

# Распознавание личности по походке: опыт использования метода главных компонент и машины опорных векторов

О.В. Струкова<sup>1</sup>, Л.В. Ширипова<sup>1</sup>, Е.В. Мясников<sup>1</sup>

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

**Аннотация.** В работе представлены результаты исследования эффективности классификации личности по походке, регистрируемой с использованием видеосъемки в оптическом диапазоне. Предлагаемый в работе метод состоит в выделении движущегося человека на видеоряде с последующей нормализацией размера и снижением размерности с использованием метода главных компонент. Классификация личности выполнялась с использованием машины опорных векторов. В рамках экспериментальных исследований, выполненных на базе данных CASIA GAIT, были определены наилучшие значения параметров метода. Полученные результаты показали, что при небольшом количестве классов может быть достигнута высокая точность классификации личности.

## 1. Введение

В настоящее время во всем мире популярна и широко применяется идентификация человека по его биометрическим параметрам. Для идентификации личности используют такие специфические особенности как изображение лица, тембр голоса, отпечатки пальцев, узор радужной оболочки глаза и даже походка. Несмотря на то, что использование отпечатков пальцев или узора радужной оболочки позволяет практически безошибочно выполнять идентификацию человека, значительный интерес представляют бесконтактные и дистанционные способы идентификации. В этой связи особенно актуальной является задача распознавания личности по походке.

Рассматривая походку, как совокупность поз и движений, можно выделить два наиболее распространенных способа регистрации (захвата) такой информации: видеосъемка [1] (например, в оптическом диапазоне) и регистрация с использованием датчиков, находящихся на теле человека [2]. Кроме того, известны работы, в которых анализ походки выполняется на основе показаний акселерометров, встроенных в смартфон [3].

Учитывая, что походка позволяет производить идентификацию личности даже в тех случаях, когда другими способами произвести ее не представляется возможным (объект находится на расстоянии, качественное изображение лица получить невозможно и т.п.), особый

интерес представляет именно использование видеосъемки, осуществляемой, например, с использованием камер видеонаблюдения.

К настоящему времени для решения задачи идентификации личности на видео по походке исследователями использовались различные методы.

В работе [4] для идентификации личности применяется подход, состоящий в сегментации фона, осуществляемой с использованием метода вычитания фона на основе смеси гауссовских распределений (GMM), снижении размерности с использованием метода главных компонент (PCA), а также классификации на основе линейного дискриминанта Фишера (FLDA). Другой особенностью работы является комбинирование признаков движения с признаками следа (отпечатка ног) человека.

Первым этапом предложенного в работе [5] подхода является улучшенная процедура вычитания фона. При этом выделенные движения описываются дескрипторами на основе метода статистического анализа формы (Procrustes analysis). Процедура управляемой классификации строится с использованием соответствующей меры (Procrustes distance measure).

В работе [1] производится анализ эффективности использования линейного (PCA) и нелинейных (ISOMAP, LLE) методов снижения размерности для формирования признаков. Для классификации с использованием сформированных признаков применяется скрытая марковская модель (HMM).

В работе [6] для решения задачи классификации личности по движению используется метод опорных векторов (SVM). В частности, в работе исследуется зависимость точности классификации личности от выбора типа ядра SVM.

В целом можно отметить, что задача распознавания личности по походке привлекает в последние годы внимание все большего числа исследователей. При этом значительное внимание уделяется, как способам признакового описания движения, так и выбору эффективных методов классификации.

В настоящей работе для решения задачи идентификации личности по походке предлагается подход, переключающийся с указанными выше исследованиями [1,4,6]. Используемый метод состоит в выделении движущегося человека на видеопоследовательности с последующей нормализацией размера и снижением размерности с использованием метода главных компонент. В качестве классификатора используется машина опорных векторов. Значительное внимание в работе уделяется вопросам подбора параметров формирования признакового описания. В рамках проведенных исследований показана важность тщательного подбора параметров при решении рассматриваемой задачи, что позволяет достичь высоких показателей качества классификации личности при относительно небольшом количестве классов.

