

Исследование методов классификации для распознавания действий человека на основе PCA и векторов ширины

Л.В. Ширипова¹, О.В. Струкова¹, Е.В. Мясников^{1,2}

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. В статье представлены результаты исследования методов классификации для распознавания действий человека на основе PCA и векторов ширины с использованием видеосъемки в оптическом диапазоне. Метод, используемый в данной статье, состоит в обнаружении движущегося человека на видеопоследовательности с последующей нормализацией размера, выделением набора подпоследовательностей и получением векторов признаков. Классификация действий человека осуществляется с использованием машины опорных векторов с различными ядрами, классификатора по ближайшим соседям и классификатора случайный лес. Полученные результаты позволили подобрать наиболее эффективные параметры для классификаторов.

1. Введение

Одной из важнейших нерешенных задач остается распознавание действий человека, особенно в неконтролируемой обстановке, например, в толпе. Некоторые движения тела не всегда можно правильно понять с точки зрения адекватной оценки агрессивного поведения. В области компьютерного зрения достаточно часто применяются методы классификации для распознавания позы тела по видеопоследовательностям. Данные алгоритмы используют для определения поведенческих параметров человека или сразу нескольких людей.

Любые исследования в данном направлении позволят внести вклад в прогнозирование поведения динамических объектов, а распознавание действий человека позволит в конечном итоге предвидеть опасные ситуации и устранять их. Распознавание в комплексе с видеоаналитикой позволит решать различные задачи, полезные для бизнеса и государства, общества в целом. У распознавания действий огромные перспективы применения на промышленных предприятиях, в здравоохранении, в области предотвращения угроз общественной безопасности и обеспечения общественного порядка.

Для решения задачи распознавания действий было предложено множество подходов, подробно описанных в работе [1]. Рассмотрим некоторые из них.

Для выделения признаковой информации авторы работы [2] предложили выделять силуэт из каждого кадра, получать изображения разницы между соседними кадрами и строить окончательное изображение, накладывая полученные изображения друг на друга.

Для устранения указанной проблемы в работе [3] предложен обобщающий подход, связанный с использованием 3D motion history volume (MHV). MHV основывается на 3D вокселях, полученных для различных углов наблюдения. Далее используется преобразование Фурье для получения признаков, инвариантных к положению и вращению.

В работе [4] в качестве признаков используются простые параметры выпуклых фигур, например, площадь, периметр, компактность, закругленность, отклонение от круглости, круговая дисперсия и эллиптическая дисперсия выпуклой оболочки.

Авторы работы [5] предлагают объединение алгоритмов HOG и PCA для получения признаков, используемых в дальнейшем для обучения классификатора k-NN.

Для классификации полученных признаков используются различные подходы, например, машина опорных векторов (SVM) [6, 7, 8], классификатор по ближайшему соседу (k-NN) [9, 10, 11], а также скрытые Марковские модели (HMM) [12, 13, 14].

В настоящей работе для решения задачи распознавания действий человека применяется подход, основанный на снижении размерности с использованием метода главных компонент и последующей классификации с использованием машины опорных векторов. Подобный подход успешно использовался нами ранее [15, 16, 17] при решении задачи идентификации человека по походке.

Работа имеет следующую структуру. В разделе 2 приводится описание разработанного метода распознавания действий человека. В разделе 3 описываются результаты экспериментальных исследований, выполненных на базе данных Weizmann. В конце работы приводится заключение и список литературы.

2. Методы

Предложенный ранее метод [15-18] включает следующие этапы:

- выделение движущегося человека на видеопоследовательности,
- нормализацию размера выделенного фрагмента видеопоследовательности,
- выделение подпоследовательностей,
- формирование векторов признаков двумя способами,
- классификацию видеопоследовательности.

2.1. Выделение движущегося человека на видеопоследовательности

На первом этапе исследования производится выделение движущегося человека на видеопоследовательности с использованием метода вычитания фона на основе смеси гауссовских распределений (Gaussian mixture model, GMM) [25].

Итогом работы первого этапа является сформированный набор масок, соответствующих отдельным кадрам видеопоследовательности.

