

PAPER • OPEN ACCESS

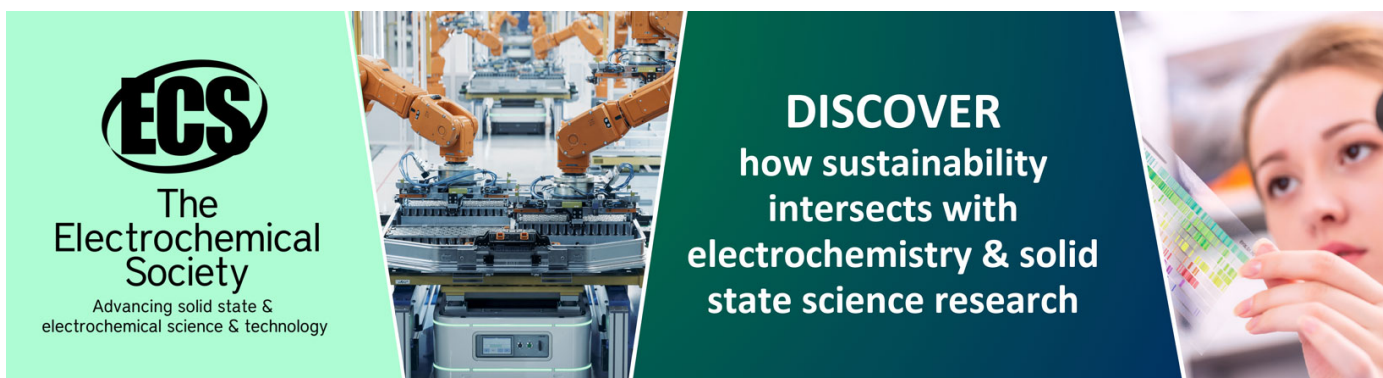
The choice of methods for the construction of PCA-based features and the selection of SVM parameters for person identification by gait

To cite this article: O V Strukova and E V Myasnikov 2019 *J. Phys.: Conf. Ser.* **1368** 032001

View the [article online](#) for updates and enhancements.

You may also like

- [Phonocardiography-based mitral valve prolapse detection with using fractional fourier transform](#)
Mahtab Mehrabbeik, Saeid Rashidi, Ali Fallah et al.
- [Dynamically weighted ensemble classification for non-stationary EEG processing](#)
Sidath Ravindra Liyanage, Cuntai Guan, Haihong Zhang et al.
- [Prediction of genetically-evaluated tumour responses to chemotherapy from breast MRI using machine learning with model selection](#)
Taiguang Yuan, Ze Jin, Yukiko Tokuda et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

The choice of methods for the construction of PCA-based features and the selection of SVM parameters for person identification by gait

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

e-mail: shnurocheck@gmail.com

Abstract. The paper presents the results of the continued research of the principal component analysis (PCA) and support vector machine (SVM) techniques for person identification by gait. The experimental studies performed using the CASIA GAIT dataset allowed us to compare two methods of work with input data for PCA. According to the results of experiments, the optimal method of forming the feature space and the most effective parameters of the SVM-classifier were selected. The classification of video sequences recorded by video cameras located frontally, at an angle and orthogonal to the direction of objects movement was carried out.

1. Introduction

In comparison with other widely used biometric systems, the identification of a person by gait is universal since it allows recognizing a person from a greater distance. The undoubted advantage of this approach is the fact that the process of identification of a person is possible without the knowledge and consent of a person. Attempts to hide a face from a surveillance camera will not interfere with the recognition of a person by gait.

However, this approach has some limitations, one of which is the change in the angle of shooting. The resistance of the personality recognition system to changes in shooting conditions is essential for widespread use of the considered technique in practice.

According to literature reviews, scientists carried out person identification by gait using various methods. In the paper [1] researchers discussed the accuracy dependence of the person classification in relation to the change the shooting angle based on data from OU-ISIR.

In modern literary sources, researchers increasingly refer to the support vector machine [2] as one of the effective classification methods. In the article [3] the efficiency of SVM classifier and the hidden Markov model (HMM) are compared.

The principal component analysis (PCA) is also often used in the tasks of human recognition by gait [4, 5]. The work [4] describes in detail the efficiency analysis of using linear (PCA) and nonlinear (ISOMAP, LLE) methods of dimensionality reduction for the formation of features.



