

УДК 004.93

*К.В. Мурыгин*

Донецкий государственный институт искусственного интеллекта, Украина

## Оптимизация габоровских вейвлет для задачи распознавания человека по изображению лица

В статье описаны результаты исследования метода распознавания человека по изображению лица, основанного на габоровских вейвлетах. Определены масштабы функций Габора, при которых достигается максимальный процент распознавания для задачи поиска лица в базе данных и минимальный процент ошибок, связанных с пропуском лиц, при решении задачи контроля доступа. Проведенные исследования показали возможность улучшения параметров работы системы распознавания в двух указанных режимах при сокращении объема используемых данных.

В настоящее время методы, основанные на габоровских вейвлетах (вейвлетом называют ядро волнового преобразования), привлекают все большее и большее внимание исследователей, занимающихся распознаванием изображений, в том числе и изображений лиц. Одним из объяснений растущей популярности данного подхода являются результаты исследований биологов, показавшие сходность двухмерных ядер Габора с формой рецепторного поля зрительных клеток коры головного мозга млекопитающих [1], [2]. Кроме того, уже накоплен достаточно большой положительный опыт использования фильтров Габора в задачах, связанных с распознаванием человека по изображению лица [3-9].

Функции Габора [6], [8], [9] локализованы в пространственной и частотной области и имеют вид плоской волны с волновым вектором  $\vec{k}$ , на которую наложена гауссовская огибающая функция шириной  $\sigma/k$ , где  $\sigma = 2\pi$ :

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

Нормировочный коэффициент, вторая экспонента в квадратных скобках, получен из условия равенства нулю интеграла:

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

Следующее волновое преобразование (wavelet transformation) дает комплексные коэффициенты, которые затем используются в качестве элементов векторов признаков, характеризующих исходное изображение  $I(\vec{x})$  в точке  $\vec{x}$ :

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

Комплексные коэффициенты  $J_j(\vec{x})$  можно записать в виде  $J_j(\vec{x}) = a_j(\vec{x}) \exp(i\varphi_j(\vec{x}))$ , где  $a_j(\vec{x})$  – медленно меняющаяся амплитуда, а фаза  $\varphi_j(\vec{x})$  изменяется с характеристической частотой соответствующего габоровского ядра (1). Коэффициенты  $J_j(\vec{x})$  имеют смысл, сходный с коэффициентами преобразования Фурье. Однако, так как функции (1) локализованы в пространственной области,  $J_j(\vec{x})$  харак-

теризуют не все изображение  $I(\bar{x})$ , а некоторую его часть, размер которой определяется параметром  $\frac{k_j^2}{\sigma^2}$ , а положение – аргументом  $\bar{x}$ .

Полученные на основе выражения (3) векторы признаков  $J_j$  и  $J'_j$  удобно сравнивать с использованием меры сходства в виде косинуса угла между ними:

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

Сочетание метрики (4) и условия (2) позволяет исключить влияние любых линейных преобразований исходных изображений на результат сравнения векторов признаков ((3) и (4)).

Каждое лицо, вводимое в базу данных, представляется в виде локальных векторов признаков, полученных в заранее определенных точках лица. Выбранный набор таких точек в совокупности с полученными в них векторами признаков называется графом лица (рис. 1).

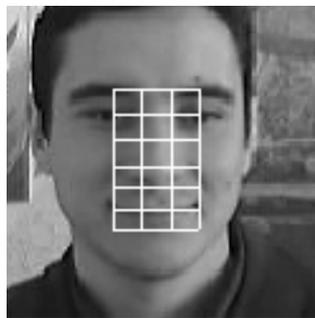


Рисунок 1 – Расположение точек графа лица

Автоматическое распознавание с использованием графов изображений лиц разделяется на два этапа. На первом этапе осуществляется позиционирование графа на изображении. На втором – выполняется сравнение найденного графа с графами лиц, сохраненными в базе данных. Сравнение графов осуществляется путем суммирования мер близости соответствующих векторов признаков, полученных согласно формуле (4), по всем узлам графа.

При выбранной конфигурации графа (расположение точек графа) лица на обоих этапах работы автоматической системы распознавания встает вопрос о выборе набора масштабов используемых габоровских вейвлет, которые обеспечивают максимальную эффективность работы алгоритмов.

