

Садртдинов И.А. Исследование имеющихся способов формирования дескрипторов наиболее подходящих для распознавания кистей рук на изображениях // Академия педагогических идей «Новация». Серия: Студенческий научный вестник. – 2018. – №6 (июнь). – АРТ 387-эл. – 0,4 п.л. - URL: <http://akademnova.ru/page/875550>

РУБРИКА: ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.051

Садртдинов Илья Айдарович

магистрант 2 курса, факультет техника и технологии

Научный руководитель: Маков С.В., к.т.н., доцент

ФГБОУ ВПО «Институт сферы обслуживания и предпринимательства»

г. Шахты, Российская Федерация

e-mail: Rino1666@gmail.com

**ИССЛЕДОВАНИЕ ИМЕЮЩИХСЯ СПОСОБОВ ФОРМИРОВАНИЯ
ДЕСКРИПТОРОВ НАИБОЛЕЕ ПОДХОДЯЩИХ ДЛЯ
РАСПОЗНАВАНИЯ КИСТЕЙ РУК НА ИЗОБРАЖЕНИЯХ**

Аннотация: В области обработки изображений задача распознавания образов является одной из широко исследуемых задач. В течение последних лет было показано, что удобным представлением изображений для их анализа является совокупность векторов, описывающих окрестности точек интереса – дескрипторов локальных особенностей изображений. В этой работе было проведено исследование имеющихся способов формирования дескрипторов, наиболее подходящих для распознавания кистей рук на изображениях, а также сравнение полученных дескрипторов на базе данных изображений кистей рук.

Ключевые слова: дескрипторы, распознавание изображений, кисти рук.

Sadrtdinov Ilya Aydarovich
2nd year master, Faculty of Engineering and Technology
Supervisor: S.V. Makov, Ph.D., Associate Professor
FGBOU HPE "Institute of Service and Entrepreneurship"
Shakhty, Russian Federation

INVESTIGATION OF AVAILABLE METHODS FOR FORMING DESCRIPTORS MOST APPROACHED FOR RECOGNIZING HANDS OF HANDS ON IMAGES

Abstract: In the image processing area, the pattern recognition task is one of the widely researched tasks. In recent years, it has been shown that a convenient representation of images for their analysis is a set of vectors describing neighborhoods of points of interest - descriptors of local image features. In this work, a study was made of the available ways of forming the descriptors most suitable for recognizing hand brushes on images, as well as comparing received descriptors on the basis of image data of the hands.

Keywords: descriptors, image recognition, hands.

1. Введение

Точное и надежное детектирование изображения кисти руки человека на изображении или в видеопотоке является необходимым и критически важным этапом при построении простых в использовании систем бесконтактного взаимодействия человека и технических устройств, например, в задачах распознавания жестов биометрической идентификации. Можно выделить три основных подхода к решению данной задачи: подход, основанный на использовании алгоритмов вычитания фона [1 – 3], подход, основанный на цветовой сегментации участков кожи [4 – 8] с последующим детектированием ладони, часто с привлечением дополнительной

информации о форме ладони [9, 10], а также подход на основе данных о глубине сцены [11].

Подход, описанный в этой работе, предлагает использовать для распознавания кистей рук дескриптор, входными данными которого являются изображения и набор особых точек, выделенных на заданном изображении. Выходными данными дескриптора являются множество векторов признаков для исходного набора особых точек. Дескрипторы должны обладать следующими свойствами: повторяемость; локальность; репрезентативность; точность; эффективность; инвариантность.

При распознавании кистей рук важно обнаружить сходство отдельных частей изображений, поэтому для данных целей применяются локальные дескрипторы, описывающие особенности областей изображений. Методы поиска по локальным дескрипторам дадут более точные результаты по сравнению с методами поиска по глобальным дескрипторам.

2. Обзор существующих дескрипторов

В ходе анализа литературы было установлено, что важными характеристиками изображений кистей рук является информация о цвете и локальных геометрических особенностях. Исходя из этого, для задачи распознавания кистей рук выбраны популярные дескрипторы, представленные ниже и учитывающие данные виды информации.

