OPTIMIZATION OF GABOR WAVELETS FOR FACE RECOGNITION

K.Murygin

Abstract: *The article describes researches of a method of person recognition by face image based on Gabor wavelets. Scales of Gabor functions are determined at which the maximal percent of recognition for search of a person in a database and minimal percent of mistakes due to false alarm errors when solving an access control task is achieved. The carried out researches have shown a possibility of improvement of recognition system work parameters in the specified two modes when the volume of used data is reduced.*

Key words: *person recognition, Gabor wavelets.*

Introduction

Now the methods based on Gabor wavelets, draw more and more attention of the researchers engaged in image recognition, including face recognition. One of explanations of growing popularity of the given approach are results of biologists researches, shown similarity of two-dimensional Gabor kernels with the form of receptor field of visual cells in the primary visual cortex [1,2]. Besides the great positive experience of Gabor filters use in tasks connected with person recognition by face image has been already saved up [3,4,5,6,7,8,9].

Gabor functions [6,8,9] are localized in spatial and frequency area and look like a plane wave with a wave vector $\vec{k}$, restricted by a Gaussian envelope function with width σ / k, where $\sigma = 2\pi$:

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

The normalizing factor, the second exponent in square brackets, is received from a condition of equality to integral zero:

$$\int \psi_j(\vec{x}) d^2\vec{x} = 0 \quad (2)$$

The following wavelet transformation gives complex factors that then are used as elements of feature vectors, describing the initial image $I(\vec{x})$ in a point $\vec{x}$:

$$J_j(\vec{x}) = \int I(\vec{x}')\psi_j(\vec{x} - \vec{x}')d\vec{x}' \quad (3)$$

Complex factors $J_j(\vec{x})$ can be written down as $J_j(\vec{x}) = a_j(\vec{x})\exp\left(i\varphi_j(\vec{x})\right)$ where $a_j(\vec{x})$ - is slowly varying amplitude, and the phase $\varphi_j(\vec{x})$ changes with characteristic frequency of corresponding Gabor kernel (1). Factors $J_j(\vec{x})$ make sense, similar to factors of Fourier transformation. However, as functions (1) are localized in spatial area, $J_j(\vec{x})$ characterize not the image $I(\vec{x})$ completely, but its some part, which size is defined by parameter $\frac{k_j^2}{\sigma^2}$, and position – by argument $\vec{x}$.

Vectors of features J_j and J'_j received on the basis of expression (3) are convenient for comparing with the use of a similarity measure as a corner cosine between them:

$$S(J_j, J'_j) = \frac{\sum_j J_j J'_j}{\sqrt{\sum_j J_j^2 \sum_j J_j'^2}} \quad (4)$$

The combination of metrics (4) and condition (2) allows excluding influence of any linear transformations of initial images on result of comparison. Thus, influence of brightness and contrast of initial images on result of comparison is minimized.

Each person entered into a database, is represented as set of local feature vectors received in beforehand-defined points of a face. The chosen set of such points in aggregate with the feature vectors received in them refers to a face graph (see fig. 1).

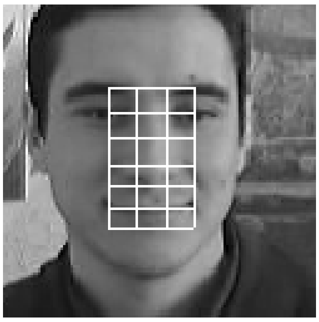

Figure 1 – Arrangement of face graph points

Automatic recognition with use of face image graphs is divided into two stages. At the first stage graph positioning on the image is carried out. On the second – comparison of the found graph with the graphs of persons kept in a database is carried out. Graph comparison is made by summation of measures of matching values of corresponding feature vectors received according to the formula (4), on all nodes of the graph.

At the chosen configuration the face graph (an arrangement of graph points) at both stages of recognition automatic system work, a question on a choice of a set of scales of used Gabor wavelets, which provide peak efficiency of algorithms work raises.

In works [6,8,9] for person recognition by face image Gabor functions of five various scales $\nu = \{0,...,4\}$, and eight orientations $\mu = \{0,...,7\}$ were used. Each function was determined by the following characteristic wave vector:

$$\vec{k}_j = \begin{pmatrix} k_\nu \cos\phi_\mu \\ k_\nu \sin\phi_\mu \end{pmatrix}, \quad k_\nu = 2^{-\frac{\nu+2}{2}}\pi, \quad \phi_\mu = \mu\frac{\pi}{8}, \qquad (5)$$

where index $j = \mu + 8\nu$. Thus, full wavelet transformation (3) gives 40 complex factors in each image point (5 scales and 8 orientations). The used set of scales of Gabor functions were offered by authors of work [7] who have carried out researches for two values of the spatial factor f determining distance between the next scales of Gabor kernels, and various values of parameter k_{max} determining the maximal scale of used kernels. Researches have shown, that optimum values are $f = \sqrt{2}$ and $k_{max} = \pi/2$. Thus, the used range of spatial frequencies is wide enough (T = 4 – 16 pixels). In work [10] the analysis of used scales of Gabor functions for reduction of number of used kernels has been carried out at preservation of an overall performance in application to a task of localization of person face graph points. Optimum results are received at use of one scale and eight orientations. And the spatial frequencies corresponding to the given scale are small enough (T = 16 – 25 points, depending on the chosen point of face graph and at radius of an eye pupil of 4 pixels).

