

Н. Г. Федотов, А. Г. Петренко, А. В. Рой, Н. С. Фионов

ПОИСК ИЗОБРАЖЕНИЯ ЧЕЛОВЕЧЕСКОГО ЛИЦА ПО ФОТОРОБОТУ В БОЛЬШОЙ БАЗЕ ДАННЫХ

Аннотация. Предложен подход к решению задачи поиска человеческих лиц в базе данных на основе стохастической геометрии и функционального анализа. Ключевым элементом этого подхода является применение нового класса конструктивных признаков распознавания – триплетных признаков. Рассмотрены этапы поиска изображений в базе данных, вытекающие из анализа физиологических принципов восприятия и распознавания совместно с теорией триплетных признаков.

Ключевые слова: распознавание лиц, база данных изображений, биометрический поиск, трейс-преобразование, стохастическая геометрия, триплетный признак.

Abstract. The authors present a new approach based on stochastic geometry and functional analysis to solve the problems of human face image retrieval from a database. A key element of this approach is the use of a new class of constructive features of recognition called triple features. An analysis of physiological principles of perception and recognition in conjunction with the theory of triple features determines a retrieval algorithm described in this paper.

Key words: face recognition, image database, biometric search, trace transform, stochastic geometry, triple feature.

Введение

Статья посвящена поиску изображений человеческих лиц в базе данных по эскизу или фотороботу. В данной работе предлагается подход к решению задачи поиска человеческих лиц на основе стохастической геометрии и функционального анализа. Ключевым элементом этого подхода является новое геометрическое трейс-преобразование изображений, введенное одним из авторов в [1]. Трейс-преобразование связано со сканированием изображений по сложным траекториям. Оно является ключевым элементом формирования конструктивных признаков распознавания нового класса – триплетных признаков, характерной особенностью которых является их структура в виде композиции трех функционалов [2]. Благодаря такой структуре возможно формирование большого числа (тысяч) признаков, причем в режиме автоматической компьютерной генерации. Опора на большое количество признаков ведет к повышению гибкости и надежности поиска.

Согласно психофизической теории восприятия и распознавания при восприятии изображения человеческого лица перемещение глаза носит стохастический характер с фиксацией на областях наибольшей информативности, в качестве которых выступают анатомические особенности лица. Детальная обработка не всего изображения, а преимущественно наиболее информативных областей, существенно увеличивает надежность поиска в базе данных, поскольку при этом реализуется принцип постепенного исключения избыточности.

1. Физиологические основы восприятия и распознавания человеческого лица

В 1930-е гг. на основе изучения зрительного аппарата человека и животных было установлено, что регулярный обзор пространства в природе не

осуществляется, и траектория движения глаз при осмотре объекта носит случайный характер.

Впоследствии экспериментальные исследования А. Л. Ярбуса по изучению механизмов выбора точек фиксации взгляда показали, что человек в процессе осмотра изображений фиксирует взгляд главным образом на наиболее информативных областях.

Выбор таких областей изображений для фиксации взгляда и детального анализа с помощью механизмов фовеального зрения зависит от множества факторов, взаимодействующих между собой.

Комплексный подход к изучению механизмов осмотра изображений, основанный на взаимодействии психофизического эксперимента и математического моделирования, описан в работе [3] научной школы НИИ нейрокбернетики им. А. Б. Когана Южного федерального университета. Наиболее информативные области, выделенные по результатам экспериментов с использованием современной системы регистрации и анализа движений глаз, оказались во многом сходны с результатами А. Л. Ярбуса.

Детальная обработка наиболее информативных областей (области глаз, носа, рта, контура лица) может значительно снизить вычислительные затраты, а удачное признаковое описание таких областей увеличит эффективность поиска и идентификации [4].

Специфика восприятия изображений лиц зрительным анализатором человека состоит также в том, что в процессе узнавания мозг выступает как активная распознающая система (с проверкой правильности решений). Этапами распознавания при этом могут быть: выделение признаков, предварительный анализ, выдвижение гипотезы, проверка гипотезы – сличение изображений с эталоном, взятым из памяти. В связи с этим принцип активного распознавания должен закладываться на этапе проектирования систем компьютерного видения изображений человеческих лиц [5].