Работа имеет следующую структуру. Раздел 2 посвящен описанию используемого в работе метода. В разделе 3 описываются результаты экспериментов. Работа завершается заключением. В конце работы приведен список использованной литературы.

## **2. Методы**

Разработанный в рамках проведенного исследования метод включает:

- выделение движущегося человека на видеопоследовательности,
- нормализацию размера выделенного фрагмента видеопоследовательности,
- снижение размерности выделенного фрагмента видеопоследовательности,
- классификацию видеопоследовательности.

### 2.1. Выделение движущегося человека на видеопоследовательности

На первом этапе разработанного метода осуществляется выделение движущегося человека на видеопоследовательности. Наиболее часто для такого выделения в случае, когда источником видеопоследовательности является камера видеонаблюдения, используются методы вычитания фона. Основная идея методов этого класса заключается в использовании некой модели фона и принятии решения о принадлежности того или иного пикселя к фону или движущемуся объекту, исходя из его соответствия модели фона. Сама модель фона с течением времени

постепенно уточняется. Хотя в простейших приложениях в качестве модели фона может быть использовано усредненное по времени наблюдения изображение, в рассматриваемой задаче лучшие результаты дают более сложные модели, например [7-9].

В настоящей работе для выделения на видеопоследовательности движущегося человека использовался метод вычитания фона на основе смеси гауссовских распределений (Gaussian mixture model, GMM) [8]. При использовании указанного метода каждый отсчет фона моделируется взвешенной суммой (смесью) гауссиан. При этом вес той или иной гауссианы определяется временем, в течение которого соответствующей гауссиане цвет присутствует на видеопоследовательности.

Отметим, что при выборе метода на основе смеси гауссовских распределений учитывались как собственные предварительные эксперименты, так и опыт использования указанного метода другими исследователями при решении рассматриваемой задачи [4, 5].

По завершении первого этапа метода формируется набор масок, соответствующих отдельным кадрам видеопоследовательности. При этом каждая маска отражает результат сегментации кадра на область переднего плана, соответствующую движущемуся человеку, и фона. Пример работы выбранного метода показан на рисунке 1.

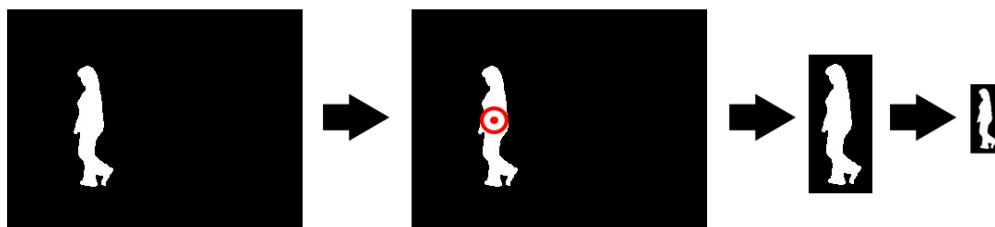


**Рисунок 1.** Выделение движущегося человека на видеопоследовательности: кадр исходной видеопоследовательности (слева), сформированная маска (справа) [10].

### 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 сжатию до заданных размеров. Описанная схема нормализации размера представлена на рисунке 2.

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 на третьем этапе работы производится снижение размерности данных, описывающих движение человека.