In the present work, researches similar with described in [10] were carried out with the purpose of reception of an optimum set of Gabor functions for the decision of a recognition task. At carrying out of experiments were used some combinations of scales suggested in work [7] and used in works [6,8,9] described above.

Carried out researches

For researches the database of 160 face images of 10 persons (database of Weizmann Institute of Science∗) was used. All photos have been received at various illuminations. For each of faces key points (the centers of pupils, the centers of eyebrows above pupils, a tip of a nose and the edge of lips) were manually marked in accordance with which face graph was built automatically, and face properties in each graph point as feature vectors received on a basis of (3) and (5) were remembered. Graph configuration described in works [6,8,9] (see fig. 1) was used. The sizes of all face images have been preliminary resulted in one scale. The interpupil distance has been chosen for a scale factor. The reference scale was equal 34 pixels.

Researches were carried out in two directions. First, dependence of efficiency of search in a database on scales of used Gabor kernels was investigated. The given task is characterized by absence of necessity to analyze mistakes connected to face admission as the most similar accordance to the input image is searched in a database. Recognition is considered correct if the found image, most similar to the input face image belongs to the same person. Thus, optimization criterion is the percent of correct recognition. As reference accidentally one of 16 images for each person was chosen. In total 16 experiments were carried out. The estimation of efficiency was calculated as the average efficiency received in all experiments. For each of ranges $\nu = \{0,...,4\}$ the results given in table 1 have been received.

Table 1 – Dependence of percent of correct recognition on used scale sets

Used Gabor functions scales					Recognition percent, %
0	1	2	3	4	
*					74.8
	*				80.8
		*			75.7
			*		64.6
				*	62.0
*	*				86.7
	*	*			84.9
		*	*		76.1
			*	*	70.5
*	*	*			87.5
	*	*	*		82.0
		*	*	*	76.6
*	*	*	*		82.8
	*	*	*	*	80.9
*	*	*	*	*	81.4

The received data testify that the best results for search in a database give scales $\nu = \{0,1,2\}$, and it's better to use all three scales that give the maximal percent of recognition – 87.5%. At use of all five scales efficiency of recognition falls up to 81.4% that's possible to explain by strong dependence of the big scales on conditions of illumination. Use of only 0 and 1 scales reduces efficiency (in comparison with maximal) less than on 1 % at reduction of volume of remembered and analyzed data more than on 30 %. It allows using only these scales in systems, critical to volume of the remembered data and an operating time of recognition algorithm.

The second direction of researches has been connected to access control task. For an estimation of recognition efficiency for each of researched scales sets, functions of density of distribution of graph comparison results (see fig. 2) have been calculated. In figure 2, more to the left, function of distribution density, received at comparison of different people faces is shown. More to the right received for face images of one person. The square of crossing area, characterizes the minimal total mistake of recognition, which is achieved at a choice of a threshold of recognition equal abscissa of crossing points of distribution density functions. The data received at a similar choice of recognition threshold, are resulted in table 2.

∗ the database is received from a server ftp://eris.wisdom.weizmann.ac.il/pub

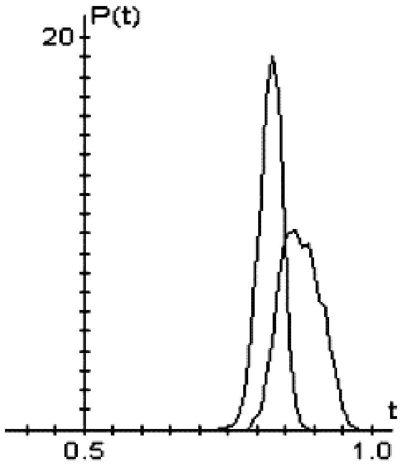

Figure 2 – Distribution density functions of graph comparison results

Table 2 – Dependence of mistakes of 1 and 2 sorts on used Gabor functions scale sets