2.2. Нормализация размера выделенного фрагмента видеопоследовательности

На втором этапе метода полученные покадрово маски обрабатываются следующим образом. Вначале производится поиск центра масс каждой области переднего плана, затем определяются линейные размеры области, и выполняется кадрирование (усечение изображения маски). После этого кадрированное изображение подвергается масштабированию (сжатию) до заданных размеров.

Очевидно, что с учетом временной координаты размерность описывающих движение человека данных остается высокой даже после их нормализации (кадрирования и сжатия) изображений. В связи с этим производится снижение размерности данных, описывающих движение человека.

2.3. Выделение подпоследовательностей

Для каждой последовательности кадров, содержащей движение, выделяется набор подпоследовательностей заданной длины. Выделение подпоследовательностей осуществляется с некоторым заданным шагом, начиная с начала исходной последовательности. Подробное описание выделения подпоследовательностей приведено в предыдущих работах [15-18].

2.4. Формирование векторов признаков

2.4.1. Первый способ: снижение размерности с использованием метода главных компонент

Для каждой выделенной подпоследовательности вектор признаков формируется следующим образом: каждый нормализованный кадр подпоследовательности разворачивается в строку, и полученные для отдельных кадров строки сцепляются между собой. Вектора признаков подпоследовательностей всех последовательностей формируют входную матрицу для метода главных компонент.

После нахождения главных компонент в качестве признакового описания берется проекция данных исходных подпоследовательностей на первые N главных компонент.

2.4.2 Второй способ: использование векторов ширины

В качестве второго способа формирования признаков использовались вектора ширины, с помощью которых можно представить любое действие. Вектор ширины определяется как разница между значениями координаты x самого левого ненулевого пикселя и самого правого ненулевого пикселя в строке кадра. Таким образом, для изображения размера $W \times H$ вектор ширины может быть записан как:

$$w(t) = [w_1(t), w_2(t), \dots, w_H(t)] \in \bullet H,$$

где $w_h(t) = x_{hR}(t) - x_{hL}(t) \geq 0$, для $\forall h = 1, 2, \dots, H$ – разность местоположения крайнего правого пикселя $x_{hR}(t)$ и крайнего левого пикселя $x_{hL}(t)$ для данной строки h в t -м кадре. Полученные вектора ширины берутся в качестве признакового описания.

2.5. Классификация видеопоследовательностей

Полученные в результате метода главных компонент признаки используются для обучения классификатора, в качестве которого в настоящей работе выступает машина опорных векторов (support vector machines, SVM) [19] с различными ядрами, классификатор по ближайшим соседям и алгоритм Random Forest. При этом классы соответствуют отдельным действиям, а вектора признаков, полученные для всех соответствующих этим действиям подпоследовательностей, соответствуют отдельным реализациям.

Отметим, что приведенное в текущем разделе описание справедливо для режима обучения, в котором производится настройка, как процедуры снижения размерности, так и классификатора. В режиме тестирования данные проходят обработку аналогичным образом, за исключением того, что при снижении размерности используются параметры полученного в режиме обучения линейного преобразования с последующей классификацией обученным классификатором.

3. Эксперименты

Предложенный в работе метод был реализован на языке C++ и Python с использованием библиотеки OpenCV и Scikit-learn. При выполнении экспериментальных исследований был использован ПК на базе Intel Core i5-3470 CPU 3.2 ГГц.

Для экспериментального исследования предложенного метода были использованы раскадрированные видеопоследовательности из открытой базы данных Weizmann (рисунок 1). Указанная база данных содержит последовательности бинарных изображений, соответствующих отдельным кадрам видеопоследовательности, на которых уже выполнено выделение движущихся объектов (сегментация переднего плана и фона). В базе данных представлены видеопоследовательности для 9 человек, выполняющих 10 различных действий. Всего база данных содержит 90 последовательностей. Таким образом, в каждом классе было 9 последовательностей. Минимальная длина последовательности составляла 28 кадров.

