящее время распознавание жестов на основе компьютерного зрения становится доминантной тенденцией. Перспективность данного направления исследования подтверждается результатами многих авторов.

Однако, существующие подходы еще не до конца решают проблемы распознавания жестов в реальном времени. Повышение скорости, точности, надежности и количества распознаваемых жестов и обеспечение возможности работы создаваемых алгоритмов в реальных условиях применения представляют собой сложные задачи и требуют проведения дальнейших исследований.

Таким образом, разработка надежного, точного и высокоскоростного алгоритма распознавания жестов в реальном времени представляет собой актуальную задачу.

Содержание работы

Во *введении* обоснована актуальность темы, сформулирована цель, поставлены задачи диссертационного исследования. Рассмотрены вопросы научной и практической значимости проведенных исследований и изложены основные положения, выносимые на защиту.

В *первой главе* приведено понятие жеста и интерфейса на основе жестов. Проведен аналитический обзор методов распознавания жестов на основе алгоритмов компьютерного зрения и выявлены достоинства и недостатки существующих решений. Сделан вывод об актуальности темы диссертационной работы, поставлена цель работы, сформулированы задачи, необходимые для решения проблемы распознавания жестов в видеопотоке в режиме реального времени.

Во *второй главе* приведено детальное описание разработанного алгоритма распознавания поз руки (*hand postures*), созданного алгоритма распознавания траектории глобального движения руки (*hand motion*) и алгоритма распознавания жестов (*hand gestures*). Алгоритм распознавания поз руки построен на основе SURF-дескрипторов, метода *k*-средних и нейронной сети. В алгоритме распознавания траектории глобального движения руки применяется нейронная сеть, способ передискретизации и алгоритм упрощения траектории. В алгоритме распознавания жестов используется метод Джонса-Виолы для обнаружения руки на видеокадре, метод *CAMShift* для слежения за рукой в последовательных кадрах и сочетается возможность распознавания автономных и интерактивных жестов.

В *третьей главе* представлены результаты тестирования разработанного алгоритма распознавания поз на созданных выборках, базе жестов Кембриджского университета и базе изображений Себастьяна

Марселя. Сделан вывод о точности, скорости и робастности созданного алгоритма распознавания поз руки.

В главе приведены результаты тестирования алгоритма распознавания движения руки в реальных условиях применения и сделан вывод о работоспособности алгоритма.

В *четвертой главе* приведено детальное описание разработанной программной системы распознавания жестов с целью демонстрации созданных алгоритмов. Изложено описание классов и главных модулей программной системы: детектор, трекер, модуль для распознавания формы руки, модуль для распознавания движения руки, механизм создания, обучения и тестирования нейронной сети. Рассмотрены все функции программы, изучен процесс работы с системой и продемонстрирована работа программы.

Научная новизна полученных результатов

Разработан новый алгоритм распознавания поз руки (*hand posture*) на основе *SURF*-дескрипторов, алгоритма *k*-средних и многослойной нейронной сети. Алгоритм отличается от других способностью функционировать в реальном времени, устойчивостью к различным аффинным преобразованиям, изменению освещения, и, частично, к шумам, при обеспечении точности распознавания в пределах 90-98%.

Создан новый алгоритм распознавания движения руки (*hand motion*) в видеопотоке на основе нейронной сети. Алгоритм предназначен для распознавания динамической компоненты жестов и обеспечивает возможность распознавания сложных деформированных траекторий с точностью выше 96% в реальных условиях применения.

Разработан новый алгоритм распознавания жестов (*hand gesture*) на основе применения детектора Джонса-Виолы, трекера *CAM-Shift*, предложенных алгоритмов распознавания поз и движения руки. Алгоритм сочетает возможности распознавания интерактивных и автономных жестов в режиме реального времени благодаря разбиению жестов на статическую компоненту (позу) и динамическую компоненту (движение руки).

Обоснованность и достоверность результатов и выводов диссертации

Обоснованность и достоверность результатов диссертации подтверждается согласованностью расчетных и экспериментальных данных, корректностью выполненных математических выкладок.

Результаты диссертационной работы прошли апробацию на международных научных конференциях и семинарах.

Значимость для науки и практики

Научную ценность работы представляет вклад в развитие области распознавания объектов и человеко-машинного взаимодействия. В диссертации разработан новый алгоритм распознавания поз руки, позволяющий распознавать статический компонент жестов с высокой точностью в реальном времени при обеспечении устойчивости к разным типам аффинных преобразований объектов, и частично, к шумам. Созданный алгоритм распознавания движения руки вместе с алгоритмом распознавания формы руки дает полноценное описание жестов руки человека для цели управления компьютером. Разработанный в работе алгоритм распознавания жестов позволяет создавать высокоэффективные интерфейсы на основе жестов для управления компьютерной системой, оборудованной веб-камерой.

Реализованная в ходе выполнения диссертации программа *Hand Recognitor* обеспечивает управление презентациями, навигацию веб-браузера, рисование, управление *Windows media center* с использованием жестов.

Рекомендации по использованию результатов диссертации

Полученные в диссертации результаты могут использоваться в научных исследованиях по распознаванию объектов, обработке изображений и человеко-машинному взаимодействию.

Разработанные в диссертации методические, алгоритмические и информационные средства предназначены для использования в системах удаленного управления компьютером с помощью жестов, управления компьютерной игрой, навигации виртуальной среды.

Результаты диссертации Нгуен Тоан Тханг следует довести до сведения следующих организаций: ИСА РАН, ИВМ СО РАН, МГТУ им. Баумана, НГТУ, СФУ.

Замечания по работе

1. В обзоре не рассматриваются работы российских авторов.
2. Во второй главе нечетко сформулировано обоснование выбора метода *CAMShift* в качестве трекера. Существуют более мощные методы трекинга.
3. В диссертации не исследуется влияние размерности *SURF*-дескрипторов на результаты работы алгоритма распознавания поз руки.
4. В диссертации имеются стилистические, синтаксические и грамматические ошибки.

Приведенные недостатки не снижают общий уровень научной работы и значимость полученных в ней результатов.

Оценка диссертационной работы в целом

В целом диссертация Нгуен Тоан Тханг выполнена на высоком научном уровне и представляет собой самостоятельную законченную научно-квалификационную работу, которая имеет важное научное и практическое значение в области распознавания объектов, человеко-машинного взаимодействия и компьютерного зрения.

Основные результаты работы опубликованы в открытой печати, в том числе, в рецензируемых научных изданиях, рекомендованных ВАК РФ (4 статьи в журналах), докладывались на международных конференциях и обсуждались на научных семинарах.

В автореферате представлены основные этапы работы, полученные результаты и сформулированные выводы. Автореферат отражает основное содержание диссертации.

Оформление диссертации и автореферата удовлетворяет требованиям соответствующих нормативно-методических документов.

Диссертация полностью удовлетворяет требованиям «Постановления Правительства РФ от 24.09.2013 № 842 “О порядке присуждения ученых степеней”», предъявляемым к кандидатским диссертациям по техническим наукам, а её автор Нгуен Тоан Тханг заслуживает присуждения ему учёной степени кандидата технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Отзыв обсуждён и одобрен на заседании научно-технического семинара кафедры системного анализа и исследования операций Сибирского

Профессор кафедры САИО,
д.т.н., профессор,
Заслуженный работник
высшей школы РФ

0,

Подпись Астамошкин А.В.
Ученый секретарь СибГАУ
г. Красноярск