

Қолданылған әдебиеттер тізімі

1. Власов, К. П. Методы исследований и организация экспериментов / К. П. Власов, П. К. Власов, А. А. Киселёва. – Х. : Гуманитарный центр, 2002. – 256 с.
2. Норенков, И. П. Информационные технологии в образовании / И. П. Норенков, А. М. Зимин. – М. : Изд-во МГТУ им. Н. Э. Баумана, 2004. – 352 с.
3. Редактор электронных курсов CourseLab/ – <http://www.courselab.ru/>
4. Евланов А. Г. Теория и практика принятия решений. — М.: Экономика, 2004. — 175 бет.
5. С.К.Құрманбаев, Г.Н.Гамарник, О.С.Сұлтанов «Муниципальный менеджмент» Алматы, 2000 ж.
6. Миронов А.А., Мордвинов В.А., Скуратов А.К. Семантико-энтропийное управление OLAP и модели интеграции xOLAP в SemanticNET (ONTONET). Информатизация образования и науки №2, 2009. С. 2130.

УДК 004.93, 004.89

МЕТОДЫ РАСПОЗНАВАНИЯ ЖЕСТОВ РУК

Калымова Кульзия Акрашевна

gulzia_kalymova@mail.ru

Докторант специальности «6D070400 – Вычислительная техника и программное обеспечение» факультета информационных технологий

Евразийский Национальный университет им. Л.Н. Гумилева, г. Нур-Султан, Казахстан
Научный руководитель – Д.Ж. Сатыбалдина

Успешные исследования по распознаванию жестов рук в течение последних двух десятилетий привели к системам естественного взаимодействия человека с компьютером. Но все еще имеются такие нерешенные проблемы, как точная идентификация фаз (этапов) жеста, чувствительность к изменениям размера, формы и скорости, а также проблемы, связанные с окклюзией. В связи с этим в настоящей работе рассмотрены количественные и качественные сравнительные характеристики алгоритмов распознавания жестов с использованием RGB и RGB-D камер. Количественное сравнение подходов выполняется с использованием набора показателей, выбранных из различных свойств метода и экспериментальной методологии, принятой для оценки алгоритма. Необходимость рассмотрения данных показателей совместно с точностью распознавания жестов позволяет сделать выводы о преимуществах тех или иных методик с целью их дальнейшего использования в реальных приложениях. Представлены также результаты собственных исследований по распознаванию статических жестов рук с использованием камеры глубины и сверточной нейронной сети.

Невербальная коммуникация, которая включает в себя общение через жесты рук, положение тела и выражения лица (мимику), составляет около двух третьей части всего общения между людьми. Жесты руками являются одной из наиболее распространенных категорий языка тела, используемых при общении и взаимодействии людей. В то время как остальные части тела указывают на более общее эмоциональное состояние, жесты рук могут иметь определенное лингвистическое содержание [1]. Благодаря скорости и выразительности во взаимодействии жесты рук широко используются в языках жестов и в системах взаимодействия человека с компьютером.

Для начала следует разделить два понятия - **поза руки** и **жест**. Под *позой* руки понимается ее статическое положение. *Жестом* называют телодвижение, преимущественно движение рукой, сопровождающее речь или имеющее *значение* какого-либо сигнала, знака. Язык жестов в широком смысле бесконечно разнообразен как по форме жестов, так и по их семантике, поэтому полномасштабное *распознавание* языка жестов представляет собой сложную задачу даже для человеческого интеллекта.

На современном уровне развития систем компьютерного зрения задача автоматического распознавания обычно ставится применительно к небольшому набору заранее определенных жестов [2]. Существует множество различных способов распознавания языка жестов, которые можно разделить на две группы: контактные (на основе прямого контакта) и бесконтактные (компьютерное зрение), которые будут описаны в дальнейшем.

Кроме того, имеется несколько способов классификации жестов:

- 1) на основе наблюдаемых особенностей»
- 2) на основе интерпретации.

В первую категорию классифицируются два типа жестов на основе временных отношений: статические и динамические жесты (рис. 1).

Статические жесты (так называемые позы рук) - это те, в которых положение рук не меняется в течение периода жестов. Статические жесты в основном зависят от формы и углов изгиба пальцев. В динамических жестах рук положение руки постоянно меняется по отношению ко времени. Динамические жесты обычно имеют три этапа движения: подготовка, ход (stroke) и возврат (retraction).