2. Триpletные признаки

Для описания каждого портретного изображения строится набор переменных – множество tripletных признаков.

Пусть $F(x, y)$ – функция изображения на плоскости (x, y) . Определим на плоскости сканирующую прямую $l(\varphi, \rho, t)$, которая задается нормальными координатами φ и ρ :

$$x \cdot \cos \varphi + y \cdot \sin \varphi = \rho,$$

параметр t задает точку на прямой. Определим функцию двух аргументов $g(\varphi, \rho) = T(F \cap l(\varphi, \rho, t))$ как результат действия функционала T при фиксированных значениях переменных φ и ρ .

При проектировании распознающих систем используют дискретный вариант трейс-преобразования. Параметры сканирующей прямой образуют два дискретных множества $\Omega = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$, $\Gamma = \{\rho_1, \rho_2, \dots, \rho_m\}$.

В результате действия функционала T получаем матрицу, элементами которой являются значения $t_{ij} = T(F \cap l(\varphi_j, \rho_i, t))$. Детерминированное сканирование позволяет однозначно определить каждый элемент матрицы. Дан-

ная матрица называется трейс-матрицей. Она является результатом трейс-преобразования, или трейс-трансформантой.

В дискретном варианте вычислений трейс-трансформанта представляет собой матрицу, элементами которой t_{ij} являются, например, значения яркости изображения F при пересечении со сканирующей линией $l(\varphi, \rho)$. Параметры сканирующей линии φ_j , ρ_i определяют позицию этого элемента в матрице. Последующее вычисление признака заключается в последовательной обработке столбцов матрицы с помощью функционала P , который называется диаметральным функционалом.

К полученному после применения P -функционала набору чисел, представляющему собой 2π -периодическую кривую, нужно применить круговой функционал Φ , чтобы получить число-признак.

Таким образом, триплетный признак вычисляется как последовательная композиция трех функционалов: $\Pi(F) = \Phi \circ P \circ T(F \cap L(\varphi, \rho, t))$, где каждый функционал (Φ , P и T) действует на функции одной переменной (φ , ρ и t) соответственно.

Варьируя свойства функционалов, входящих в триплетный признак, можно получить признаки с заданными свойствами. В частности, при определенном выборе функционалов можно получить признаки, инвариантные по отношению к группе движений и линейных деформаций распознаваемых изображений. Это повышает надежность поиска изображений в базе данных, так как система идентификации должна устойчиво функционировать в условиях линейных преобразований изображения и ограниченных ракурсных трансформаций объекта идентификации.

Кроме того, благодаря источнику триплетных признаков – трейс-преобразованию – в этих признаках отражаются свойства окрестностей точек пересечения сканирующей линией, что также говорит о перспективности выбора триплетных признаков для задачи поиска в базе данных изображений.

Таким образом, анализ физиологических принципов восприятия и распознавания совместно с теорией триплетных признаков приводит к идее поиска изображений в базе данных в несколько этапов, одним из которых является обучение. В данной работе поиск в базе данных изображений лиц производится по фотороботу, представленному в виде бинарного изображения.

3. Предварительная обработка изображений

Поскольку черно-белое изображение фоторобота сравнивается с изображением лиц из базы данных, необходима предварительная обработка изображений с целью приведения их к бинарному виду. Исходные цветные фотографические портреты приводятся специальной процедурой к полутоновым изображениям. Затем полутоновое изображение преобразуется в монохроматическое путем нелинейной пороговой фильтрации.

На этапе сегментации происходит выделение (разбиение изображения) наиболее информативных областей: глаза, нос, рот, овал лица и др. (рис. 1). В данном случае применен оригинальный метод, основанный на рекурсивной функции заливки заданным цветом произвольной замкнутой области изображения.