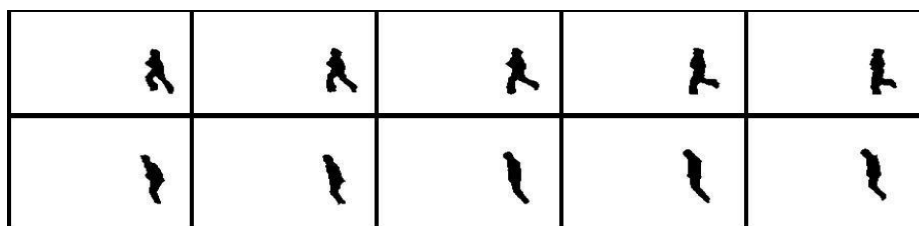


Рисунок 1. Примеры последовательностей из базы данных Weizmann: бег и прыжки в длину.

Классы были разделены на обучающую и тестовую выборки. Последовательности каждого класса были предварительно обработаны с использованием описанного ранее алгоритма. Шаг, используемый при выделении подпоследовательностей, был выбран в соответствии с предыдущими исследованиями [15-18] и составлял 2 кадра, длина подпоследовательности – 28 кадров. При проведении эксперимента применялась полная кросс-валидация по всем последовательностям. Для этого использовались все сочетания последовательностей для обучения и для тестирования. Для определения качества работы метода оценивалась точность классификации, определяемая как доля правильно классифицированных объектов.

В первом эксперименте исследовалась зависимость точности классификации от используемого ядра метода опорных векторов. При проведении эксперимента использовались следующие ядра: Inter, RBF (Radial Basis Function), Polynomial, Linear. Для получения векторов признаков использовались способы, рассмотренные в п. 2.4. а выходная размерность после применения PCA равнялась 64 в соответствии с результатами предыдущих работ [15-18]. Количество последовательностей для обучения составляло 5, для тестирования - 4. Результаты представлены на рисунке 2. Экспериментально было определено, что наилучшие значения показателя точности классификации достигаются при использовании линейного ядра, которое применяется в последующих экспериментах.

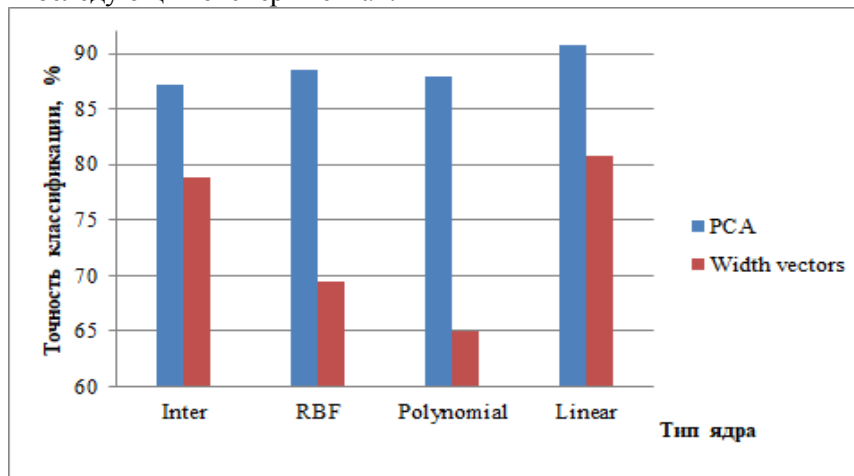


Рисунок 2. Зависимость точности классификации от выбранного типа ядра.

Во втором эксперименте исследовалась зависимость точности классификации от используемого классификатора и от количества последовательностей для обучения и тестирования. Для получения векторов признаков использовался способ, рассмотренный в п. 2.4.1 а выходная размерность после применения PCA равнялась 64. Количество последовательностей для обучения и тестирования варьировалось от 3 до 6. В качестве классификатора использовалась машина опорных векторов с ядром Linear, классификатор по ближайшим соседям с $k=1, 3$ и алгоритм случайный лес. Результаты представлены на рисунке 3.

В третьем эксперименте исследовалась зависимость точности классификации от используемого классификатора и от количества последовательностей для обучения и тестирования. Для получения векторов признаков использовались вектора ширины,

рассмотренные в п. 2.4.2. Количество последовательностей для обучения и тестирования варьировалось от 3 до 6. В качестве классификатора использовалась машина опорных векторов с ядром Linear, классификатор по ближайшим соседям с $k=1, 3$ и алгоритм случайный лес. Результаты представлены на рисунке 4.