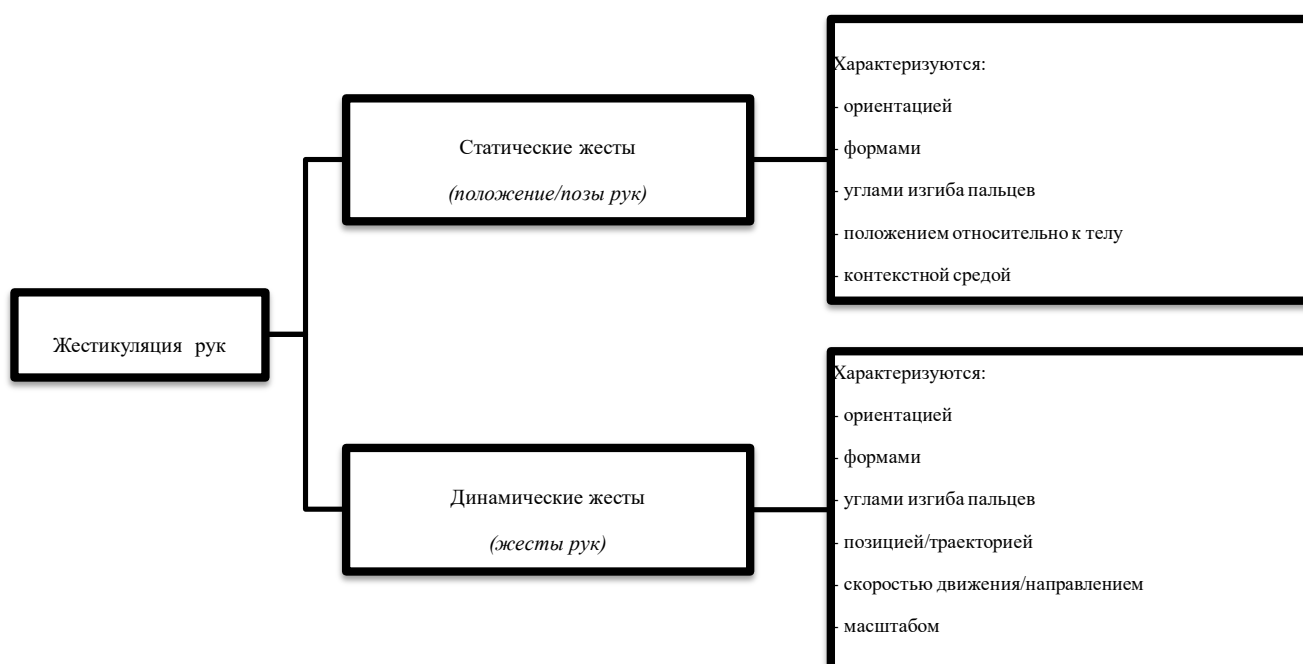


Рисунок 1. Классификация жестов рук, основанная на временных отношениях

Информация в динамическом жесте в основном содержится во временной последовательности на этапе движения. Динамические жесты основываются на траекториях и ориентациях рук, а также на формах и углах изгиба пальцев.

Во вторую категорию классифицируются жесты на основе интерпретируемого значения. Например, эмблемы, иллюстрации, регуляторы, отображение эффектов и адаптеры являются типичными классами для описания жестов [3].

На рис. 2 показана блок-схема типичной системы распознавания бесконтактного жеста. В системе распознавания жестов на основе зрения сенсор представлен в виде камеры. Берман и соавт. [4] рассмотрели различные сенсоры, используемые в системах распознавания жестов, и предоставили всесторонний анализ интеграции датчиков в системы распознавания жестов и влияние их на производительность.

Системы распознавания жестов на основе зрения разделены на две категории (рис.3.):

- методы, основанные на извлечении признаков по внешнему виду жеста;
- методы, основанные на извлечении признаков по трехмерной модели руки (3D).

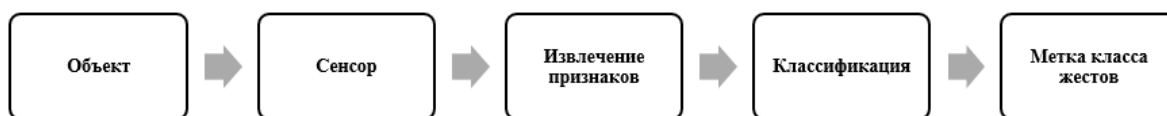


Рисунок 2. Блок-схема системы распознавания жестов

Особенностью методов, основанных на внешнем виде, является анализ только внешнего вида (формы, позиции и т.д.) целевого объекта. Для распознавания не хранится никакой информации о физических свойствах рассматриваемого объекта. Данные методы по распознаванию жестов руки человека на основе анализа внешних признаков жеста рассмотрены в работе [5].