Гистограмма направленных градиентов (англ. Histogram of Oriented Gradients, HOG) [12]. Дескриптор основан на подсчете количества направлений градиента в локальных областях изображения и вычисляется на плотной сетке равномерно распределенных ячеек, используя нормализацию перекрывающегося локального контраста для увеличения точности.

Дескрипторы на основе анализа гистограмм. Гистограммы представляют собой одномерный вектор, элементы которого (бины) являются количеством пикселей изображения, яркость которых (или другая числовая характеристика) попадает в заданный диапазон. Собранные данные могут иметь любой признак, который будет полезным для описания изображения.

Гистограмма может вести учет не только интенсивности цвета, но и другие параметры изображения, которые можно измерить (например, градиенты, направления и т.д.) и использовать в качестве дескрипторов [13].

Дескриптор LBP используется для поиска объекта на изображении, а также проверки этого объекта на принадлежность некоторому классу (верификация, распознавание эмоций, пола по лицу). Впервые был предложен Т.Оjala в 1996 году [14]. Он представляет собой эффективный дескриптор, который представляет каждый пиксель изображения в виде бинарного числа, зависящего от интенсивностей соседних пикселей изображения. LBP дескриптор является эффективным в вычислительном плане, так как работает только с целочисленной арифметикой (это позволяет достигать real-time производительности в некоторых задачах), а также он инвариантен к изменениям яркости на изображении, которые вызваны съемкой в различных условиях освещения.

Цель создания BRIEF-дескриптора (Binary Robust Independent Elementary Features) [4] состояла в том, чтобы обеспечить распознавание одинаковых участков изображения, которые были сняты с разных точек зрения. При этом ставилась задача максимально уменьшить количество выполняемых вычислений. Вычисление BRIEF-дескрипторов занимает очень мало времени, поскольку сводится к вычислению расстояния Хэмминга между двумя последовательностями битов. Сами по себе

дескрипторы BRIEF не инвариантны к повороту. Однако такой инвариантности можно добиться, если предварительно повернуть фрагмент вокруг точки интереса на угол, соответствующий, например, доминирующему направлению градиента яркости, как это делается для дескрипторов SIFT и SURF.

Дескриптор Haralick [16] основан на статистике, отражающей частоту присутствия в текстуре различных пар уровней яркости пикселя. Составляется матрица частот, число строк и столбцов которой равны числу уровней яркости в текстуре.



Рисунок 2 – Алгоритм исследования конкретного дескриптора

Алгоритм исследования конкретного дескриптора приведен на рисунке 2. Данная последовательность действий осуществляется для каждого дескриптора из числа выбранных.

3. Программы и оборудование

В качестве среды для программирования использовалась Anaconda, установленная на персональный компьютер. Это бесплатный дистрибутив языка python, который включает в себя множество научных и инженерных

библиотек. К этой среде была подключена библиотека алгоритмов компьютерного зрения OpenCV, применяемая для обработки изображений и численных алгоритмов общего назначения с открытым кодом. Scikit-image - это библиотека с открытым исходным кодом для обработки изображений в языке программирования Python. Она включает в себя алгоритмы сегментации, геометрических преобразований, пространства манипуляций с цветом, анализа, фильтрации, морфологии, обнаружения признаков, и многое другое [17].

Для классификации изображений был использован метод опорных векторов (Support Vector Machines).

Для визуализации многомерных переменных был выбран алгоритм t-SNE [18], который также относят к методам множественного обучения признаков.

Так же была построена ROC-кривая [19] — график, позволяющий оценить качество бинарной классификации.

Для более наглядного анализа был сформирован случайный (Random) дескриптор – это массив, заполненный случайными элементами, который используется классификатором в качестве дескриптора. Этот дескриптор нужен только для проведения сравнительного анализа и не пригоден для практического применения. Он представлен в работе в подтверждение того, что классификатор и ROC кривая работают правильно.