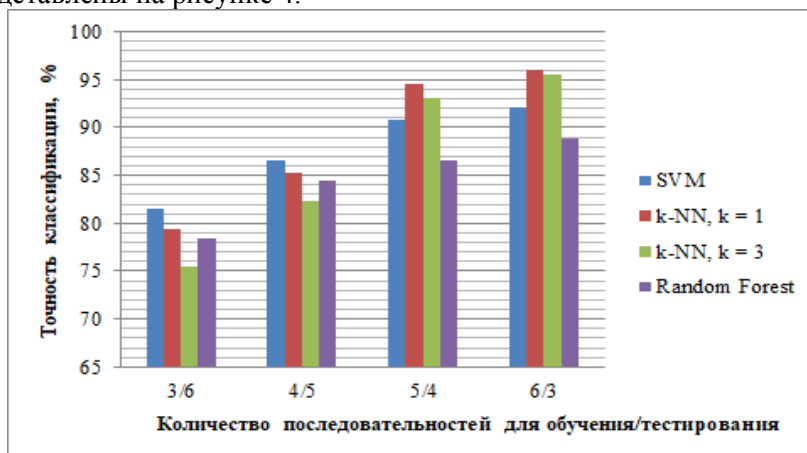


Рисунок 3. Зависимость точности классификации от используемого классификатора и от количества последовательностей для обучения и тестирования с использованием снижения размерности.

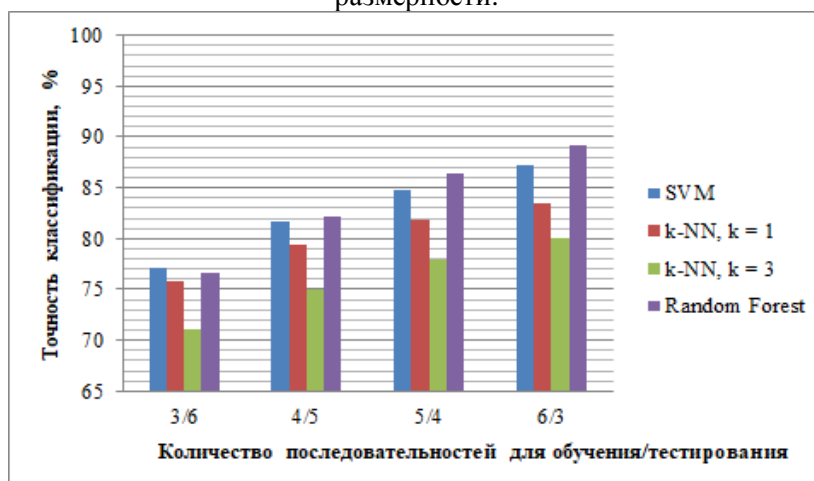


Рисунок 4. Зависимость точности классификации от используемого классификатора и от количества последовательностей для обучения и тестирования с использованием векторов ширины.

Экспериментально было определено, что наилучшие значения показателя точности классификации достигаются при формировании признаков с использованием снижения размерности.

4. Заключение

Предложенный нами метод распознавания действий человека состоит в выделении движущегося человека на видеоряде, нормализации размера, выделении подпоследовательностей, формировании векторов признаков и классификации.

В рамках предложенного ранее подхода [15-18] был исследован новый способ формирования признаков с использованием векторов ширины.

Эксперименты, выполненные на базе данных Weizmann, позволили определить наилучшие значения параметров разработанного метода. В ходе исследований было проведено сравнение четырех типов ядер классификатора SVM. Самый высокий результат был достигнут с использованием линейного ядра.

В будущем планируется расширить список алгоритмов, используемых для формирования признакового описания и список используемых методов классификации.

5. Благодарности

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 17-29-03190-офи-м в частях «1. Введение» - «2. Методы» и Министерства науки и высшего образования РФ в рамках госзадания ФНИЦ «Кристаллография и фотоника» РАН в частях «3. Эксперименты» и «4. Заключение».