В работах [6], [8], [9] для распознавания человека по изображению лица использовались функции Габора пяти различных масштабов,  $\nu = \{0, \dots, 4\}$ , и восьми углов поворота,  $\mu = \{0, \dots, 7\}$ . Каждая функция определялась следующим характеристическим волновым вектором:

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

где индекс  $j = \mu + 8\nu$ . Таким образом, полное волновое преобразование (3) дает 40 комплексных коэффициентов в каждой точке изображения (5 масштабов и

8 углов). Используемый набор масштабов функций Габора был предложен авторами работы [7], которые провели исследования для двух значений пространственного фактора  $f$ , определяющего расстояние между соседними масштабами ядер Габора, и различных значений параметра  $k_{\max}$ , определяющего максимальный масштаб используемых ядер:

$$k_v = \frac{k_{\max}}{f^v}, \quad \phi_\mu = \mu \frac{\pi}{8}.$$

Исследования показали, что оптимальными значениями являются  $f = \sqrt{2}$  и  $k_{\max} = \pi/2$ . Таким образом, используемый диапазон пространственных частот достаточно широк ( $T = 4 - 16$  пикселей). В работе [10] был проведен анализ используемых масштабов функций Габора для сокращения числа используемых ядер при сохранении эффективности работы в применении к задаче локализации точек графа лица. Оптимальные результаты получены при использовании одного масштаба и восьми углов поворота. Причем пространственные частоты, соответствующие данному масштабу, достаточно малы ( $T = 16 - 25$  точек в зависимости от выбранной точки графа лица и при радиусе зрачка 4 пикселя). Сопоставление результатов работ [7] и [10] показывает существенную разницу в выборе масштабов функций Габора для задачи распознавания и задачи локализации графа лица.

В настоящей работе исследования, похожие на описанные в [10], проводились с целью получения оптимального набора масштабов функций Габора для решения задачи распознавания. При проведении экспериментов использовались несколько комбинаций масштабов, предложенных в работе [7] и используемых в работах [6], [8], [9], описанных выше.

Для исследований использовалась база данных из 160 изображений лиц 10 человек (16 изображений для каждого человека; использовалась база данных Weizmann Institute of science\*). Все фотографии были получены при различном освещении. Для каждого из лиц были вручную отмечены ключевые точки (центры зрачков, центры бровей над зрачками, кончик носа и края губ), по которым строился граф лица, и запоминались свойства лица в каждой точке графа в виде векторов признаков, полученных на основе (3) и (5). Использовалась конфигурация графа, описанная в работах [6], [8], [9] (рис. 1). Размеры изображений всех лиц были предварительно приведены к одному масштабу. В качестве масштабного коэффициента было выбрано межзрачковое расстояние. Эталонный масштаб был равен 34 пикселям.

Исследования проводились по двум направлениям. Во-первых, исследовалась зависимость эффективности поиска в базе данных от масштабов используемых ядер Габора. Данная задача характеризуется отсутствием необходимости анализировать ошибки, связанные с пропуском лиц, так как в базе данных ищется эталон, наиболее похожий на входное изображение. Распознавание считается правильным, если найденным наиболее похожим на граф входного изображения лица является граф изображения лица того же человека. Таким образом, критерием оптимизации является процент правильного распознавания. В качестве эталонных случайным образом выбирались по одному из 16 изображений для каждого человека. Всего проводилось

\* База данных получена с сервера <ftp://eris.wisdom.weizmann.ac.il/pub>.

16 экспериментов. Оценка эффективности вычислялась как средняя эффективность, полученная по всем экспериментам. Для каждого из исследуемых наборов масштабов  $\nu = \{0, \dots, 4\}$ , были получены результаты, приведенные в табл. 1.

Таблица 1 – Зависимость процента правильного распознавания от используемых масштабных диапазонов

Использованные масштабы функций Габора					Процент распознавания, %
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

Полученные данные свидетельствуют о том, что наилучшие результаты для поиска в базе данных дают масштабы  $\nu = \{0, 1, 2\}$ , причем эффективнее всего использовать все три масштаба, что дает максимальный процент распознавания – 87.5 %. При использовании всех пяти масштабов эффективность распознавания падает до 81.4 %, что можно объяснить сильной зависимостью больших масштабов от условий освещения. Использование только масштабов 0 и 1 снижает эффективность (по сравнению с максимальной) менее чем на 1 % при сокращении объема запоминаемых и анализируемых данных более чем на 30 %. Это позволяет использовать только эти масштабы в системах, критичных к объему запоминаемых данных и времени работы алгоритма распознавания.

Второе направление исследований было связано с задачей контроля доступа. Для оценки эффективности распознавания для каждого из исследуемых наборов масштабов были рассчитаны функции плотности распределения величин сравнения графов (рис. 2). На рис. 2 левее показана плотность распределения, полученная при сравнении лиц разных людей. Правее – для изображений лица одного человека. Площадь области пересечения характеризует минимальную суммарную ошибку распознавания, которая достигается при выборе порога распознавания, равного абсциссе точки пересечения функций плотности распределения. Данные, полученные при аналогичном выборе порога распознавания, приведены в табл. 2.