4. Результаты проведения эксперимента

Для обучения машины опорных векторов используется обучающая выборка, состоящая из 4870 положительных примеров (содержащих кисть руки) и 4870 отрицательных примеров (не содержащих изображения кисти руки).

Примеры обучающей и тестовой выборки изображений представлены на рисунке 3.



Рисунок 3 – Примеры изображений из обучающей и тестовой выборки.

Исследование эффективности классификатора производится на тестовой выборке, отличной от выборки на которой производилось обучение. Тестовая выборка содержит 610 изображение с кистями рук и около 633 изображений без кистей рук. Сравнительный анализ проводится по полученным в ходе эксперимента данным ROC-кривой и проекции векторов из многомерного пространства в двумерное, которая отражает разделимость классов.

Дескриптор HOG показал точность классификации кисти руки 77%. На ROC-кривой (рисунок 4), можно увидеть, что кривая расположена очень близко к левому углу, а значит у модели на основе этого дескриптора очень большая предсказательная способность. На визуализации методом t-SNE для данного дескриптора хорошо выражен класс «есть рука» (точки синего цвета), который отделен от класса «нет руки» (точки красного цвета).

Дескриптор на основе трехмерной гистограммы (HIST3D) показал точность классификации кисти руки 16.4%. Посмотрев на ROC-кривую (рисунок 5), можно увидеть, что кривая расположена почти по диагонали и имеет хорошо выраженный изгиб. Модель на основе этого дескриптора не будет эффективной для данной задачи. На проекции векторов из

многомерного пространства в двумерное для данного дескриптора видно, что векторы класса «есть рука» (точки синего цвета) имеют не такую хорошую кучность, как у предыдущего дескриптора (HOG). Но этот класс отделим от класса «нет руки» (точки красного цвета).

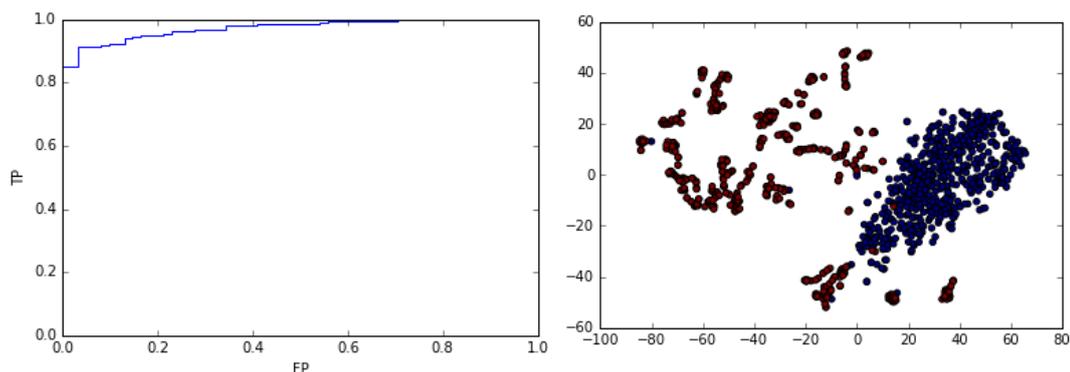


Рисунок 4 – Иллюстрация ROC-кривой и визуализация методом t-SNE для дескриптора HOG

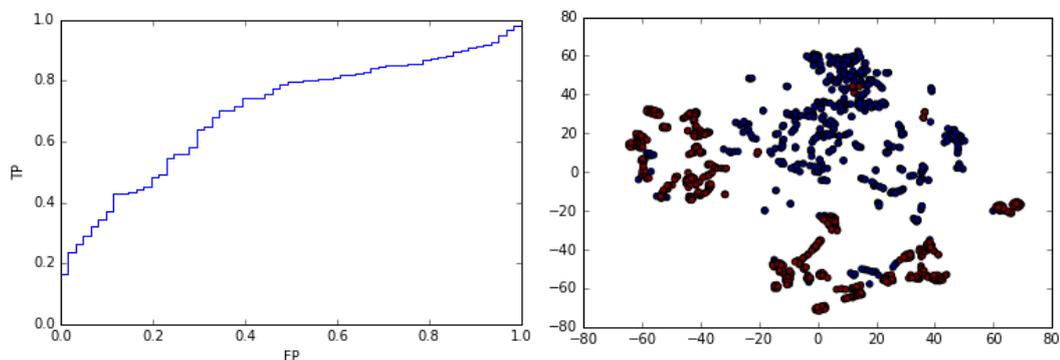


Рисунок 5 – Иллюстрация ROC-кривой и визуализация методом t-SNE для дескриптора HIST3D