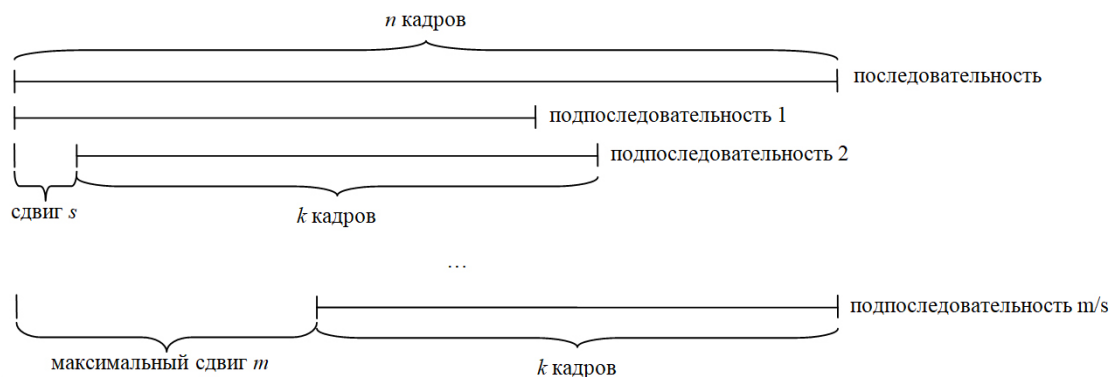
**Рисунок 2.** Нормализация размера выделенного фрагмента видеопоследовательности.

### 2.3. Снижение размерности с использованием метода главных компонент

Для снижения размерности многомерных данных используются как линейные, так и нелинейные методы. Наиболее часто используются линейные методы, такие как метод главных компонент (principal component analysis, PCA) [11] и анализ независимых компонент (ICA). Нелинейные методы снижения размерности (например, нелинейное отображение [12], ISOMAP [13], LLE [14]) используются реже в связи с высокой вычислительной сложностью таких методов.

В настоящей работе применяется метод главных компонент, как наиболее часто используемый в таких случаях (см., например [1, 4]). Этот метод осуществляет поиск линейной проекции в подпространство меньшей размерности, максимизирующей разброс данных. Метод PCA часто рассматривается, как линейный метод снижения размерности данных, минимизирующий потерю информации.

Перед снижением размерности для каждой последовательности кадров, содержащей движение, выделяется набор подпоследовательностей заданной (фиксированной) длины. Выделение подпоследовательностей осуществляется с некоторым заданным шагом, начиная с начала исходной последовательности (см. рисунок 3).



**Рисунок 3.** Параметры выделяемых подпоследовательностей.

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. 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) [15]. При этом классы соответствуют отдельным людям (личностям), а вектора признаков, полученные для всех соответствующих этим личностям 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 SVM.

### **3. Эксперименты**

Предложенный в работе метод был реализован на языке C++ с использованием библиотеки OpenCV. 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 CASIA GAIT [16]. 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 124 человек с различными вариациями угла съемки, стиля одежды и переноса сумок.

В настоящей работе из всего множества видеопоследовательностей были случайным образом взяты 15 классов, в которых угол съемки составляет 90 градусов, люди изображены в обычной не зимней одежде и без сумок. В каждом классе было 6 последовательностей. Длина каждой последовательности составляла не менее 60 кадров.

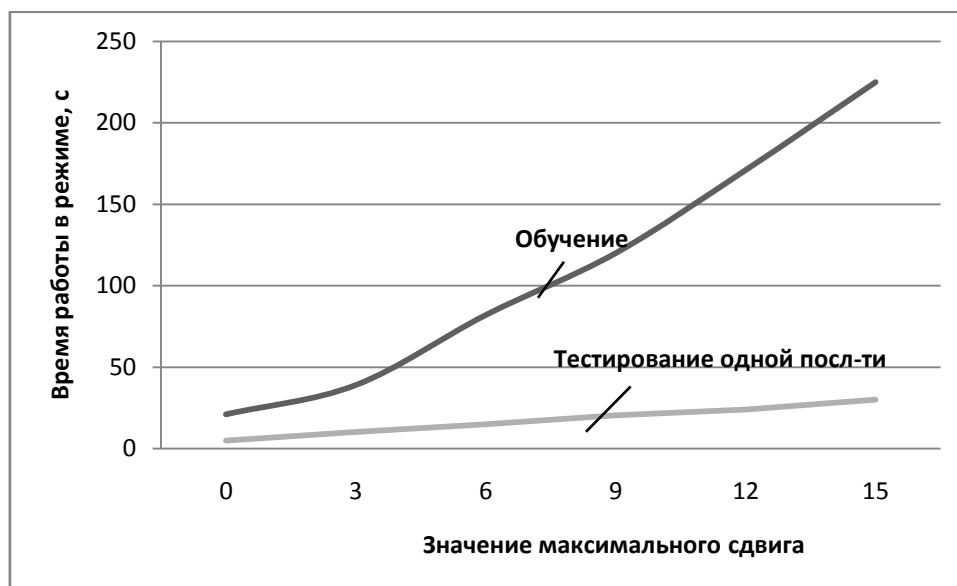
Классы были разделены на обучающую и тестовую выборки по 3 последовательности каждая. Последовательности каждого класса были предварительно обработаны с использованием алгоритма, описанного в п.2.2. Затем производилось снижение размерности данных с использованием метода, описанного в п.2.3 и классификация с использованием алгоритма, описанного в п.2.4. Для определения качества работы метода оценивалась точность классификации, определяемая как доля правильно классифицированных объектов. Во время проведения эксперимента также фиксировалось время, затрачиваемое в режиме обучения и тестирования.