При сегментации портретного изображения с учетом физиологии человеческого лица происходит анализ окрестности заданных координат. Если пиксель в окрестности не является фоном, то координаты стартовой точки меняются и процедура вызывает себя с новыми измененными координатами. Это будет происходить до тех пор, пока не будут зафиксирована вся замкнутая область.

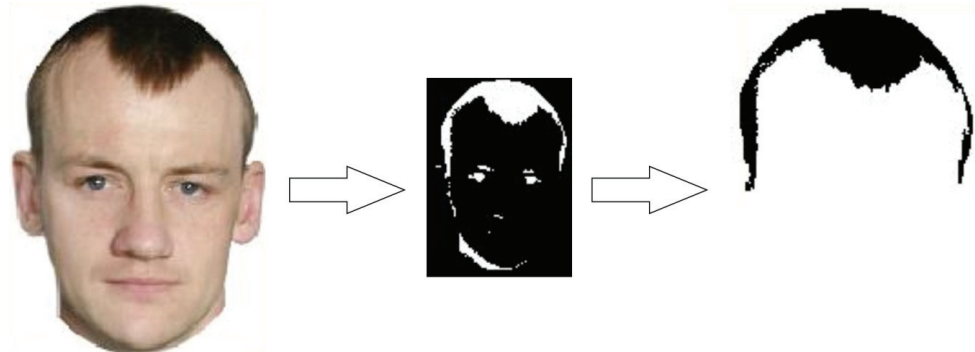


Рис. 1. Сегментация изображения из базы данных: выделение высоты лба

4. Этап обучения

В режиме обучения путем генерации триплетных признаков каждого портретного изображения, хранящегося в базе данных, строится набор поисковых триплетных признаков. Вычисления осуществляются по всем областям наибольшей информативности. Признаки, значения которых попадают в непересекающиеся или частично пересекающиеся интервалы обучающей совокупности, отбираются для поиска. Остальные, малоинформативные признаки, исключаются из рассмотрения, их влияние при принятии решения невелико.

Этапы вычисления триплетного признака схематично показаны на рис. 2.

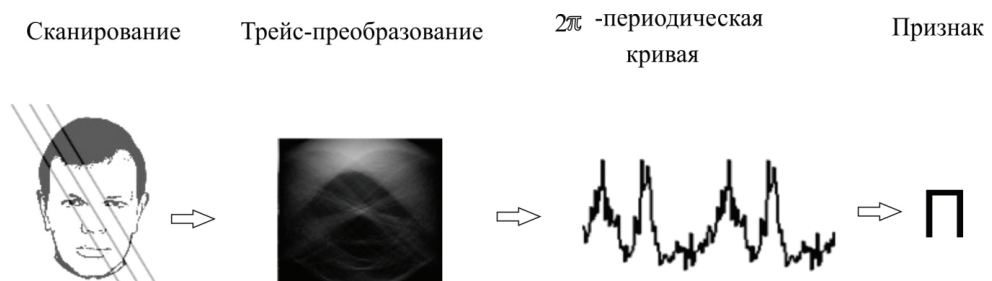


Рис. 2. Вычисление триплетного признака

Иногда для распознавания изображения достаточно только лишь одного трейс-преобразования, осуществляемого с помощью Т-функционалов. Однако при распознавании человеческих лиц в большинстве случаев одного трейс-преобразования недостаточно. Поэтому для всех изображений лиц целесообразно производить структурный анализ биометрических элементов портретного изображения (расположение глаз, носа, овала лица, определение контуров подбородка, усов). И уже для каждого элемента в отдельности выполнять процесс сканирования и вычисления триплетных признаков.

Например, важным этапом анализа человеческих лиц является различение лиц в зависимости от величины открытых глазных щелей.

На рис. 3 приведены примеры такой классификации, а также соответствующие трейс-преобразования, полученные с помощью наиболее различающихся функционалов.