Рисунок 3. Обзор жестов на основе зрительного восприятия (Bourke et al. 2007) [6]

Особенностью методов распознавания жестов на основе модели является сравнение проекций трехмерной модели руки с входными изображениями. По входному двумерному изображению руки, на основе разных гипотез модель руки приводится в определенную конфигурацию. Трехмерная модель руки проецируется на плоскость, тем самым, получается двумерное изображение, которое сравнивается с входным изображением. Если расстояние между изображениями меньше заданного порога, то жест распознан, в противном случае конфигурация модели руки меняется и заново сравнивается с входным изображением. Гипотезы, используемые в данных системах разные, но в большинстве случаев предполагается, что конфигурация руки в каждом кадре видеоряда совпадает или незначительно отличается от найденной конфигурации руки в предыдущем кадре.

В данной работе более подробно остановимся на методах второй категории, которые используются в компьютерном зрении для распознавания детальной трехмерной конфигурации руки при наличии на входе одного или нескольких изображений жеста, полученных с использованием двух видов сенсоров: RGB камеры и камеры глубины, или RGB-D. Видео потоки с камер глубины подобны цветным, за исключением того, что каждый пиксель имеет значение, представляющее расстояние от камеры, а не информацию о цвете [7].

Глубинные камеры уже несколько лет используются в компьютерном зрении. Однако применение камер глубины было ограничено из-за его высокой цены и низкого качества. Выпуск недорогой камеры глубины цвета (RGB-D) Kinect [8,9] от Microsoft произвела революцию в распознавании жестов, предоставляя высококачественные изображения глубины, решая такие проблемы, как сложные фоны и изменение освещенности. Устройство рассчитывает трехмерную карту сцены, используя комбинацию RGB и ИК камеры. В работе Хан и соавт. [10] предоставляется обзор того, как Kinect полезен для решения фундаментальных проблем в компьютерном зрении. Датчики, такие как Microsoft Kinect (R) и ASUS Xtion PRO LIVE (R) обеспечивают надежное отслеживание положения тела человека в игровых сценариях. На основе отслеживания эти устройства предоставляют такие функции, как координаты скелетной модели, которые используются для распознавания жестов.

Данные скелета от этих датчиков RGB-D должны быть преобразованы в более значимые и высокоуровневые объекты, а также разработаны алгоритмы для надежной классификации этих объектов (жестов). Распознавание жестов рук особенно сложны из-за сложной артикуляции и относительно небольшой площади области кисти. Кроме того, надежный алгоритм распознавания жестов рук должен быть инвариантным по отношению к размеру ладони и скорости жеста, а также ориентации жеста. Рафаэль и другие [11] оценили влияние информации о глубине в процессе распознавания жестов и пришли к выводу, что использование силуэтов глубины значительно увеличивает точность распознавания. Доминио и соавт. [12] предложил алгоритм для объединения нескольких дескрипторов на основе глубины для распознавания жестов рук.

В таблице 1 представлены сравнительные характеристики нескольких опубликованных систем распознавания жестов рук с использованием камер глубины.

Таблица 1 - Сравнение методов распознавания жестов на основе использования камер глубины

Работа	Статический (S) или динамический (D) жест	Сенсор	Признаки распознавания	Метод классификации	Точность распознавания, %
[13]	D	CSEM Swissranger SR-2	Примитивы движений	Классификатор вероятностного редактирования расстояния	92,9
[14]	S	ToF и RGB камеры	Трехмерные модели положений рук	Метод ближайших соседей (Nearest Neighbors)	99,07
[15]	S	Kinect	Формы рук и пальцев	Сопоставление шаблонов с использованием FEMD (finger earth mover's distance)	93,9
[16]	D	Kinect	Изображение истории движений (Motion History Image)	Метод на основе вычисления максимума коэффициента корреляции	Не сообщается
[17]	S, D	Kinect	Значения пикселей глубины	Классификация, основанная на принципах случайного леса	91, 84
[18]	S	RealSense SR300	Значения пикселей глубины и пикселей цвета	Сверточная нейронная сеть, содержащая два входных канала: цветные изображения и изображения глубины	99,4

Как видно, из таблицы 1, использованные в работах методы выделения признаков и классификации жестов обеспечивают высокую точность распознавания как статических, так и динамических жестов. Следует отметить, что в большинстве работ получены эффективность

систем распознавания оценивается по небольшому набору жестов, что ограничивает их массовое применение. Это подтверждается тем фактом, что имеется множество разработок с открытым исходным кодом для распознавания лица, рта и глаз (OpenCV), но нет надежных детекторов рук. Поэтому остается актуальной задача совершенствования методов распознавания статических и динамических жестов рук на фотоизображении или в видео потоке.