В первом эксперименте исследовалась зависимость точности классификации от максимального сдвига подпоследовательностей от начала последовательностей (параметр  $m$  на рисунке 3). В рассматриваемом эксперименте для сокращения времени обучения используемый при выделении подпоследовательностей шаг сдвига  $s$  составлял 3 кадра. Таким образом, максимальный шаг подпоследовательностей  $m$  также изменялся с шагом 3, принимая значения от 0 до 15. Последнее значение определялось исходя из длины  $k$  формируемых подпоследовательностей, составляющей  $k=45$  кадров, и минимальной длины отображенных для экспериментов последовательностей, составляющей  $n=60$  кадров.

Результаты эксперимента представлены на рисунках 4 и 5. Экспериментально было определено, что точность классификации в целом возрастает при увеличении максимального сдвига. Это наблюдение вполне ожидаемо, так как чем больший максимальный сдвиг используется при формировании признаков, тем более «полным» становится признаковое описание видеопоследовательности. Указанный рост сопровождается и увеличением времени обработки, так как растет количество обрабатываемых подпоследовательностей.



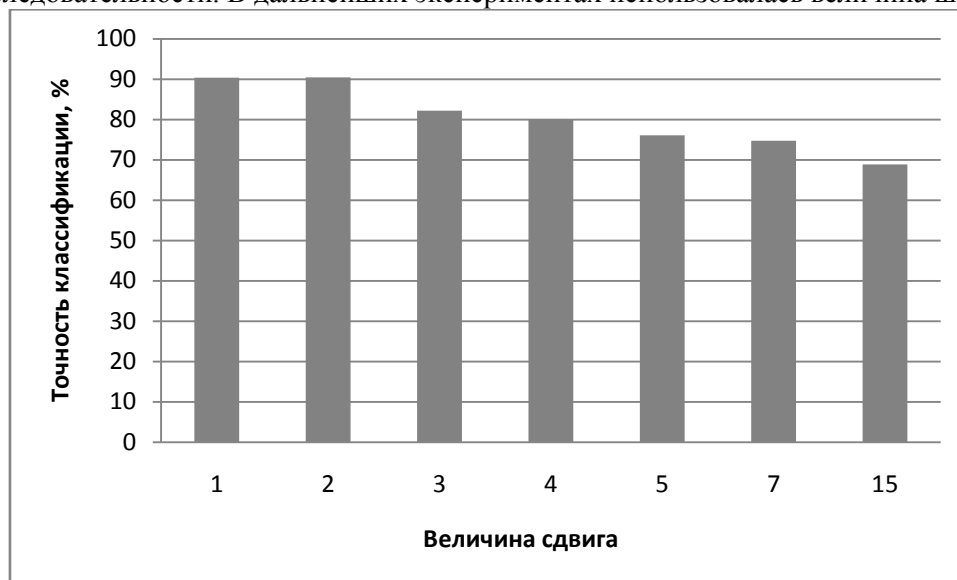
Рисунок 4. Зависимость точности классификации от максимального сдвига.