Из рис. 3 видна чувствительность различных T -функционалов и соответствующих трейс-матриц к сегментированной области глаз. Однако согласно алгоритму распознавания решающим критерием при распознавании изображений является триплетный признак. T -функционал, формирующий трейс-преобразование, является лишь фрагментом триплетного признака. Дальнейшая свертка информации осуществляется путем постолбцовой обработки трейс-матрицы с помощью диаметрального функционала. Итогом будет 2π -периодическая кривая. Изучая ее поведение, можно продолжить отбирать самые информативные поисковые признаки, а также оценить влияние P -функционалов на всю процедуру распознавания, используемую при поиске изображений в базе данных.

После обработки 2π -периодической кривой круговым функционалом получаем триплетный признак изображения.

Видимая на глаз чувствительность T -функционалов лишь облегчает работу по формированию набора информативных признаков в процессе обучения. На практике, даже если между трейс-трансформантами нет заметных различий для разных исходных изображений, различия между этими изображениями все же можно выявить, если заранее знать, что геометрически они существуют. Чтобы достичь этой цели, необходимо целенаправленное формирование триплетных признаков для самой главной области наибольшей информативности.

Например, для глаз может быть характерна либо вытянутость по горизонтали, либо круглая форма. Следовательно, нужно сформировать триплетные признаки, оценивающие различие форм: эллипсовидной и круглой. При округлой форме число нулевых элементов в каждом из столбцов трейс-матрицы будет примерно одинаково. Это происходит потому, что значение длин отрезков, высекаемых на различных сканирующих прямых, примерно одинаково. Для вычисляемого объекта наибольшее число ненулевых элементов расположено в том столбце, который соответствует направлению сканирования, совпадающему с направлением вытянутости.

5. Оценка информативности признакового пространства

Для оценки информативности пространства признаков используется подход, основу которого составляет гипотеза компактности [6]. Из нее следует, что для хорошего распознавания образов желательно, чтобы расстояния между своими точками каждого образа были малыми, а расстояния до точек других образов по возможности большими.

На рис. 4 приведены 3 исходных портретных изображения, предварительная обработка, сегментация и результаты вычислений триплетных признаков. Для данного примера в качестве трейс-преобразования T_1 было выбрано число сегментов, получаемых при пересечении прямой и образа. Также были взяты один диаметральный функционал P_1 – сумма всех элементов p -го столбца трейс-матрицы; и три круговых функционала: Φ_1 – площадь, образующаяся функцией на отрезке $[0; 2\pi]$, Φ_2 – среднее значение и Φ_3 – дисперсия функции.