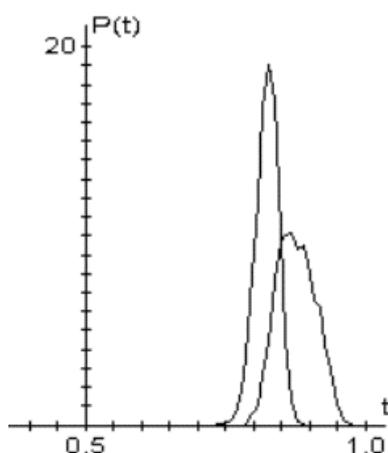


Рисунок 2 – Функции плотности распределения результатов сравнения графов

Таблица 2 – Зависимость ошибок первого и второго рода от используемых масштабов функций Габора

Использованные масштабы функций Габора					Пропуск лица	Неправильное распознавание	Суммарная ошибка
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

Анализ полученных данных показал, что наилучшие результаты по суммарной ошибке наблюдаются при использовании тех же масштабов, что и в экспериментах по поиску в базе данных. Однако в случае задачи контроля доступа большую важность имеют ошибки, связанные с неправильным распознаванием (это эквивалентно возможности несанкционированного доступа). Поэтому оптимальные масштабы необходимо выбирать, главным образом основываясь именно на данных, полученных для этих ошибок.

Если процент допустимых ошибок распознавания жестко определен, то параметром оптимизации становится частота распознавания, которая определяется ошибками, связанными с пропуском лиц. В табл. 3 приведены экспериментальные данные, полученные при априорной установке допустимой ошибки распознавания.

Таблица 3 – Зависимость ошибок пропуска лиц от используемых масштабов функций Габора и заданного ограничения на ошибку распознавания

Используемые масштабы функций Габора					Пропуск лиц (%) при допустимой ошибке распознавания:				
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
				*	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

Как показывают результаты, приведенные в табл. 3, использование только масштабов 0 и 1 является оптимальным и для задачи контроля доступа. Таким образом, сокращение числа используемых масштабов функций Габора с 5 до 2 (использование только масштабов  $v = \{0,1\}$ ) улучшает показатели работы системы распознавания как в случае решения задачи поиска в базе данных, так и для задачи контроля доступа. При этом существенно сокращается объем запоминаемых и анализируемых данных, повышается скорость работы алгоритма сравнения лиц. Для дальнейшего повышения эффективности работы алгоритма распознавания, основанного на использовании функций Габора, необходимо провести исследования в двух направлениях. Во-первых, расширить область поиска в пространстве параметров функций Габора для достижения максимальной эффективности работы системы распознавания. Во-вторых, исследовать возможность компенсации влияния условий освещения на результат автоматической идентификации.

## Литература

1. Jones C. P., Palmer L. An evaluation of the two-dimensional Gabor-filter model of simple receptive fields in cat striate cortex // J. Neurophysiol. – 1987. – P. 1233-1258.
2. Burr D., Morrone M., Spinelli D. Evidence for edge and bar detectors in human vision // Vision Res. – 1989. – P. 419-431.
3. Padgett C., Cottrell G. Representing face images for emotion classification / Ed. by M. Mozer, M. Jordan, T. Petsche // Advances in Neural Information Processing Systems. – Vol. 9. – Cambridge, MA, 1997.
4. Bartlett M.S. Face Image Analysis by Unsupervised Learning and Redundancy Reduction: PhD thesis. – San Diego: University of California, 1998.
5. Bruce V. Human face perception and identification / Ed. by H. Wechsler, P.J. Phillips, V. Bruce, F. Fogelman-Soulie, T. Huang // Face Recognition: From Theory to Application, NATO ASI Series F. – Springer-Verlag, in press.

6. Wiskott L., Fellous J.M., Krueger N., 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. Lades M. et al. Distortion invariant object recognition in the dynamic link architecture // IEEE Trans. Comput. – 1993. – Vol. 42(3). – P. 300-311.
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., 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. Fasel Ian R., Bartlett Marian S., Movellan Javier R. A Comparison of Gabor Filter Methods for Automatic Detection of Facial Landmarks // Proc. of the 7th Symposium on Neural Computation. – 2000. – P. 44-50.

In this paper the researches in the sphere of face recognition method based on the Gabor wavelets are described. The Gabor function scales are determined that maximize recognition rate in the face database search task and minimize face admission error rate in the access control task. The research results show availability to improve quality of the face recognition system in those two modes with significant reduction of the used data size.

У статті описані результати досліджень методу розпізнавання людини за зображенням обличчя, який базується на габоровських вейвлетах. Визначені масштаби функцій Габора, при використанні яких досягається максимальний відсоток розпізнавання для задачі пошуку заданного обличчя в базі даних і мінімальний відсоток помилок, пов'язаних з пропуском облич, при вирішенні задачі контролю доступу. Проведені дослідження показали можливість покращання параметрів роботи системи розпізнавання у двох указаних режимах при скороченні обсягу даних, що використовуються.

*Статья поступила в редакцию 22.08.03.*