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

As part of the previously proposed approach [6], we continued to search the optimal method of forming the feature vector of human movement and to select the best parameters of SVM-classifier to improve the classification efficiency. The studies were performed in the extended database. The results of these studies demonstrated the importance of the careful selection of characteristics for solving the problem, which can achieve high levels of classification quality.

The paper has the following structure. Section 2 is devoted to the description of the methods used in the paper. Section 3 describes the results of experiments. The paper ends up with the conclusion and reference list.

2. Methods

In brief, the methods developed in this study are based on our previous paper [6]. On the whole, both considered methods consist of the detection and segmentation of a moving person in a video sequence with the subsequent size normalization and dimensionality reduction of the selected sequence of frames. Finally, the classification of video sequences is performed to identify a moving person. A brief description of the above steps is given below.

2.1. Detection and segmentation of a moving person in the video sequence

At the first stage of the developed method, the moving person is extracted in the video sequence using an algorithm based on the mixture of Gaussian distributions (Gaussian mixture model, GMM) [7]. The result of the first stage is a set of masks corresponding to individual frames of the video sequence (figure 1).

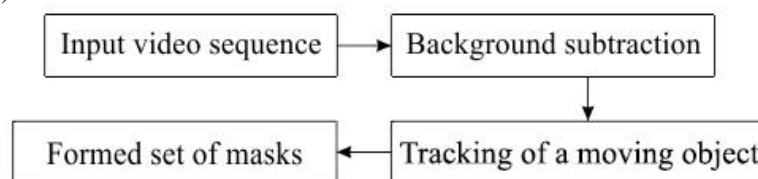


Figure 1. Detection of a moving person.

2.2. The frame size normalization of the selected video sequence fragment

At the second stage, the masks obtained at the previous stage are processed step by step according to figure 2.

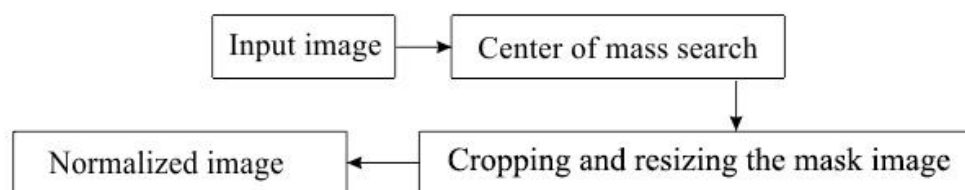


Figure 2. Normalization of the mask image.

2.3. Dimensionality reduction

The total dimensionality of one sequence of mask images is defined by $k \times w \times h$, where k is the number of images in a sequence, w and h are the width and height of the normalized image, correspondingly. As the dimensionality of a sequence remains high even after cropping and resizing, we apply a dimensionality reduction technique to acquire more compact representation.

The principal component analysis technique was chosen as the most commonly used method of this class [3-6], while other linear [8] and nonlinear [4, 9] methods exist. This method searches for a linear projection into a subspace of a smaller dimension maximizing the data variance.

According to our previous paper [6], we use the set of fixed-length subsequences of a sequence to achieve the time-invariant representation of a movement, and a feature vector is generated for each subsequence using PCA. In this paper, we compare two methods for the construction of PCA-based features, which are described below.

2.3.1. Dimensionality reduction of subsequences

According to the first method, each normalized frame of a subsequence expands into a vector and the vectors obtained for individual frames are concatenated. Each row of an input matrix for the PCA represents one subsequence of expanded and concatenated frames. The input matrix contains rows for all expanded subsequences of all sequences for all persons (classes) as it is shown in figure 3. The projection of input data onto the first N principal components is taken as a feature description after finding of the principal components.

