

**Заключение диссертационного совета Д 212.267.08 на базе
федерального государственного бюджетного образовательного учреждения
высшего профессионального образования
«Национальный исследовательский Томский государственный университет»
Министерства образования и науки Российской Федерации
по диссертации на соискание ученой степени кандидата наук**

аттестационное дело № _____

решение диссертационного совета от 27.03.2014 г., № 135

О присуждении **Нгуен Тоан Тханг**, гражданину Социалистической республики Вьетнам, ученой степени кандидата технических наук.

Диссертация **«Алгоритмы распознавания жестов на видеопоследовательностях»** по специальности **05.13.11** – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, принята к защите 22.01.2014 г., протокол № 132, диссертационным советом Д 212.267.08 на базе федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Национальный исследовательский Томский государственный университет» Министерства образования и науки Российской Федерации (634050, г. Томск, пр. Ленина, 36, приказ о создании диссертационного совета № 937-648 от 23.05.2008 г.).

Соискатель **Нгуен Тоан Тханг**, 1983 года рождения.

В 2008 году соискатель окончил государственное образовательное учреждение высшего профессионального образования «Томский политехнический университет».

В 2014 году соискатель очно окончил аспирантуру федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Национальный исследовательский Томский политехнический университет».

В настоящее время не работает.

Диссертация выполнена на кафедре вычислительной техники федерального государственного бюджетного образовательного учреждения высшего

профессионального образования «Национальный исследовательский Томский политехнический университет» Министерства образования и науки Российской Федерации.

Научный руководитель – д-р техн. наук, **Спицын Владимир Григорьевич**, федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Национальный исследовательский Томский политехнический университет», кафедра вычислительной техники, профессор.

Официальные оппоненты:

Кориков Анатолий Михайлович, д-р техн. наук, проф., федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Томский государственный университет систем управления и радиоэлектроники», кафедра автоматизированных систем управления, заведующий кафедрой

Протасов Константин Тихонович, д-р техн. наук, ст. науч. сотр., федеральное государственное бюджетное учреждение науки Институт оптики атмосферы им. В.Е. Зуева Сибирского отделения Российской академии наук, лаборатория распространения оптических сигналов, старший научный сотрудник
дали положительные отзывы на диссертацию.

Ведущая организация – федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «**Сибирский государственный аэрокосмический университет имени академика М.Ф. Решетнева**», г. Красноярск, в своём положительном заключении, подписанном **Антамошкиным Александром Николаевичем** – д-р техн. наук, проф., кафедра системного анализа и исследования операций, профессор, указала, что исследование распознавания жестов на основе компьютерного зрения в настоящее время становится доминантной тенденцией. Работа выполнена на высоком научном уровне, имеет важное научное и практическое значение. Автором получены новые научные результаты, обоснованность и достоверность которых подтверждается согласованностью расчетных и экспериментальных данных, корректностью выполненных математических выкладок. Полученные

результаты могут использоваться в научных исследованиях по распознаванию объектов, обработке изображений и человеко-машинному взаимодействию. Методические, алгоритмические и информационные средства предназначены для использования в системах удаленного управления компьютером с помощью жестов, управления компьютерной игрой, навигации виртуальной среды.

Соискатель имеет 9 опубликованных работ, в том числе по теме диссертации – 9 работ, опубликованных в рецензируемых научных изданиях – 4, статей в научных журналах – 2, свидетельство на программу для ЭВМ – 1, публикаций в материалах всероссийских и международных конференций – 2. Общий объем работ – 2,93 п.л., авторский вклад – 1,95 п.л.

Наиболее значимые научные работы по теме диссертации:

1. Нгуен Т.Т. Алгоритмическое и программное обеспечение для распознавания фигур с помощью Фурье-дескрипторов и нейронной сети // Известия Томского политехнического университета. – 2010. – Т. 317, № 5. – С. 122–125. – 0,26 п.л.