В связи с этим, нами был предложен подхода для распознавания жестов, основанного на комбинированном использовании глубинного сенсора Intel RealSense D435 для захвата жеста и предварительно обученной сверточной нейронной сети с архитектурой VGG-16 для выделения признаков и классификации объектов. Для программной реализации системы распознавания жестов на языке программирования Python использованы библиотеки RealSense от компании Intel, OpenCV и DL - фреймворки с открытым исходным кодом Keras и TensorFlow.

Мы подготовили базу данных, которая содержит изображения с сегментированными жестами, представленными на рисунке 4. Камеру глубины Intel RealSense D400 разместили на столе, участники эксперимента представляли жесты, в положении сидя, располагая руку перед сенсором на расстояние от 10 до 75 см перед ним (см. Рисунок 5).



Рисунок 4. Образцы жестов

Наш набор данных имеет в общей сложности 2000 изображений, в том числе 1000 изображений RGB и 1000 карт глубины, собранных при разных фонах в нескольких комнатах с изменениями освещенности. Чтобы увеличить разнообразие базы данных, мы также попытались увеличить количество изображений с помощью вариации расстояния до сенсора, углов наклона, поворота ладони и т.д.



Рисунок 5. Пример захвата жеста руки человека сенсором RealSense D400

В таблице 2 представлена матрица средних вероятностей классификаций, полученных с использованием предложенной системы распознавания статических жестов на стадии

тестирования. Как видно, для всех жестов значение элементов главной диагонали составляет более 97%, и лишь небольшой процент выборок определяется как принадлежащий другим жестам (менее 0,2%). Это указывает на то, что предложенный подход имеет высокую производительность как по показателю *Precision_c* (точность), так по *Recall_c* (полнота). (см. последний столбец и последнюю строку в Таблице 2). Исключением является распознавание жеста «Palm», где ошибка классификации превышает 2,5%, что можно объяснить определенным сходством этого жеста с жестом «Peace». Учитывая, что результирующая точность классификатора рассчитывается как среднее арифметическое его точности по всем классам, производительность предлагаемого подхода по точности распознавания получена на уровне 0,9935, или $\approx 99,4\%$.

Таблица 2 - Матрица ошибок для задачи распознавания статических жестов рук

Фактический жест	Решение классификатора					<i>Precision_c</i>
	First	L	Ok	Palm	Peace	
First	99,9798	0,0018	0,0131	0,005	0,0003	0,9998
L	0	99,9978	0,0021	0	0,0001	1,0000
Ok	0,0008	0,0255	99,9726	0	0,0011	0,9997
Palm	0,0194	0,0159	0,0151	97,4405	2,5091	0,9744
Peace	0	0,0004	0,0032	0,6245	99,3719	0,9937
<i>Recall_c</i>	0,9998	0,9996	0,9997	0,9936	0,9754	

Жесты уже давно считаются техникой взаимодействия, которая потенциально может предоставить более естественные, творческие и интуитивно понятные методы для общения с компьютерами, роботами. Эта статья предоставляет анализ сравнительных исследований сделанных в этой области. Также в работе представлена система распознавания статических жестов рук, которая использует пиксели цвета и глубины от сенсора Intel® Real Sense™ D435. Уникальные возможности камеры глубины использованы, чтобы сегментировать жесты рук и удалить фон на изображениях с жестом. Выделение признаков и классификация жестов были выполнены с использованием глубоких сверточных нейронных сетей архитектуры VGG, предварительно обученных на базе данных ImageNet. Полученные оценки распознавания на этапе тестирования для 1000 образцов жестов подтверждают эффективность предварительно представленной структуры распознавания и указывают на возможности камеры глубины D435 для будущих приложений на основе человеко-компьютерного взаимодействия.

Однако распознавания статических жестов рук недостаточно для эффективных систем взаимодействия людей и компьютерных систем [19]. Поэтому будущие исследования связаны с разработкой методов распознавания динамических жестов рук, а также и расширением базы данных и классов объектов.

