MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
SUMY STATE UNIVERSITY
UKRAINIAN FEDERATION OF INFORMATICS

PROCEEDINGS

OF THE V INTERNATIONAL SCIENTIFIC CONFERENCE

ADVANCED INFORMATION SYSTEMS AND TECHNOLOGIES

AIST-2017

(Sumy, May 17–19, 2017)

SUMY
SUMY STATE UNIVERSITY
2017

Gait recognition on base of representation in spatio-temporal area

Mihail Babiy
Sumy State University, mbabiy@id.sumdu.edu.ua

***Abstract* – New method of automatic gait recognition from video is proposed. The method works with a sequence of silhouettes derived from the video after background subtraction, decreasing shadows and noise. Two-dimensional silhouette shape is converted into one-dimensional signal presenting distance from center of gravity to outline of this silhouette. A set of signals extracted from a sequence of silhouettes forms a two-dimensional picture.**

The features extraction is performed using Gabor wavelets. Testing the method on the samples of CASIA gait database showed high recognition accuracy.

***Keywords* – gait, recognition, Gabor wavelets, CASIA database.**

I. INTRODUCTION

Automated person identification is important research problem in the area of computer vision. Traditional face recognition technology is applicable only to frontal or nearly frontal faces. In addition, the recognition capabilities are very limited at night, when only silhouette of a person can be observed. In these cases, gait recognition can be a promising method.

Gait is a spatio-temporal process that typifies the motion characteristics of an individual. As biometric characteristic, gait has next advantages: (1) invisibility for the object under surveillance, (2) no need to improve the quality of the image, (3) the complexity of disguise, (4) the possibility of identifying at a significant distance from the place of observation, when other attributes are still indistinguishable.

There are two main ways to represent gait in recognition problems: model-based approach and model-free approach. In the first case, static and dynamic body parameters are determined in modeling the movement of the arms, legs and other body components. However, model-based approach is sensitive to the quality of the gait image sequences. Another disadvantage is large amount of computation.

Model-free approach focus on shape of silhouette. This approach is insensitive to the quality of silhouettes. It requires less computation and is more often used in practice.

In the framework of this approach, some authors apply the hidden markov models (HMM). Iwamoto et al. [1] regards the walking operation as a periodic signal. A period (gait cycle) is defined as the time of heel strike for the same leg. Person's outline information in each frame is used as the feature parameter. The feature vector is compressed by principle component analysis (PCA). Training and identification processes are performed using continuous HMM

In [2] the discrete HMM is taken into consideration for recognition process. Vector quantization and clustering are used to convert continuous sequence of gait signals into a discrete sequence of symbols for discrete HMM. A general drawback of this approach is ambiguity in specifying the number of clusters and the number of hidden states.

Topological approach is used in [3] and [4]. 3D digital picture is obtained by gluing silhouettes through their gravity centers. A border simplicial complex associated with picture is constructed. Authors define eight directions to obtain eight filtrated simplicial complex. According to each filtration a persistence barcodes are calculated. Using barcodes the similarity value for two gait subsequences can be computed. A triangulation is supported by existing software, but number of required triangles is very large and the computational analysis by topological approach is correspondingly slow.

II. CREATING THE GAIT IMAGES

The proposed algorithm includes next phases: silhouette estimation, transformation of gait sequence to gait picture, feature extraction using wavelets, near-neighbor gait recognition "Fig. 1".

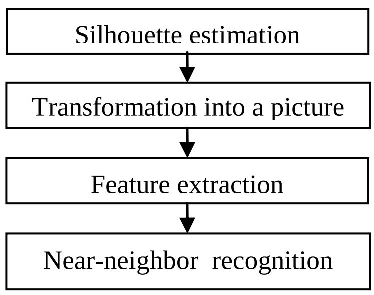

Figure 1. Flowchart of algorithm

Background modeling uses median method. This method incorporates each foreground pixel so that each occurrence of foreground objects leads to corruption of background image. To obtain the silhouette, background subtraction and binarization is performed.

At the next stage, the gravity center of silhouette is determined. By choosing the gravity center as reference origin, we unwrap the outer contour clockwise. For this purpose, we find distances from the gravity center to the contour points lying on the rays emerging from this center “Fig. 2”.