**Рисунок 5.** Зависимость времени работы в режимах обучения и тестирования от максимального сдвига.

Во втором эксперименте исследовалась зависимость точности классификации от величины шага  $s$ , используемого при выделении подпоследовательностей. При выполнении эксперимента максимальный сдвиг  $t$  был фиксирован и равен 15.

Результаты эксперимента представлены на рисунке 6. Экспериментально было определено, что наилучшая точность классификации достигается при малых значениях сдвига  $s$ . Это также хорошо согласуется с ожиданиями, так как чем меньший шаг сдвига используется при формировании признаков, тем более «плотным» становится признаковое описание видеопоследовательности. В дальнейших экспериментах использовалась величина шага 2.



**Рисунок 6.** Зависимость точности классификации от величины сдвига.

В третьем эксперименте исследовалась зависимость точности классификации от выходной размерности  $N$  векторов признаков, формируемых на шаге 2.3 рассматриваемого метода. При выполнении эксперимента выходная размерность варьировалась от 2 до 256, а другие параметры оставались фиксированными. В частности, шаг  $s$ , используемый при выделении

подпоследовательностей, составлял 2 кадра, а максимальный сдвиг  $m$  начала выделяемых подпоследовательностей - 15 кадров.

Результаты эксперимента представлены на рисунках 7 и 8. Как видно из приведенных результатов, наилучшие значения показателя точности классификации достигаются для размерностей 16 и 64. Рост размерности сопровождается ожидаемым возрастанием времени обработки, хотя в целом изменения не являются значительными. Тем не менее, более предпочтительным представляется выбор меньшей размерности ( $N=16$ ), обеспечивающей сравнимые показатели качества.

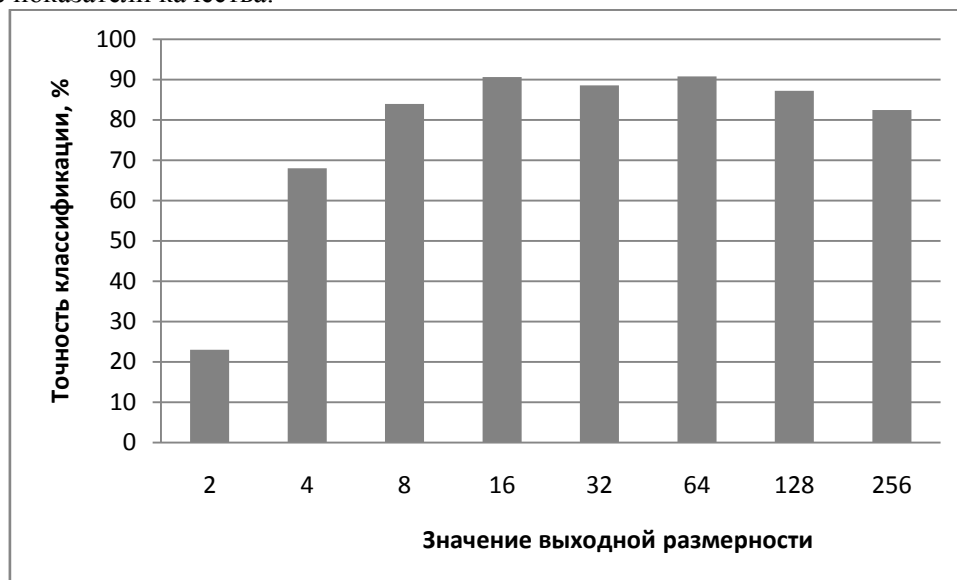


Рисунок 7. Зависимость точности классификации от размерности вектора признаков.