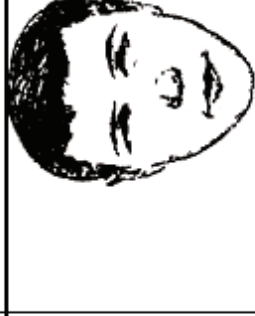
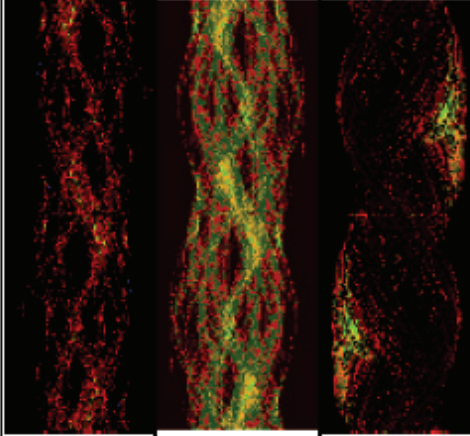
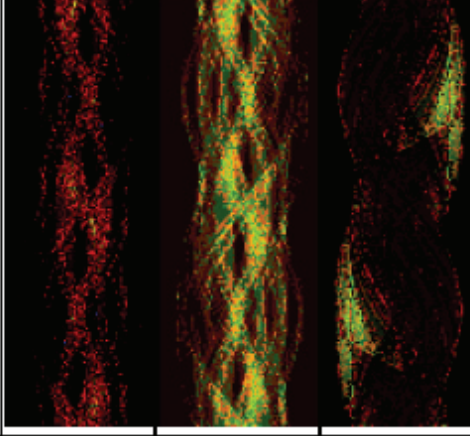
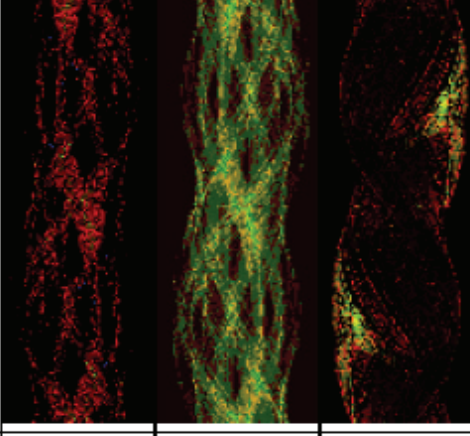
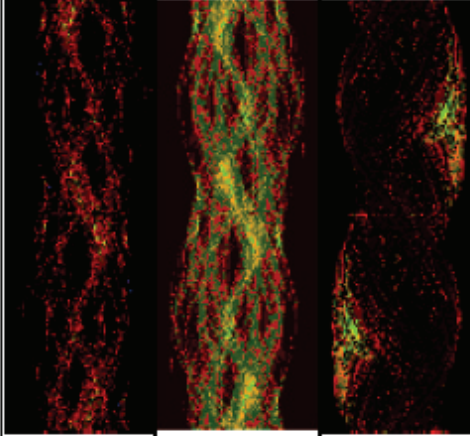
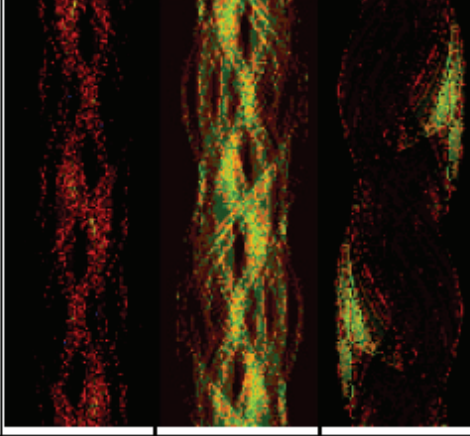
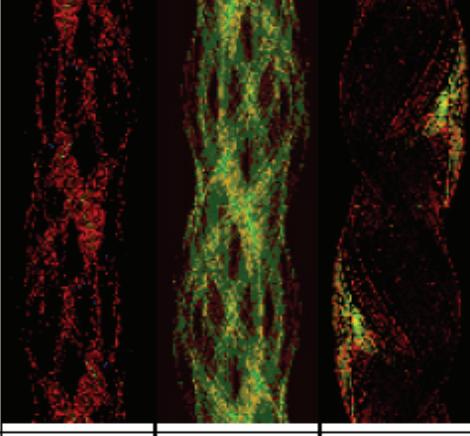
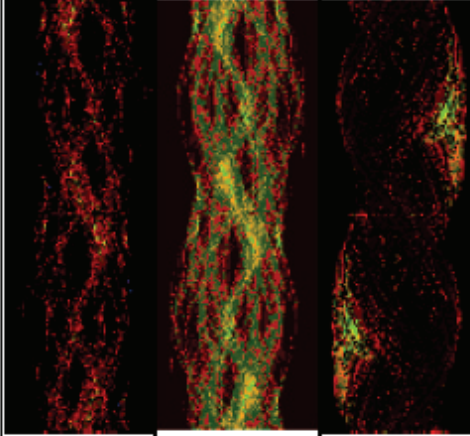
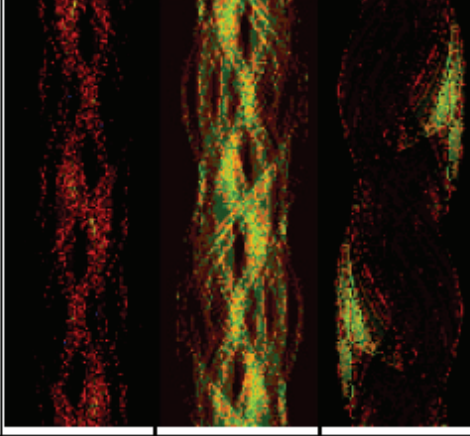
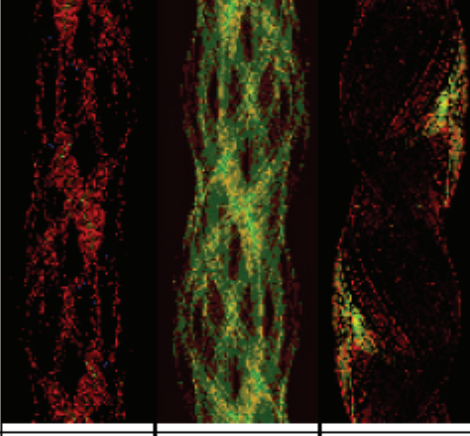
		Величина открытых глазных щелей		
		Средняя	Большая	Малая
	T-функционал			
	Длина высекаемых отрезков			
	Число пересечений сканирующей линии с фигурой			
	Максимальная длина высекаемых отрезков			