6. Литература

- [1] Kong, Y. Human action recognition and prediction: a survey / Y. Kong, Y. Fu // *Journal of Latex class files*. – 2018. – Vol. 13(9). – P. 1-20.
- [2] Bobick, A. The Recognition of Human Movement Using Temporal Templates / A.F. Bobick, J.W. Davis // *IEEE Trans Pattern Analysis and Machine Intelligence*. – 2001. – Vol. 23. – P. 257-267.
- [3] Weinland, D. Free viewpoint action recognition using motion history volumes / D. Weinland, R. Ronfard, E. Boyer // *Computer Vision and Image Understanding*. – 2006. – Vol. 104. – P. 249-257.
- [4] Gosciewska, K. Silhouette-based action recognition using simple shape descriptors / K. Gosciewska, D. Frejlichowski – Springer, 2018.
- [5] Jahagirdar, A.S. Silhouette-based human action recognition by embedding HOG and PCA features / A.S. Jahagirdar, M.S. Nagmode // *Intelligent Computing and Information and Communication*. – 2018. – Vol. 673. – P. 363-371.
- [6] Laptev, I. Recognizing Human Actions: A Local SVM Approach / C. Schuldt, I. Laptev, B. Caputo // *Proc. ICPR'04 – Cambridge, UK, 2004*.
- [7] Marszalek, M. Actions in context / M. Marszalek, I. Laptev, C. Schmid // *Computer Vision and Pattern Recognition, 2009*.
- [8] Laptev, I. Learning realistic human actions from movies / I. Laptev, M. Marszalek, C. Schmid, B. Rozenfeld // *CVPR, 2008*.
- [9] Blank, M. Actions as space-time shapes / M. Blank, L. Gorelick, E. Shechtman, M. Irani, R. Basri // *Proc. ICCV, 2005*.
- [10] Laptev, I. Retrieving actions in movies / I. Laptev, P. Perez // *ICCV, 2007*.
- [11] Tran, D. Human activity recognition with metric learning / D. Tran, A. Sorokin // *ECCV, 2008*.
- [12] Duong, T.V. Activity recognition and abnormality detection with the switching hidden semi-markov model / T.V. Duong, H.H. Bui, D.Q. Phung, S. Venkatesh // *CVPR, 2005*.
- [13] Rajko, S. Real-time gesture recognition with minimal training requirements and on-line learning / S. Rajko, G. Qian, T. Ingalls, J. James // *CVPR, 2007*.
- [14] Ikizler, N. Searching video for complex activities with finite state models / N. Ikizler, D. Forsyth // *CVPR, 2007*.
- [15] Strukova, O. Gait analysis for person recognition using principal component analysis and support vector machines / O. Strukova, L. Shiripova, E. Myasnikov // *CEUR Workshop Proceedings*. – 2018. – Vol. 2391. – P. 170-176.
- [16] Shiripova, L. Comparative analysis of classification methods for human identification by gait / L. Shiripova, E. Myasnikov // *CEUR Workshop Proceedings*. – 2018. – Vol. 2268. – P. 118-128.
- [17] Strukova, O. The choice of methods for the construction of PCA-based features and the selection of SVM parameters for person identification by gait / O. Strukova, E. Myasnikov // *Journal of Physics: Conference Series*. – 2019. – Vol. 1368. – P. 032001
- [18] Shiripova, L. Human action recognition using dimensionality reduction and support vector machine / L. Shiripova, E. Myasnikov // *CEUR Workshop Proceedings*. – 2018. – Vol. 2391. – P. 48-53.
- [19] Cortes, C. Support-vector networks / C. Cortes, V. Vapnik // *Machine Learning*. – 1995. – Vol. 20(3) – P. 273-297.

Study of classification techniques for human action recognition based on PCA and width vectors

L.V. Shiripova¹, O.V. Strukova¹, E.V. Myasnikov^{1,2}

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. The article presents the results of research of the classification techniques for human action recognition based on PCA and width vectors using a video recorded in the optical range. The method used in this paper consists in the detection of a moving person on a video sequence with size normalization, formation a set of subsequences and feature vectors. The classification of the human action is carried out using support vector machine with different kernels, classifier K-Nearest Neighbors and random forest classifier. The obtained results allowed us to pick up the most effective parameters for the classifiers.