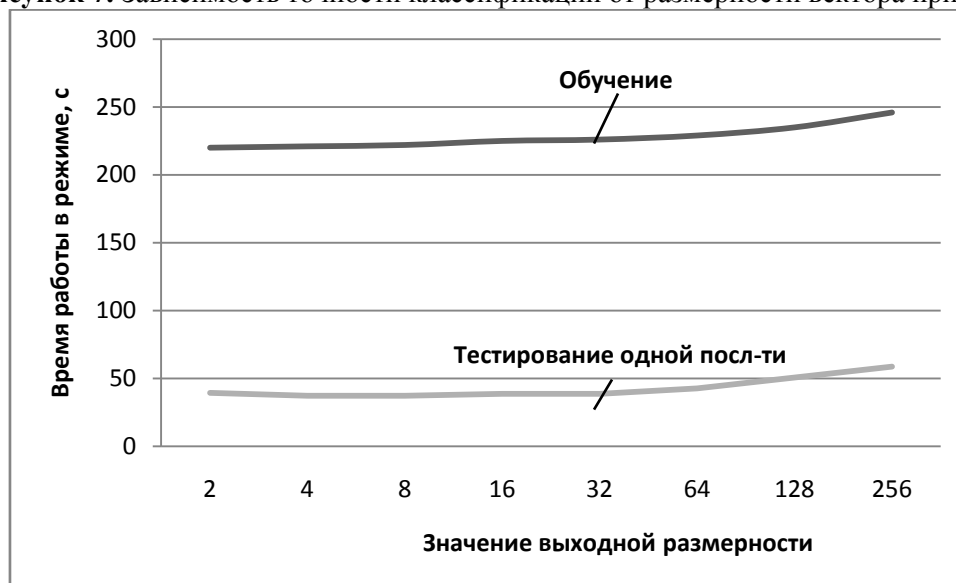


Рисунок 8. Зависимость времени работы в режимах обучения и тестирования от размерности вектора признаков.

В четвертом эксперименте рассматривалось разное количество классов (личностей). Эксперимент проводился для 5, 10, 15, 20 и 25 человек, а другие параметры оставались фиксированными, в частности, шаг  $s$ , используемый при выделении подпоследовательностей, составлял 2 кадра, а максимальный сдвиг  $m$  начала выделяемых подпоследовательностей - 15 кадров. Результаты эксперимента представлены на рисунках 9 и 10.

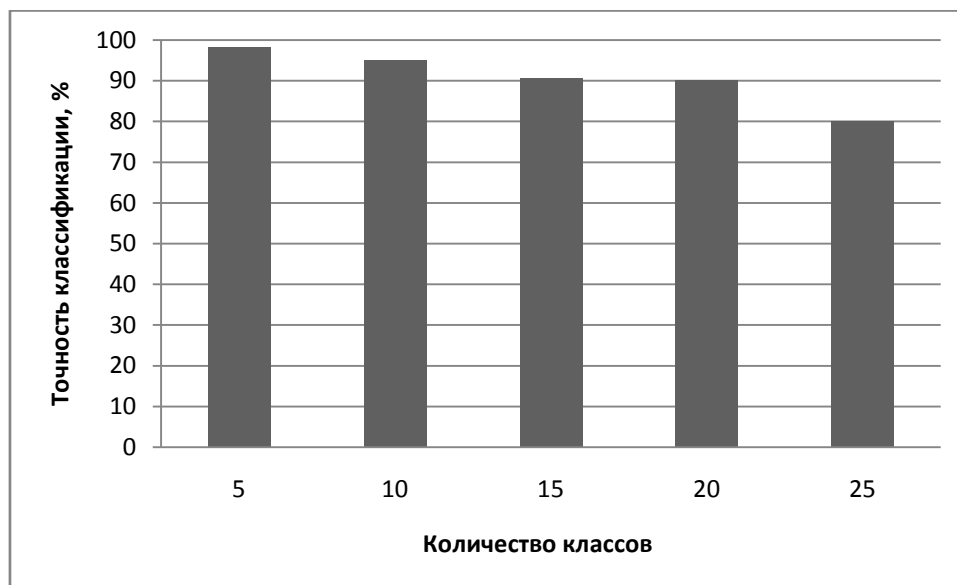


Рисунок 9. Зависимость точности классификации от количества классов.