Дескриптор на основе двумерной гистограммы (HISTHS) показал точность классификации кисти руки 13.11%. Точность хуже, чем у трехмерной гистограммы. Посмотрев на ROC-кривую (рисунок 6), можно увидеть, что кривая расположена почти по диагонали и имеет меньший изгиб, чем HIST3D. Значит, модель на основе этого дескриптора менее эффективна, чем HIST3D и также не пригодна для данной задачи. По проекции векторов для данного дескриптора проекция можно заметить, что

она схожа с проекцией для дескриптора HIST3D. Но разделение классов «есть рука» (точки синего цвета) и «нет руки» (точки красного цвета), а также подразделение классов на локальные группы выражено сильнее.

Дескриптор LBP показал точность классификации 81.96% на тестовой выборке. Посмотрев на ROC-кривую (рисунок 7), можно увидеть, что кривая расположена очень близко к левому углу, даже ближе чем у HOG дескриптора, а значит у модели на основе этого дескриптора очень хорошая предсказательная способность. На проекции векторов из многомерного пространства в двумерное для дескриптора LBP хорошо выражен класс «есть рука» (точки синего цвета), который отделен от класса «нет руки» (точки красного цвета).

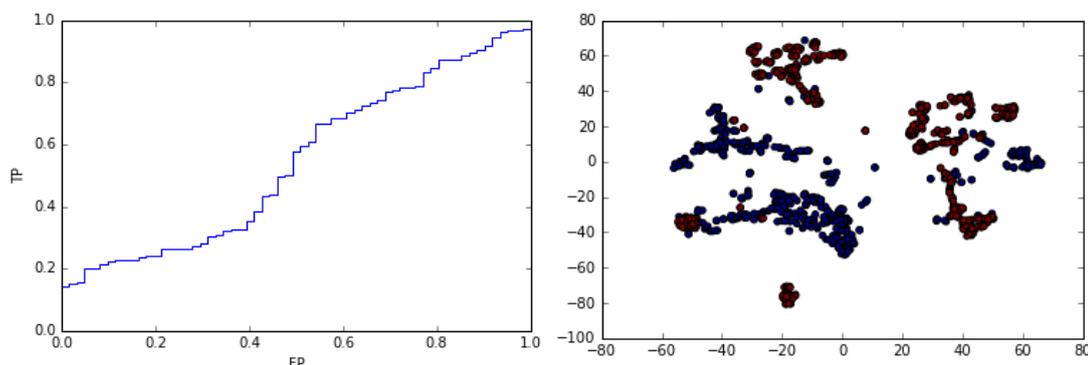


Рисунок 6 – Иллюстрация ROC-кривой и визуализации методом t-SNE для дескриптора HISTHS