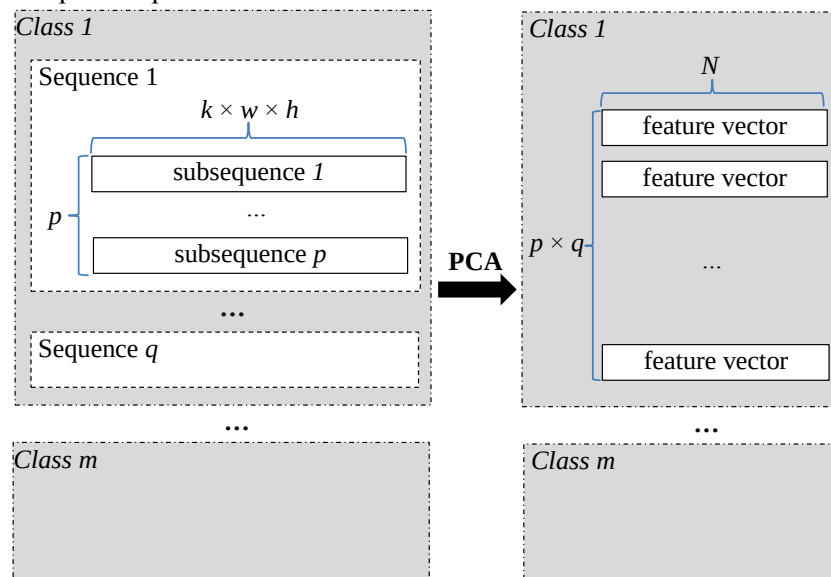


Figure 3. Method 1: dimensionality reduction of subsequences.

2.3.2. Dimensionality reduction of frames

In the second method, each frame is also normalized and expanded into a vector. But, contrary to the first method, we consider expanded frames of each subsequence as rows of input matrix for the PCA. So, the input matrix contains rows for all expanded frames of all subsequences generated for all sequences and all persons (classes) as it is shown in figure 4. Then we apply PCA and obtain the projection of expanded frames onto the first N principal components. Then we obtain the required feature descriptors by concatenating the rows of the output matrix corresponding to individual subsequences.

2.4. Classification of video sequences

The feature vectors obtained by one of the described above methods are used to train the support vector machine (SVM) [2, 10] classifier. In the considered case, the classes correspond to individual persons (individuals), and feature vectors obtained for all the subsequences correspond to individual observations (examples).

3. Experiments

The described above methods were implemented in C++ using the OpenCV library. We used a PC based on Intel Core i5-3470 CPU 3.2 GHz to perform experimental studies.

For the experimental study, the video sequences from the open CASIA GAIT dataset [11] were used. This dataset contains the sequences of binary images containing the silhouettes of moving persons.

For the experiments, we used three sets of video sequences with shooting from three cameras according to the direction of human movement: frontally (0 degrees), at an angle of 36 degrees and in the orthogonal direction (90 degrees). 25 classes were taken for each angle randomly. There were 6

sequences in each class. The length of each sequence was not less than 60 frames. Classes were divided into training and test samples of 3 sequences each.

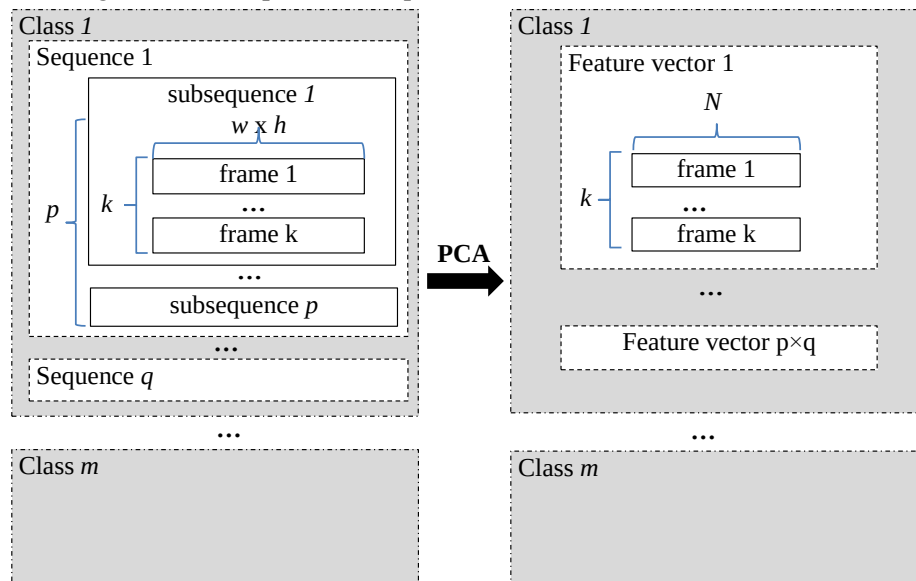


Figure 4. Method 2: dimensionality reduction of frames.

To estimate the quality of the considered methods, we used the classification accuracy, defined as the proportion of correctly classified sequences. The classification accuracy was calculated using the k-fold cross-validation.