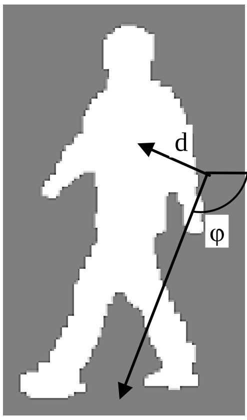

Figure 2. Determining distances from the gravity center to the outer contour of silhouette

In our research, the angle between neighboring rays $\Delta\varphi = \pi / 36$. As a result, the distance signal $D = \{d_1, d_2, \ldots, d_{72}\}$ is obtained. This signal indirectly represents the initial 2D silhouette shape into 1D space.The algorithm for determining a distance is as follows. Starting from the gravity center, we move along the ray with increment equal to one. The following actions are performed at each step. Using the current coordinates, coordinates of the neighboring pixel $\{x_p, y_p\}$ are calculated as $x_p = \text{floor}(x)$, $y_p = \text{floor}(y)$. If the brightness of the pixel $\{x_p, y_p\}$ becomes less than the threshold value, the distance is taken equal to the number of steps. Otherwise the movement along the ray continues. For noisy silhouettes, other algorithm can be used. The motion is performed to the image boundary. The distance is defined as the number of steps, by which the brightness of the current pixel is greater than the threshold value.Wang and Tan [5] indirectly use obtained signals for person identification. They apply PCA training to represent the original gait features from high-dimensional measurement space to a low-dimensional eigenspace. Nonzero eigenvalues and the associated eigenvectors based on singular vector decomposition (SVD) are calculated. After that an original distance signal can be projected into a point in the eigenspace.

Other approach is used in this article. For each frame in the time sequence of the silhouettes, a corresponding 1D signal is constructed. We introduce the coordinate s, representing the numbers of the border points in which the distance signal is measured. By grouping the silhouettes at sequenced times t, the gait can be represented as a graph of the signal D in the spatio-temporal area (s, t). We convert this graph into the normalized grayscale image $I(s, t)$ by formula

$$I(s,t) = \frac{255D(s,t)}{d_{\max}},$$

where $d_{\max}$ is the maximal value of $D(s, t)$.

In the following text, instead of the notations (s, t), we will use the more habitual symbols (x, y) to address the image points. We will call these images “gait images”. Examples of images are shown in “Fig. 3”.

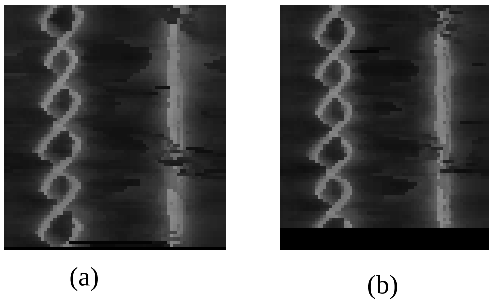

Figure 3. Examples of gait images from CASIA database: (a) gait fyc/00_1; (b) gait fyc/00_3.

III. Feature Vector Extraction

Gait images are converted to standard size 128*128 pixels. Feature vector is constructed from the responses of the gait image to Gabor filters.

Two-dimensional Cabor wavelet has the shape of a plane wave bounded in amplitude by a Gaussian function. Each wavelet ψ_j from a given set of wavelets is determined by its wave vector $\vec{k}_j$ [6]:

$$\psi_j(\vec{r}) = \frac{k_j^2}{\sigma^2}\exp(-\frac{k_j^2 r^2}{2\sigma^2})\left[\exp(i\vec{k}_j\vec{r}) - \exp(-\frac{\sigma^2}{2})\right], \quad (1)$$

where $\vec{r} = \vec{r}(x, y)$. We define the vector $\vec{k}_j$ for five different frequencies with indices $p = 0, \ldots, 4$ and eight different orientations with indices $q = 0, \ldots, 7$:

$$\vec{k}_j = (k_{jx}, k_{jy}) = (k_p \cos\varphi_q, k_p \sin\varphi_q);$$

$$k_p = 2^{-\frac{p+2}{2}}\pi; \quad \varphi_q = q\frac{\pi}{8}; \quad j = q + 8p.$$

Our image is given by array $I(\vec{r})$ of pixel brightness in the points $\vec{r}(x, y)$. Wavelet transform of image can be represented as convolution $R_j(\vec{r}_0)$ of image with j-th wavelet from the set of Gabor wavelets:

$$R_j(\vec{r}_0) = \int I(\vec{r})\,\psi_j(\vec{r} - \vec{r}_0)\,dxdy. \quad (2)$$