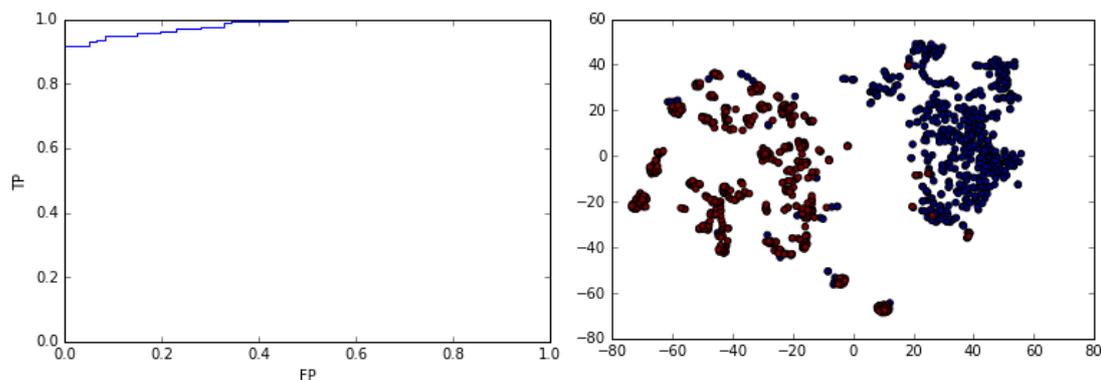


Рисунок 7 – Иллюстрация ROC-кривой и визуализации методом t-SNE для дескриптора LBP

Дескриптор BRIEF показал точность классификации кисти руки 62.3%. Посмотрев на ROC-кривую (рисунок 8), можно увидеть, что кривая расположена по диагонали и имеет хорошо выраженный изгиб. Модель на основе этого дескриптора будет менее эффективна, чем HOG или LBP, но даст лучшие результаты, чем дескрипторы, основанные на гистограммах. На проекции векторов для дескриптора BRIEF класс «есть рука» (точки синего цвета) и класс «нет руки» (точки красного цвета) не имеют кучности, по сравнению с результатами, рассмотренными ранее. Так же на визуализации отсутствует ясно выраженное разделение на 2 класса, как в предыдущих результатах, но, благодаря отсутствию явно выраженного деления на локальные группы, точность классификации составляет 62.3%

Дескриптор RANDOM показал точность классификации кисти руки 45.9%. Посмотрев на ROC-кривую (рисунок 9), можно увидеть, что кривая расположена ровно по диагонали. Такой дескриптор не пригоден для применения. На визуализации методом t-SNE для этого дескриптора проекции векторов обоих классов находятся вперемешку и не имеют четкого разделения.

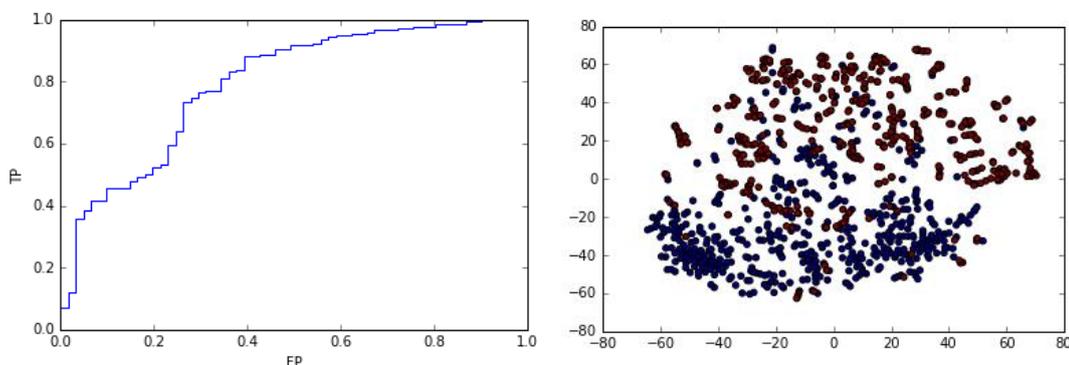


Рисунок 8 – Иллюстрация ROC-кривой визуализации методом t-SNE для дескриптора BRIEF