Список использованных источников

1. Pramod Kumar Pisharady, M. Saerbeck. Recent methods and databases in vision-based hand gesture recognition: A review // Computer Vision and Image Understanding, Volume 141, December 2015, p. 152-165.
2. Куракин А.В. Распознавание динамических жестов на основе медиального представления формы изображений // МГУ им. Ломоносова.- Москва, 2014. – 26 с.
3. A. Kendon, Gesture and speech: how they interact, in: John M. Wiemann, Randall P. Harrison (Eds.), Nonverbal Interaction, Sage Publications, Beverly Hills, 1983.
4. S. Berman, H. Stern, Sensors for gesture recognition systems // IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. 42 (3) (2012) 277–290.
5. Нагапетян В.Э. Методы распознавания жестов руки на основе анализа дальностных изображений [Текст]: автореф. дис. на соиск. учен. степ. канд. физ.-мат. наук (05.13.17) / Нагапетян Ваагн Эдвардович; РУДН. – Москва, 2013. – 108 с.

6. A. Bourke, J. O'Brien, G. Lyons. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & Posture* 26(2) 2007:194–199.
7. M. Carfagni, R. Furferi, L. Governi, C. Santarelli, M. Servi, F. Ucheddu, Y. Volpe. Metrological and critical characterization of the Intel D415 stereo depth camera // *Sensors*. – 2019. – Т. 19. – №. 3. – p. 489-508.
8. Z. Zhang. Microsoft kinect sensor and its effect, *IEEE MultiMed.* 19 (2) (2012) 04–10.
9. J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, A. Blake, Real-time human pose recognition in parts from single depth images, in: *Proceedings of IEEE Computer Vision and Pattern Recognition, CVPR, Colorado Springs, 2011.*
10. J. Han, L. Shao, D. Xu, J. Shotton, Enhanced computer vision with Microsoft Kinect sensor: a review, *IEEE Trans. Cybern.* 43(5) (2013) 1318–1334.
11. R. Munoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, A. Carmona-Poyato, Depth silhouettes for gesture recognition, *Pattern Recognit. Lett.* 29 (3) (2008) 319–329.
12. F. Dominio, M. Donadeo, P. Zanuttigh, Combining multiple depth-based descriptors for hand gesture recognition, *Pattern Recognit. Lett.* 50 (2014) 101–111.
13. Holte M. B., Moeslund T. B., Fihl P. Fusion of range and intensity information for view invariant gesture recognition // *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. – IEEE, 2008. – С. 1-7.
14. Van den Bergh M. et al. Real-time 3D hand gesture interaction with a robot for understanding directions from humans // *2011 Ro-Man*. – IEEE, 2011. – Pp. 357-362.
15. Ren Z., Yuan J., Zhang Z. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera // *Proceedings of the 19th ACM international conference on Multimedia*. – 2011. – Pp. 1093-1096.
16. Wu D., Zhu F., Shao L. One shot learning gesture recognition from rgb-d images // *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. – IEEE, 2012. – Pp. 7-12.
17. Keskin C., Kirac F., Kara Y., Akarun L. Randomized decision forests for static and dynamic hand shape classification // *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. – IEEE, 2012. – Pp. 31-36.
18. Liao, B., Li, J., Ju, Z., Ouyang, G. Hand gesture recognition with generalized hough transform and DC-CNN using realsense // *2018 Eighth International Conference on Information Science and Technology (ICIST)*. – IEEE, 2018. – Pp. 84-90
19. Сатыбалдина Д.Ж., Калымова К.А. Разработка приложения, управляемого жестами, с использованием MICROSOFT KINECT SENSOR // «Цифровая обработка сигналов и ее применение – DSPA-2019»: Сборник докладов 21-й Международной конференции, Москва, Россия, 2019.- стр. 525-529.

ӘОЖ 001.378

ӨНЕРКӘСІП КӘСІПОРЫНДАРЫНЫҢ АҚПАРАТТЫҚ ҚАУІПСІЗДІГІН ҚАМТАМАСЫЗ ЕТУ

Касимов Асылбек Агыбаевич

a.kasimov-81@mail.ru

Л.Н.Гумилев атындағы ЕҰУ Информатика және ақпараттық жүйе

магистранты, Астана, Қазақстан

Ғылыми жетекшісі – т.ғ.к., доц. Туребаева Р.Д.

Қазіргі уақытта өнеркәсіптік кәсіпорындардың ақпараттық ағындардың едәуір көлеміне тәуелділігі өсуде. Ақпараттық және коммуникациялық технологиялардың дамуымен субъектілер арасында әртүрлі ақпарат түрлерімен алмасу саласында өзгеріс болды, бұл ақпараттық технологиялардың көмегімен қазіргі қоғамның көптеген міндеттерін жедел шешуге мүмкіндік береді. Дегенмен, олардың негізінде ішкі ұйымдастыру, ұйымаралық,