The feature vector for the image will be constructed from the values of R_j at lattice nodes with a step of 8 pixels horizontally and vertically. This step is quite acceptable for images with a side length of about a hundred pixels.To enter the metric, we agree to consider the distance between neighboring pixels equal to one. Then, if we choose σ = 2π, then for the wavelet with the index p = 2 the standard deviation of the Gaussian function is $\sigma / k_p = 8$, which is equal to the distance to the next lattice node. Accordingly, for p = 0, the deviation will be two times smaller, and for p = 4 it will be twice as large. Direct calculation of the convolution

$$R(\vec{r}) = I(\vec{r}) * \psi(\vec{r}) \tag{3}$$

is quite laborious. We apply the discrete Fourier transform F to both parts of (3). As a result, the convolution operation is transformed into multiplication of spectra

$$F[R(\vec{r})] = F[I(\vec{r})]\ F[\psi(\vec{r})] \tag{4}$$

Performing the inverse Fourier transform, we obtain

$$R(\vec{r}) = F^{-1}\{F[I(\vec{r})]\ F[\psi(\vec{r})]\} \tag{5}$$

The discrete Fourier transform for the two-dimensional case has the form

$$f_{k_x k_y} = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} x_{n_x n_y} \times$$

$$\times \exp\left(-\frac{2\pi i}{N_x} k_x n_x\right) \exp\left(-\frac{2\pi i}{N_y} k_y n_y\right)$$

For its implementation, it is convenient to use the Fast Fourier Transform (FFT) algorithm. The processing of graphic images is carried out using the non-commercial OpenCV library. It includes a function *cvDFT*, which implements the algorithm for direct and inverse FFT, including for two-dimensional input data. A trait of the FFT is the reduction of the problem for N numbers to the problem for $N1 = N / m$, where m is the divisor of the number N. In the implementation of *cvDFT*, the optimal size of the image side is the product of integer powers of numbers 2, 3 and 5. Accordingly, the size of the side of the image is equal to such a product. To reduce computational complexity, we confine ourselves to the real (even) component of the Gabor wavelet. In this case, the amount of input data is reduced by half, and for complex output data, a special packed format is used, which also reduces the output data by half.To relate an object to a particular recognition class, the nearest-neighbor method is used.

IV. Experimental Results

In accordance with the described algorithm, a program for gait recognition is developed. The program is written in C++ for the Visual Studio environment. Computer image processing is performed using the additional OpenCV library compiled for Visual Studio. For the testing of the program, CASIA Gait Database (DatasetA) was taken [7]. DatasetA includes gait samples for 20 persons. Each person has 12 image sequences, 4 sequences for each of the three directions. The lengths of the sequences are not identical due to the different walker's speed. For testing, the first three persons selected in alphabetical order were taken. The format of the image filename in DatasetA is “xxx-mm_n-ttt.png”, where xxx – subject id, mm – direction, n – sequence number, ttt – frame number in a sequence. As a measure of closeness, the similarity function was taken. The results of the calculation are presented in the Table I.

TABLE II. Results of Recognition

Test set	Values of the similarity function for the training set		
	fyc/00_1	hy/00_1	ljg/00_1
fyc/00_3 (fyc-00_3-001.png – fyc-00_3-069.png)	<u>0,893</u>	0,874	0,865
hy/00_3 (hy-00_3-001.png – hy-00_3-060.png)	0,817	<u>0,907</u>	0,867
ljg/00_3 (ljg-00_3-001.png – ljg-00_3-069.png)	0,866	0,851	<u>0,869</u>

Thus, all three gaits are correctly recognized.

V. Conclusions

In this paper we propose the new method of automatic gait recognition from video. Two-dimensional silhouette shape is converted into one-dimensional signal. A set of signals extracted from a sequence of silhouettes forms a two-dimensional picture. The features extraction is performed using wavelet analysis. Testing the method on the samples of CASIA gait database showed high recognition accuracy.

References

[1] K. Iwamoto, K, Sonobe, N. Komatsu, “A gait recognition method using hmm,” *Proceeding of the SICEAnnual Conference Roy. Soc. London,* August 2003, pp. 1936–1941,

[2] A. Kolawole, A. Akintola, “A novel gait recognition system based on hidden markov models,” *Proceedings of the 8th International Symposium on Visual Computing*, 2012.

[3] J. Lamar-Leon, E. Garcia-Reyes, R. Gonzalez-Diaz, “Human gait identification using persistent homology,” *CIARP*, *Lecture Notes in Computer Science*, vol. 7441, pp. 244–251, Springer, 2012.

[4] J. Lamar, E. Garcia-Reyes, R. Gonzalez-Diaz, R. Alonso-Baryolo, “An application for gait recognition using persistent homology,” Univ. of Seville, Spain, 2013.

[5] L. Wang, T. Tan, “Silhouette analysis-based gait recognition for human identification,” *IEEE Trans. on pattern analysis and machine intelligence*, vol. 25, no. 12, pp. 740–741, December 2003.

[6] L. Wiskott, J. M. Fellous, N. Kruger, C. Malsburg, “Face Recognition by Elastic Bunch Graph Matching,” *IEEE Trans. on pattern analysis and machine intelligence*, 19(7), pp.775-779, 1997.

[7] CASIA Gait Database (DatasetA) http://www.cbsr.ia.ac.cn/GaitDatasetA-silh.zip