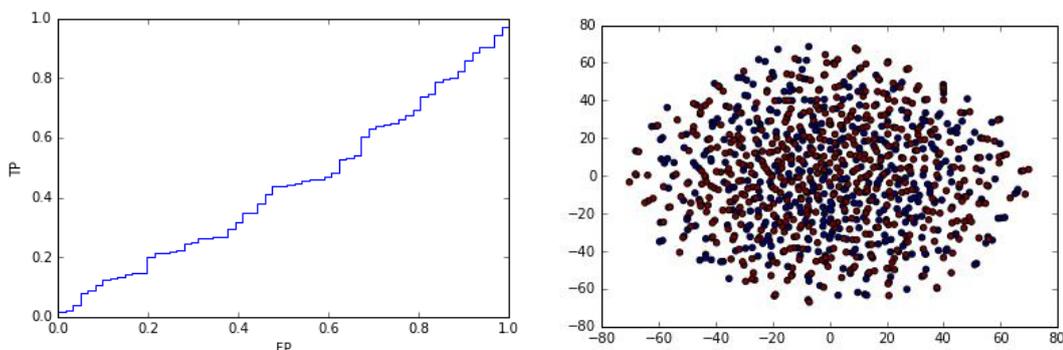


Рисунок 9 – Иллюстрация ROC-кривой и визуализации методом t-SNE для дескриптора RANDOM

Дескриптор Haralick показал точность классификации кисти руки 65.57%. Посмотрев на ROC-кривую (рисунок 10), можно увидеть, что кривая расположена довольно близко к левому углу, но не так близко, как у HOG или LBP дескриптора. Исходя из ROC-кривой модель, на основе этого дескриптора имеет неплохую предсказательную способность т.к. кривая расположена близко к верхнему левому углу. На полученной проекции векторов из многомерного пространства в двумерное для дескриптора Haralick видна хорошая разделимость классов. Но из-за большого деления на локальные группы точность классификации ниже, чем у HOG и LBP.

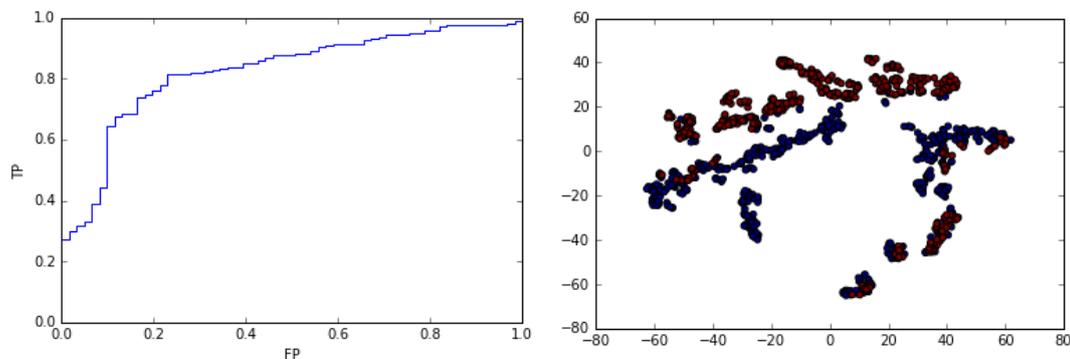


Рисунок 10 – Иллюстрация ROC-кривой и визуализации методом t-SNE для дескриптора Haralick

В таблице 2 содержится информация сколько раз, во время проведения эксперимента, система приняла верное, и сколько раз неверное решение при использовании дескриптора для классификации.

Таблица 2 – Таблица сопряжённости алгоритма классификации.

Исследуемый дескриптор	Ответ классификатора	
	Положительно (TP)	Отрицательно (TN)
HOG	77%	90%
HIST3D	16.4%	78%
HISTHS	13.11%	90%
Naralick	65.57%	70.49%
LBP	81.96%	94%
BRIEF	62.3%	73.77%
Random	45.9%	42.62%

5. Заключение

В результате исследования было установлено, что для задачи распознавания кистей рук лучше всего подходят два дескриптора: LBP и HOG. Самый низкий положительный ответ классификатора дали дескрипторы на основе анализа двухмерной (HISTHS) и трехмерной гистограммы (HIST3D). Так же было установлено, что дескриптор Naralick и BRIEF применим для данной задачи, но имеет меньшую точность классификации кисти руки, чем HOG и LBP в условиях поставленного эксперимента.

Дальнейшие исследования будут сосредоточены на исследовании и сравнительном анализе современных методов распознавания образов с использованием нейронных сетей.