2. Нгуен Т.Т., Спицын В.Г. Распознавание формы руки на видеопоследовательности в режиме реального времени на основе SURF-дескрипторов и нейронной сети // Электромагнитные волны и электронные системы. – 2012. – Т. 16, № 7. – С. 31–39. – 0,8 / 0,4 п.л.

3. Нгуен Т.Т., Болотова Ю.А., Спицын В.Г. Распознавание жестов на видеопоследовательностях в режиме реального времени на основе иерархически-временной сети // Научный вестник Новосибирского государственного технического университета. – 2012. – №. 2. – С. 33–42. – 0,5 / 0,2 п.л.

4. Нгуен Т.Т., Спицын В.Г. Алгоритмическое и программное обеспечение для распознавания формы руки в реальном времени с использованием SURF-дескрипторов и нейронной сети // Известия Томского политехнического университета. – 2012. – Т. 320, №. 5. – С. 48–54. – 0,56 / 0,28 п.л.

На автореферат поступили 7 положительных отзывов. Отзывы представили:
1) **В.Г. Редько**, д-р физ.-мат. наук, заместитель руководителя Центра оптико-нейронных технологий НИИ системных исследований РАН, г. Москва, *без замечаний*. 2) **Г.С. Плесневич**, канд. физ.-мат. наук, профессор кафедры

прикладной математики Национального исследовательского университета «МЭИ», г. Москва, *с замечанием*: не показано, как с помощью разработанного алгоритма распознавания жестов построить интерфейс на основе жестов. 3) **М.Г. Гриф**, д-р техн. наук, проф., зав. кафедрой автоматизированных систем управления Новосибирского государственного технического университета, *с замечаниями*: в автореферате недостаточное внимание уделено описанию разработанного программного обеспечения и результатам практического внедрения; представляет интерес оценка применения разработанного математического обеспечения для задач распознавания жестовой речи глухих. 4) **А.В. Петров**, д-р техн. наук, профессор Иркутского государственного технического университета *с замечаниями*: не приведено описание архитектуры разработанной программной системы; не указан способ выбор параметров нейронных сетей (в алгоритме распознавания поз и алгоритме распознавания движения). 5) **В.А. Эттель**, канд. техн. наук, профессор кафедры «Информационные технологии и безопасность» Карагандинского государственного технического университета, Республика Казахстан, *с замечаниями*: нет обоснования выбора порогового значения для определения «чужих» признаков; термины «признак», «вектор» и «дескриптор» используются бессистемно. 6) **С.В. Панин**, д-р техн. наук, проф., зам. директора по научной работе, и **П.С. Любутин**, канд. техн. наук, младший научный сотрудник лаборатории полимерных композиционных материалов (Институт физики прочности и материаловедения СО РАН, г. Томск), *с замечаниями*: недостаточно полно описаны методики тестирования алгоритмов (в части полученных результатов желательно было привести примеры кадров, для которых была бы проиллюстрирована выявленная траектория движения руки, в особенности в условиях сложного фона); изображения тестовых выборок недостаточно информативны (в частности, не видно явных отличий выборки на рис. 4 и рис. 5). 7) **Д.М. Сонькин**, канд. техн. наук, и. о. директора ООО «ИНКОМ», г. Томск, *с замечаниями*: в автореферате не приведены схемы разработанных алгоритмов; в результатах тестирования алгоритмов явно не представлены ошибки первого и второго рода.

Выбор официальных оппонентов и ведущей организации обосновывается тем, что: А.М. Кориков является признанным специалистом в области отслеживания объектов и распознавания образов; К.Т Протасов является опытным специалистом по обработке изображений и распознаванию объектов на изображениях; Сибирский государственный аэрокосмический университет имени академика М.Ф. Решетнева – один из ведущих научно-исследовательских центров России, в котором работают высококвалифицированные специалисты в области обработки изображений, распознавания образов и компьютерного зрения.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

Предложен алгоритм распознавания поз руки (*hand posture*) на основе SURF-дескрипторов, алгоритма k -средних и многослойной нейронной сети, предназначенный для распознавания статической компоненты жестов и обладающий способностью функционировать в режиме реального времени, устойчивостью к различным аффинным преобразованиям, изменению освещения, и, частично, к шумам, при обеспечении точности распознавания в пределах 90-98%. Отличительной особенностью алгоритма является сокращение времени обучения, высокая скорость и точность распознавания при отсутствии требования к наличию большой базы данных для обучения.