Рис. 3. Сравнение изображений лиц с разными величинами открытыми глазными щелями по результатам применения T-функционалов

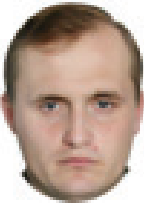








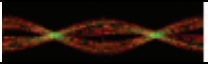
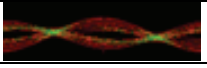
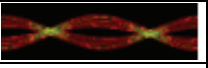
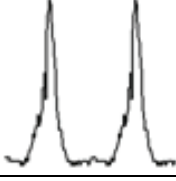
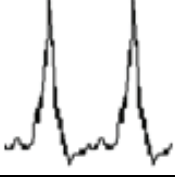
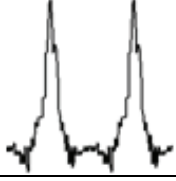
Исходный вид изображений				Π_{cp}
Приведение к бинарному виду				
Пример сегментации (область глаз в качестве области наибольшей информативности)				
T_1				
$T_1 \circ P_1$				
$\Pi_{111} = T_1 \circ P_1 \circ \Phi_1$	3568	2738	3194	3203,2
$\Pi_{112} = T_1 \circ P_1 \circ \Phi_2$	1,0295	1,0280	1,0889	1,04084
$\Pi_{113} = T_1 \circ P_1 \circ \Phi_3$	161,70	83,94	131,61	126,618

Рис. 4. Результаты вычислений триплетных признаков

Для полученных триплетных признаков необходимо выбрать весовые коэффициенты. Это может сделать или оператор в режиме обучения, или программа в соответствии с гипотезой компактности. Для последнего случая, если за значения признака принять расстояние между точками сопоставляемых образов, необходимо сначала определить величину разнесенности образов средним расстоянием между точками образа.

Информативность признака вычисляется как отношение разнесенности каждого признака к их средней величине. Для данного набора изображений самым информативным из трех является признак Π_{111} . Полученные оценки информативности используются при сравнении изображений фоторобота с изображениями из базы данных.

6. Этап поиска

Процесс поиска в базе данных основан на том, чтобы предложить эксперту для рассмотрения в первую очередь те портретные изображения, которые имеют наибольшую степень близости к задаваемому фотороботу. Таким образом, происходит упорядочение всех изображений по выбранному критерию близости к объекту поиска – наименьшему расстоянию между сравнива-

емыми наборами поисковых признаков. Используя значения критерия R с учетом оценок информативности, программа упорядочит изображения по степени близости к выбранному фотороботу следующим образом:





Фоторобот	Изображение	Критерий $R_i, \%$
		0,03255
		0,02603
		0,02066

Рис. 5. Результат применения решающего правила

В итоге мы получаем список, в начало которого попадают самые похожие на фоторобот изображения. Результаты поиска для 200 фотороботов и 1500 изображений из базы данных таковы: в 70 % случаев на первое место списка ставится изображение, соответствующее фотороботу; в 95 % случаев изображение, сопоставляемое фотороботу, попадает в первую двадцатку списка.