Used Gabor functions scales					Face admission	Wrong recognition	Total mistake
0	1	2	3	4			
*					30.83	8.03	38.86
	*				22.67	14.85	37.52
		*			24.00	20.07	44.07
			*		44.75	7.22	51.97
				*	51.25	6.26	57.51
*	*				27.50	6.40	33.90
	*	*			29.08	9.41	38.49
		*	*		32.83	13.20	46.03
			*	*	43.00	7.67	50.67
*	*	*			19.33	14.24	33.57
	*	*	*		27.08	13.62	40.70
		*	*	*	35.83	11.48	47.32
*	*	*	*		25.67	11.19	36.86
	*	*	*	*	29.58	13.72	43.30
*	*	*	*	*	26.33	14.46	40.79

The analysis of received data, has shown, that the best results by a total mistake are observed at use of the same scales, as in experiments on search in a database. However in a case of access control task the mistakes connected to wrong recognition (it is equivalent to an opportunity of the non-authorized access) have the great importance. Therefore optimum scales are necessary for choosing, mainly being based on the data received for these mistakes.

If the percent of allowable mistakes of recognition is rigidly determined, parameter of optimization becomes frequency of recognition, which is determined by mistakes connected to face admission. In table 3 the experimental data received at aprioristic setting of an allowable mistake of recognition are given.

Table 3 – Dependence of mistakes of face admission on used Gabor functions scales and restrictions for a recognition mistake

The used Gabor functions scales					Face admission (%) at an allowable recognition mistake:				
0	1	2	3	4	0 %	1 %	2 %	3 %	4 %
*					75	54	43	43	43
	*				78	53	44	44	44
		*			77	57	57	49	49
			*		79	59	52	52	52

The used Gabor functions scales					Face admission (%) at an allowable recognition mistake:				
0	1	2	3	4	0 %	1 %	2 %	3 %	4 %
				*	79	63	57	57	57
*	*				69	38	38	38	28
	*	*			72	56	48	48	39
		*	*		69	55	47	47	47
			*	*	75	56	56	50	50
*	*	*			76	50	40	40	29
	*	*	*		67	52	45	45	45
		*	*	*	69	56	51	51	44
*	*	*	*		61	44	44	36	36
	*	*	*	*	67	53	46	46	46
*	*	*	*	*	67	52	45	45	36

As the results given in table 3 show, use of only 0 and 1 scales is also optimum for a task of access control.

Conclusion

The carried out experiments have shown, that reduction of a number of Gabor functions scales from 5 up to 2 (use of only 0 and 1 scales) improves parameters of recognition system work as for task of search in a database, so for access control task. Thus the volume of the remembered and analyzed data is essentially reduced, speed of work of algorithm of face comparison raises.

References

1. C. PJ. Jones, L. Palmer An evaluation of the two-dimensional Gabor-filter model of simple receptive fields in cat striate cortex. // J. Neurophysiol., 1987.-p. 1233-1258.
2. D. Burr, M. Morrone, D. Spinelli Evidence for edge and bar detectors in human vision. // Vision Res., 1989,-p. 419-431.
3. C. Padgett and G. Cottrell. Representing face images for emotion classification. In M. Mozer, M. Jordan, and T. Petsche, editors, Advances in Neural Information Processing Systems, volume 9, Cambridge, MA, 1997. MIT Press.
4. M.S. Bartlett. Face Image Analysis by Unsupervised Learning and Redundancy Reduction. PhD thesis, University of California, San Diego, 1998.
5. V. Bruce. Human face perception and identification. In H. Wechsler, P.J. Phillips, V. Bruce, F. Fogelman-Soulie, and T. Huang, editors, Face Recognition: From Theory to Application, NATO ASI Series F. Springer-Verlag, in press.
6. Wiskott L., Fellous J.M., Krueger N. and von der Malsburg C. Face Recognition and Gender Determination. // Proc. of the Int. Workshop on Automatic Face-and Gesture-Recognition, Zuerich, 1995.-p. 92-97.
7. M. Lades et al. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput., 42 (3):-p. 300-311,1993.
8. Wiskott, L. Phantom Faces for Face Analysis. // Proc. 7th Intern. Conf. on Computer Analysis of Images and Patterns, CAIP ' 97, Kiel, 1997.-p. 480-487.
9. Wiskott L., Fellous J.M., Krueger N. and von der Malsburg C. Face Recognition by Elastic Bunch Graph Matching. // Proc. 7th Intern. Conf. on Computer Analysis of Images and Patterns, CAIP ' 97, Kiel, 1997.-p. 456-463.
10. 10.Ian R. Fasel, Marian S. Bartlett, Javier R. Movellan, A Comparison of Gabor Filter Methods for Automatic Detection of Facial Landmarks, Proceedings of the 7th Symposium on Neural Computation, 2000,-p. 44-50.

Authors information

Murygin Kirill Valerievich – Institute of Artificial Intelligence, B.Hmelnitsky avenue, 84, Donetsk - 83050, Ukraine; e-mail: kir@iai.donetsk.ua