Предложен алгоритм распознавания движения руки (*hand motion*) в видеопотоке на основе нейронной сети, предназначенный для распознавания динамической компоненты жестов в режиме реального времени. В основе алгоритма лежит идея упрощения и передискретизации траектории, полученной после трекинга, что обеспечивает возможность распознавания сложных деформированных траекторий с точностью выше 96% в реальных условиях применения.

Разработан новый алгоритм распознавания жестов (*hand gesture*) на основе детектора Джонса-Виолы, трекера *CAM-Shift*, предложенных алгоритмов распознавания поз и движения руки, позволяющий распознавать жесты на видеопоследовательностях в режиме реального времени. Особенностью

предложенного алгоритма является сочетание возможности распознавания интерактивных и автономных жестов благодаря разбиению жестов на статическую компоненту (позу) и динамическую компоненту (движение руки). Одно из достоинств алгоритма состоит в том, что он обеспечивает высокую скорость работы, независимо от размера видеокадра.

Теоретическая значимость исследования обоснована тем, что: в работе предложен способ представления признакового пространства на основе сочетания SURF-детектора и метода k-средних, который позволяет создать вектор с фиксированной длиной, применимый для различных алгоритмов классификации.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

разработанный алгоритм распознавания жестов позволяет создавать высокоэффективные интерфейсы на основе жестов для управления компьютерной системой, оборудованной веб-камерой;

реализованная программа *Hand Recognitor* обеспечивает управление презентациями, навигацию веб-браузера, рисование, управление *Windows Media Center* с использованием жестов.

Рекомендации об использовании результатов диссертационного исследования. Результаты диссертационной работы могут быть использованы организациями, занимающимися созданием дружественных интерфейсов для человеко-машинных компьютерных систем, организациями высшего профессионального образования, 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 управления компьютером с помощью жестов, управления компьютерной игрой, навигации виртуальной среды.

Оценка достоверности и новизны результатов исследования выявила:

показана высокая точность и робастность разработанных алгоритмов при использовании в реальных условиях;

теория подтверждается строгим применением математического аппарата нейронной сети, численных методов, методов компьютерного зрения и обработки изображений;

установлено качественное и количественное совпадение авторских результатов с ранее известными результатами.

Все результаты, полученные автором диссертации, **являются новыми**.

Личный вклад соискателя состоит в доказательстве и обосновании полученных в диссертации результатов, математических выкладках, численных расчетах, подготовке публикаций по выполненной работе и личном участии в апробации результатов исследования.

Диссертация соответствует п. 9 Положения о присуждении ученых степеней, является научно-квалификационной работой, в которой содержится решение задачи распознавания жестов в видеопотоке в реальном времени, имеющей существенное значение для развития методов компьютерного зрения и распознавания образов, создания и повышения эффективности разрабатываемых интерфейсов взаимодействия с компьютером на основе жестов.

На заседании 27.03.2014 г. диссертационный совет принял решение присудить **Нгуен Тоан Тханг** учёную степень кандидата технических наук.

При проведении тайного голосования диссертационный совет в количестве 15 человек, из них 7 докторов наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, участвовавших в заседании, из 21 человек, входящих в состав совета, проголосовали: за – 14, против – 1, недействительных бюллетеней – нет.

Председатель
диссертационного совета

Ученый секретарь
диссертационного совета



Поддубный
Поддубный
Василий Васильевич

Скворцов
Скворцов
Алексей Владимирович

27 марта 2014 г.