Заключение

Для поиска изображений в базе данных использован математический аппарат стохастической геометрии, что позволило достигнуть высоких показателей эффективности: гибкости, универсальности, надежности поиска, большего быстродействия.

При определенном выборе свойств функционалов, входящих в триплетный признак, можно добиться инвариантности поисковых признаков по отношению к группе движений и линейным искажениям изображений.

Для повышения надежности поиска изображений предложено определение триплетных признаков по областям изображений, соответствующим элементам человеческого лица. С этой целью разработана процедура сегментации. Алгоритм поиска на основе стохастической геометрии включает предварительную обработку изображения, его сегментацию, формирование геометрических трейс-преобразований изображений и вычисление по ним триплетных признаков, решающую процедуру.

Список литературы

1. **Федотов, Н. Г.** Методы стохастической геометрии в распознавании образов / Н. Г. Федотов. – М. : Радио и связь, 1990. – 143 с.
2. **Федотов, Н. Г.** Теория признаков распознавания образов на основе стохастической геометрии и функционального анализа / Н. Г. Федотов. – М. : Физматлит, 2009. – 304 с.
3. **Подладчикова, Л. Н.** Комплексный подход к исследованию механизмов осмотра изображений / Л. Н. Подладчикова и др. // Нейроинформатика–2009 : сборник научных трудов 11-й Всероссийской научно-технической конференции. – М. : МИФИ, 2009. – Т. 1. – С. 249–257.
4. **Шапошников, Д. Г.** Оценка параметров информативных областей изображений лиц / Д. Г. Шапошников и др. // Материалы Международной конференции по нейрокибернетике. – Ростов-на-Дону, 2002. – С. 163–166.
5. **Глазунов, А. С.** Компьютерное распознавание человеческих лиц / А. С. Глазунов // Открытые системы. – 2000. – № 3. – URL: <http://www.osp.ru/os/2000/03/>.
6. **Загоруйко, Н. Г.** Прикладные методы анализа данных и знаний / Н. Г. Загоруйко. – Новосибирск : Изд-во Инст. математики, 1999. – 270 с.

Федотов Николай Гаврилович

доктор технических наук, профессор,
заведующий кафедрой экономической
кибернетики, Пензенский
государственный университет

E-mail: fedotov@pnzgu.ru

Fedotov Nikolay Gavrilovich

Doctor of engineering sciences, professor,
head of sub-department of economical
cybernetics, Penza State University

Петренко Александр Григорьевич

кандидат технических наук, доцент,
кафедра экономической кибернетики,
Пензенский государственный
университет

E-mail: alxndrpetrenko@mail.ru

Petrenko Alexander Grigoryevich

Candidate of engineering sciences, associate
professor, sub-department of economical
cybernetics, Penza State University

Рой Алексей Владимирович

кандидат технических наук, старший
преподаватель, кафедра экономической
кибернетики, Московский
государственный агроинженерный
университет им. В. П. Горячкина

E-mail: ec@pnzgu.ru

Roy Aleksey Vladimirovich

Candidate of engineering sciences,
senior lecturer, sub-department of economic
cybernetics, Moscow State Agroengineering
University named after V. P. Goryachkin

Фионов Никита Сергеевич

аспирант, Пензенский
государственный университет

E-mail: finik86@rambler.ru

Fionov Nikita Sergeevich

Postgraduate student,
Penza State University

УДК 004.931:004.932

Федотов, Н. Г.

Поиск изображения человеческого лица по фотороботу в большой базе данных / Н. Г. Федотов, А. Г. Петренко, А. В. Рой, Н. С. Фионов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2011. – № 3 (19). – С. 65–74.