In the first experiment, we compared two methods of work with input data for PCA described in sections 2.3.1 and 2.3.2. The experiment was carried out for 5, 10, 15, 20 and 25 classes, which were taken at the right angle. Other parameters remained fixed. In particular, the step s , with which the subsequences were extracted from a sequence, was equal to $s=2$ frames. The maximum shift m of the beginning of extracted subsequences relative to the beginning of a sequence was equal to $m=15$ frames. The optimal value of the output dimensionality for the principal component analysis technique was determined in our previous works [3, 6]. In all the experiments, the projection of feature vectors onto the first N principal components was taken as a feature description for the further classification. The results are shown in figures 5, 6.

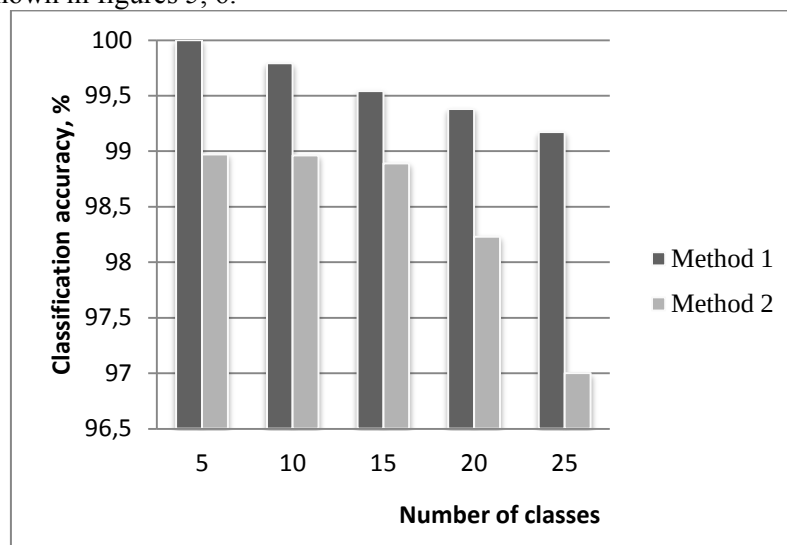


Figure 5. Dependence of the classification accuracy on the number of classes using the proposed methods described in 2.3.1 (Method 1) and 2.3.2 (Method 2).

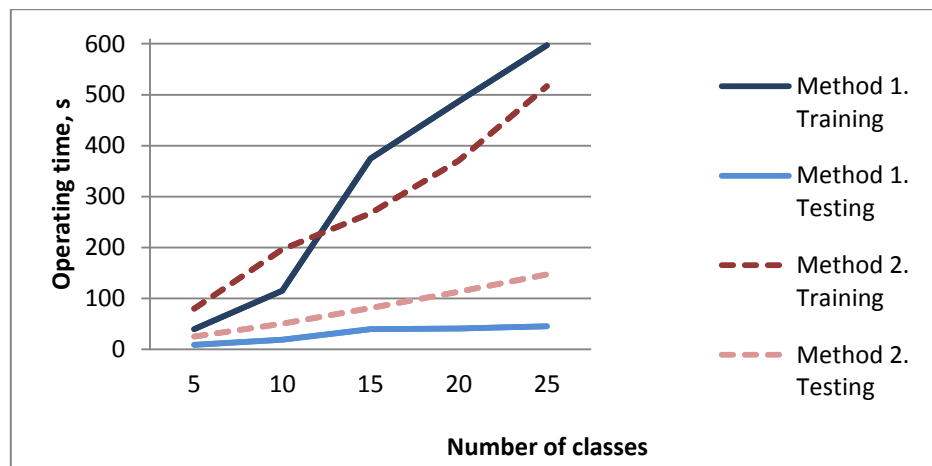


Figure 6. Dependence of training and testing time on the number of classes using the proposed methods described in 2.3.1 (Method 1) and 2.3.2 (Method 2).

As seen from the above results, the best classification accuracy was achieved using the method proposed in section 2.3.1. More than that, this method processed much faster in testing mode (figure 6).

Another research direction was the study of the parameters of the SVM classifier. In the experiments, we used video sequences taken from three shooting angles (0° , 36° , 90°). Table 1 shows the results of the classifier with LINEAR and INTERS (histogram intersection) kernels. The linear kernel is the scalar multiplication of any two given feature vectors. The histogram intersection kernel measures the similarity between the histograms describing the features of an object. Both kernels do not require additional configuration and optimization of classifier parameters. Therefore, linear and histogram intersection kernels are the fastest. It was experimentally determined that the classification accuracy is higher using the LINEAR kernel.