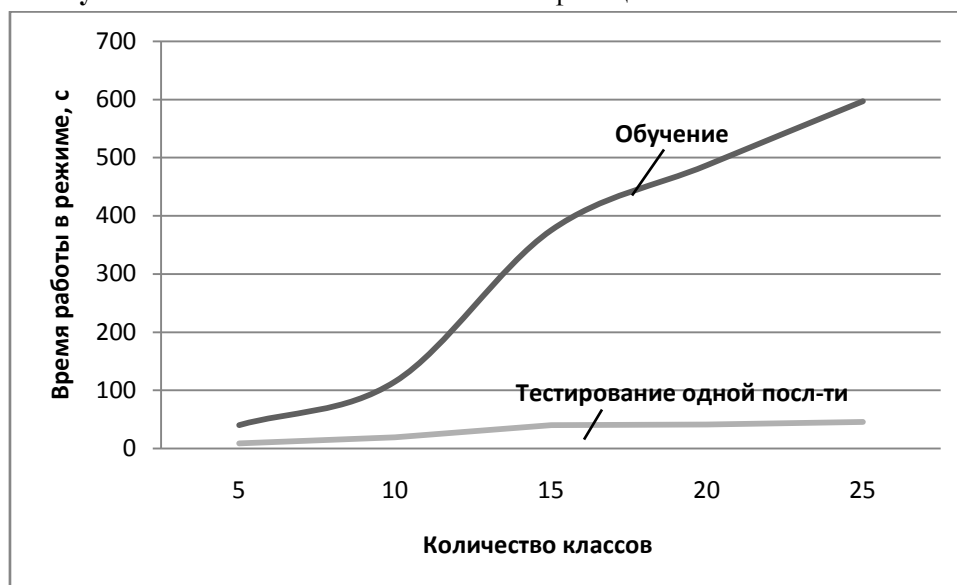


Рисунок 10. Зависимость времени работы в режимах обучения и тестирования от количества классов.

Как видно из результатов приведенных экспериментов, при относительно небольшом количестве классов (20 классов и менее) обеспечивается высокая (не менее 90%) точность классификации. В частности, при использовании 5 классов точность составила более 90%, при использовании 10 классов – 95%. При использовании 25 классов точность составила 80%.

Хотелось бы отметить, что прямое экспериментальное сравнение с работами других исследователей представляется достаточно сложной задачей в связи с разными используемыми наборами данных, а также потенциальными отличиями в условиях проведения эксперимента. Наиболее близкой к представленным в настоящей работе результатам представляется статья [6], в которой в зависимости от настроек используемого классификатора авторами заявлена точность от 92,08 до 98,79% для единственного рассмотренного случая с десятью объектами. Таким образом, можно говорить о том, что полученные результаты соответствуют современному состоянию работ в рассматриваемой области исследований.



Как видно из рисунка 9 с ростом числа классов обработка занимает все большее время. Значительное время при этом занимает обработка в режиме обучения, что становится особенно важным в тех сценариях использования, когда состав классов изменяется и требуется выполнять регулярное переобучение системы.

#### 4. Заключение

Предложенный в настоящей работе метод распознавания личности по походке состоит в выделении движущегося человека на видеоряде с последующей нормализацией размера, снижением размерности с использованием метода главных компонент и классификацией с использованием машины опорных векторов.

Выполненные на базе данных CASIA GAIT эксперименты позволили определить наилучшие значения параметров разработанного метода. Было показано, что при небольшом количестве классов (не более 20) предложенный метод обеспечивает на указанной базе данных высокую (не менее 90%) точность классификации личности. В частности, при использовании 5 классов точность составила более 98%, при использовании 10 классов – 95%.

В будущем планируется расширить как список алгоритмов, используемых для формирования признакового описания, так и список используемых методов классификации. Другим важным направлением исследований представляется снижение времени обработки.