Список использованной литературы:

1. Y. Benezeth. Review and evaluation of commonly-implemented background subtraction algorithms // Conf Pat. Recong.2008. С. 1 – 4.
2. M. Picardi. Background subtraction techniques: a review // IEEE Int. Conf. Man Cybern. (IEEE Cat. No.04CH37583). 2004 Vol. 4. С. 3099 – 3104.
3. S. Shiravandi, M. Rahmat, F. Mahmoudi. Hand gestures recognition using dynamic Bayesian networks // 3rd Joint Conference of AI & Robotics and 5th RoboCump Iran Open International Symposium. IEEE, 2013. С. 1 – 6.
4. M. Jones, J. Rehg. Statistical Color Models with Application to Skin Detection // Int. J. Comput. Vis. 2002. Vol. 46, № 1. С. 81 – 96.
5. P. Kakumanu, S. Makrogiannis, N. Bourbakis. A survey of skin-color modeling and detection methods // Pat. Recogn. 2007. Vol. 40, № 3. С. 1106 – 1122.
6. V. Vezhnevets. A Survey on Pixel-Based Skin color Detection Tech-niques // Cybernetics. 2003. Vol. 85, № 0896-6273 SB-IM. С. 85 – 93.
7. M. Wimmer. Adaptive Skin Color Classifier // Proc. first ICGST Int. Conf. Graph. Vis. Image Process. GVIP-05. 2005. № December. С. 324–327.
8. S. Phung, A. Bouzerdoum, D. Chai. Skin segmentation using color pixel classification: analysis and comparison. // IEEE Trans. Pattern Anal. Mach. Intell. 2005. Vol. 27. № 1. С. 148 – 154.
9. W. Junqiu, Y. Yagi. Integrating Color and Shape-Texture Features for Adaptive Real-Time Object Tracking // Image Process. IEEE Trans. 2008. Vol. 17, № 2. С. 235–240.
10. Малашин Р.О., Луцив В.Р. Восстановление силуэта руки в задаче распознавания жестов с помощью адаптивной морфологической фильтрации бинарного изображения // Оптический журнал. 2013. Vol. 80. № 11. С. 54 – 61.
11. J. Suarez, R. Murphy. Hand gesture recognition with depth images: A review // Proc. IEEE Int. Work. Robot Hum. Interact. Commun. 2012. № August 2016. С. 411 – 417.
12. Dalal N., Triggs B. Histograms of oriented gradients for human detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE. 2005. V. 1. С.886-893.
13. OpenCV 2.4.13.0 documentation // Histogram calculation URL: http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_calculation/histogram_calculation.html (дата обращения 2.02.18).
14. T. Ojala, M. Pietikäinen, and D. Harwood. A Comparative Study of Tex-ture Measures with Classification Based on Feature Distributions // Pattern Recog-nition. 1996 vol. 29, С. 51-59.
15. Alpaslan F., Apolloni B., Ghosh A., Jain L.C., Patnaik S. Machine Learn-ing and Robot Perception // Springer, 2005.
16. Robert M. Haralick, et al. Textural Features for Image Classification // IEEE Transactions on Systems, Man and Cybernetics. Vol. SMC-3, No. 6, Nov. 1973, С. 610-621.
17. Scikit-image // Материал из Википедии — свободной энциклопедии. URL: <https://en.wikipedia.org/wiki/Scikit-image> (дата обращения 18.02.18).

18. L.J.P. van der Maaten and G.E. Hinton. Visualizing High-Dimensional Data Using t-SNE // Journal of Machine Learning Research 9(Nov): C. 2579-2605, 2008.
19. Yang Z., Ai H. Demographic classification with local binary patterns. Advances in Biometrics // Springer Berlin Heidelberg, 2007. – С. 464-473.

Дата поступления в редакцию: 20.06.2018 г.

Опубликовано: 25.06.2018 г.

© Академия педагогических идей «Новация». Серия «Студенческий научный вестник», электронный журнал, 2018

© Садртдинов И.А., 2018