Table 1. Dependence of the classification accuracy on the number of classes and kernel type.

Number of classes	Shooting angle 90°		Shooting angle 36°		Shooting angle 0°	
	LINEAR	INTER	LINEAR	INTER	LINEAR	INTER
5	100.00	98.67	82.33	76.39	81.5	80.83
10	99.79	98.96	79.79	75.21	80.56	74.17
15	99.54	98.89	80	74.67	78.24	72.33
20	99.38	98.23	77.43	70	77.5	72.08
25	99.17	97.00	77.16	69.17	77.33	69.72

Similar experiments were carried out by the authors of the article [2], where only 10 objects with a right angle shooting were taken to train the parameters of the SVM classifier. The classification accuracy obtained in our research corresponds to the current state in the considered field of study. It exceeds the result demonstrated in [2] for experiments with similar conditions (number of classes, type of core, database, and angle of shooting).

4. Conclusion

In the framework of our previously developed approach [19, 21], we proposed and compared two methods for the formation of input data for the principal component analysis technique. The

experiments performed on the CASIA GAIT database showed that the method based on the dimensionality reduction of subsequences outperforms the method based on the dimensionality reduction of frames both in classification accuracy and testing time.

We compared two types of kernels for the support vector machine classifier and achieved 99.17% accuracy for 25 classes at 90° shooting angle. The experiments with different shooting angles showed that the accuracy reduces when shooting in nonorthogonal directions (0° and 36°).

The possible direction of further research is the study of nonlinear kernels.

5. References

- [1] Takemura N, Makihara Y, Muramatsu D, Echigo T and Yagi Y 2018 Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition *IPSN Transactions on Computer Vision and Applications* **10(4)**
- [2] Shelke P B and Deshmukh P R 2014 Person Identification Using Gait: SVM Classifier Approach *International Journal of Emerging Technologies and Engineering (IJETE)* **1(10)**
- [3] Shiripova L and Myasnikov E 2018 Comparative analysis of classification methods for human identification by gait *CEUR Workshop Proceedings* **2268** 118-128
- [4] Josiński H, Świtoński A, Michalczyk A, Kostrzewa D and Wojciechowski K 2013 Feature Extraction and HMM-Based Classification of Gait Video Sequences for the Purpose of Human Identification *Vision Based Systems for UAV Applications. Studies in Computational Intelligence* (Heidelberg: Springer) **481** 233-245
- [5] Murukesh C, Thanushkodi K, Padmanabhan P and Feroze Naina Mohamed D 2018 Secured Authentication through Integration of Gait and Footprint for Human Identification *Electrical Engineering and Technology* **9(6)** 2118-2125
- [6] Strukova O V, Shiripova L V and Myasnikov E V 2018 Gait analysis for person recognition using principal component analysis and support vector machines *CEUR Workshop Proceedings* **2210** 170-176
- [7] Zivkovic Z 2004 Improved adaptive Gaussian mixture model for background subtraction *Proc. of the 17th Int. Conf. on Pattern Recognition* Cambridge, UK (IEEE) **2** 28-31
- [8] Dmitriev E A and Myasnikov V V 2018 Comparative study of description algorithms for complex-valued gradient fields of digital images using linear dimensionality reduction methods *Computer Optics* **42(5)** 822-828 DOI: 10.18287/2412-6179-2018-42-5-822-828
- [9] Myasnikov E V 2018 Evaluation of Nonlinear Dimensionality Reduction Techniques for Classification of Hyperspectral Images *CEUR Workshop Proceedings* **2268** 147-154
- [10] Cortes C and Vapnik V 1995 Support-vector networks *Machine Learning* **20(3)** 273-297
- [11] URL: <http://www.cbsr.ia.ac.cn/english/Databases.asp>

Acknowledgments

The work was partly funded by RFBR according to the research project 17-29-03190 in parts of «1. Introduction» – «2. Methods» and by the Russian Federation Ministry of Science and Higher Education within a state contract with the "Crystallography and Photonics" Research Center of the RAS under agreement 007-Г3/Ч3363/26 in part of «3. Experiments».