#### 5. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-29-03190 офи\_м.

#### 6. Литература

- [1] Josiński, H. Feature Extraction and HMM-Based Classification of Gait Video Sequences for the Purpose of Human Identification / H. Josiński, A. Świtoński, A. Michalczyk, D. Kostrzewa, K. Wojciechowski // *Vision Based Systems for UAV Applications. Studies in Computational Intelligence*. – Heidelberg: Springer. – 2013. – Vol. 481.
- [2] Suutala, J. Gaussian Process Person Identifier Based on Simple Floor Sensors / J. Suutala, K. Fujinami, J. Röning // *European Conference on Smart Sensing and Context*. – 2008. – P.55-68.
- [3] Dingbo, D. Automatic Person Identification in Camera Video by Motion Correlation / D. Dingbo, G. Guangyu, H. Chi, J. Ma // *Hindawi Publishing Corporation Journal of Sensors*. – 2014. – Vol. 2014.
- [4] Murukesh, C. Secured Authentication through Integration of Gait and Footprint for Human Identification // C. Murukesh, K. Thanushkodi, P. Padmanabhan, Mohamed D. Feroze, Naina // *Journal of Electrical Engineering and Technology*, 2014.
- [5] Wang, L. Automatic Gait Recognition Based on Statistical Shape Analysis / L. Wang, T. Tan, W. Hu, H. Ning // *Transactions on image processing*. – 2003. – Vol. 12.
- [6] Shelke, P.B. Person Identification Using Gait: SVM Classifier Approach / P.B. Shelke, P.R. Deshmukh // *International Journal of Emerging Technologies and Engineering (IJETE)*. – 2014. – Vol. 1(10).
- [7] Kadew, P. An improved adaptive background mixture model for real-time tracking with shadow detection / P. Kadew, T.K. Pong, R. Bowden. – 2001.
- [8] Zivkovic, Z. Improved adaptive Gaussian mixture model for background subtraction / Z. Zivkovic. – 2004.
- [9] Godbehere, A.B. Visual Tracking of Human Visitors under Variable-Lighting Conditions for a Responsive Audio Art Installation / A.B. Godbehere, A. Matsukawa, K. Goldberg. – 2012.
- [10] Background Subtraction [Электронный ресурс]. – Режим доступа: [https://docs.opencv.org/3.3.0/db/d5c/tutorial\\_py\\_bg\\_subtraction.html](https://docs.opencv.org/3.3.0/db/d5c/tutorial_py_bg_subtraction.html) (13.11.2017).
- [11] Fukunaga, K. Introduction to Statistical Pattern Recognition / K. Fukunaga. – London: Academic Press, 2nd edn., 2003.

- [12] Sammon, J.W. A nonlinear mapping for data structure analysis / J.W. Sammon // Transactions on Computers. – 1969 – Vol. 18. - P. 401-409.
- [13] Tenenbaum, J.B. A Global Geometric Framework for Nonlinear Dimensionality Reduction / J.B. Tenenbaum, V. de Silva, J.C. Langford // Science. – 2000 – P. 2319-2323.
- [14] Roweis, S.T. Nonlinear Dimensionality Reduction by Locally Linear Embedding/ S.T. Roweis, L.K. Saul // Science. – 2000. – Vol. 290. – P. 2323-2326.
- [15] Cortes, C. Support-vector networks / C. Cortes, V. Vapnik // Machine Learning. – 1995 – Vol. 20(3). – P. 273-297.
- [16] База данных CASIA GAIT [Электронный ресурс]. – Режим доступа: <http://www.cbsr.ia.ac.cn/english/Databases.asp> (13.11.2017).

## Gait analysis for person recognition using principal component analysis and support vector machines

O.V. Strukova<sup>1</sup>, L.V. Shiripova<sup>1</sup>, E.V. Myasnikov<sup>1</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

**Abstract.** The paper is devoted to the problem of the recognition of a person by gait using a video recorded in the optical range. The method proposed in this paper consists in the detection of a moving person on a video sequence with the subsequent size normalization and dimensionality reduction using the principal component analysis technique. The person classification was carried out using the support vector machine. The experimental studies performed using the CASIA GAIT dataset allowed us to determine the best values of the method parameters. The obtained results showed that with a small number of classes, high classification accuracy can be achieved.

**Keywords:** gait analysis, person identification, dimensionality reduction, SVM.