

УДК 004.93'1: 004.932

**А. В. Гороховатский**

Харьковский национальный университет радиоэлектроники  
проспект Ленина, 14, 61166 Харьков, Украина

## **Верификация изображения лица в пространстве инвариантных признаков с применением лучевых преобразований**

*Исследована проблема верификации изображения человеческого лица. Предложен метод вычисления инвариантных признаков для распознавания на основе анализа систем проекционных функций изображения. Проведенные эксперименты подтверждают высокое быстродействие и работоспособность подхода.*

**Ключевые слова:** компьютерное зрение, распознавание, верификация лица, инвариантные признаки, проекционное преобразование, Trace/Radon.

### **Введение**

Распознавание лиц в эпоху развитой IT-индустрии приобретает распространение во всех компьютеризированных областях, таких как Интернет-технологии (Picasa, iPhoto, Windows Live Photo Gallery), производство техники, стационарных и портативных компьютеров (VeriFace, SmartLogon, Toshiba Face Recognition Software). Достаточно большой сегмент подобных систем насыщен развивающимися коммерческими продуктами (NeoFace, VOCORD, JustLook, C-TRUE Super, Blink!), которые решают актуальные прикладные задачи распознавания [1, 2].

Практически все системы верификации лиц часто получают негативные отзывы пользователей из-за недостаточного качества их работы. Легко находятся способы, с помощью которых можно обмануть даже самые известные и популярные из них. Необходимо также учитывать психологические аспекты внедрения соответствующего программного обеспечения, которые связаны с предубеждением пользователей. Несмотря на активное развитие и внедрение методов идентификации лиц, проблему распознавания лиц до сих пор нельзя считать окончательно решенной.

Существует несколько основных групп методов для анализа и распознавания лиц на фото- и видеофрагментах [1]. Наиболее проблемными местами остаются изменчивость освещения, угла поворота головы и ее местоположение, локальные объекты (очки, борода и т.п.), затрудняющие распознавание. Большинство методов

© А. В. Гороховатский

ориентированы на выделение характерных признаков и сравнение их между собой как в числовом, так и в структурном контексте, что зачастую представляет собой реализацию довольно сложных в вычислительном плане методов и алгоритмов. Способам, в основе которых лежит анализ характеристик цельного изображения, уделено недостаточно внимания. Один из таких подходов связан с преобразованиями Trace/Radon [7, 8], основанными на вычислении проекционных функционалов изображения.

Целью исследования является разработка метода быстрой верификации (либо распознавания) изображения лица человека с использованием инвариантных проекционных признаков.

Под верификацией мы подразумеваем частный случай распознавания, т.е. проверку гипотезы о том, являются ли входное и эталонное изображения представителями одного класса эквивалентности. Распознавание дает ответ на вопрос, какому из классов эквивалентности принадлежит входное изображение. Задачи верификации и распознавания на ограниченном наборе (2–3 эталона) считаем равносложными. Отдельно отметим, что исследование посвящено сравнению изображений, в то время как процедура локализации лица считается решенной.

### Инвариантные признаки проекционного типа

Разложение изображения на совокупность проекций получило распространение в задачах компьютерного зрения, т.к. проекции содержат значительный потенциал для получения необходимого набора инвариантных признаков без существенных вычислительных затрат в сравнении с другими методами [3].

В качестве инвариантных значений для верификации будем использовать признаки на основе проекционного анализа изображения и геометрических моментов [3–5]:

$$I_k = (m_0)^{-(k+1)} \int_p (R(p, \theta))^{k+1} \left(p - \frac{m_1}{m_0}\right)^k dp, \quad m_0 = \int_p R(p, \theta) dp, \quad m_1 = \int_p p R(p, \theta) dp, \\ R(p, \theta) = \iint_{x, y} B(x, y) \delta(p - x \cos \theta - y \sin \theta) dx dy, \quad (1)$$

где  $B(x, y)$  — анализируемое изображение;  $R(p, \theta)$  — преобразование Радона;  $k$  — порядок момента (зависит от решаемой задачи, а также от вычислительных и временных ограничений проекта);  $\theta$  — угол проецирования. Признаки (1) объединяют структурную информацию об изображении с численным описанием, что позволяет выполнять обучение и распознавание на базе сравнения числовых данных [2]. Признаки (1) устойчивы к искажениям масштабирования и смещений объекта в пределах поля зрения.

Инвариантные признаки могут быть построены как непосредственно на базе выражения (1) с использованием функции яркости  $B(x, y)$ , так и с помощью модификаций вида  $\Psi(B(x, y))$ , где  $\Psi$  — некоторая преобразующая функция. Первый метод соответствует преобразованию Радона, второй называют Trace-пре-

образованием [7, 8]. Применение различных вариаций  $\Psi$  расширяет набор возможных инвариантных признаков.

### Нормализация и сопоставление признаков

Основной характеристикой инвариантов  $I_k$  согласно (1) является порядок  $k$ . Например, признаки первого и седьмого порядка могут различаться по значению более чем в тысячу раз, что не дает возможности сравнивать их без предварительной нормализации. Для нормирования инвариантов  $I_k$  введем преобразование к некоторому фиксированному интервалу  $[-10;10]$ :

$$F = 10^{f-1}, \quad f = \begin{cases} \min_i(I_k \times 10^i < -10), & I_k < 0, \\ \min_i(I_k \times 10^i > 10), & I_k \geq 0, \end{cases}$$

где  $i \in [0;100]$  — целочисленный параметр. За счет предложенных преобразований получаем нормированные признаки одного порядка.

Одним из наиболее простых в вычислительном аспекте для сопоставления проекционных описаний является манхэттенская метрика

$$\rho_M(B, B_0) = \sum_k \sum_{\theta} |I_k(\theta) - I_k^0(\theta)|, \quad (2)$$

где  $B_0$  — шаблон изображения;  $B$  — входное изображение;  $\{I_k^0(\theta)\}$  — множество признаков  $B_0$ ;  $\{I_k(\theta)\}$  — множество признаков  $B$ .

Для снижения уровня FAR (False Acceptance Rate — «ложный пропуск») целесообразно использовать некоторое пороговое значение  $\delta$ , при котором два набора инвариантов считаем эквивалентными. Порог может быть установлен в абсолютном либо относительном виде. Значимость уровня FRR (False Rejection Rate — «ложный отказ») зависит от конкретных условий эксплуатации системы верификации. Он регулируется тем же значением  $\delta$ . В то же время можем им пренебречь ввиду возможности пользователя показать лицо в видеокамеру необходимое число раз.

Во многих системах верификации используется также параметр, который численно характеризует уровень доверия полученному результату. В качестве такого параметра можно привести

$$\varepsilon(B_i, B_0) = \frac{\min_2(\rho_M(B_i, B_0))}{\min_1(\rho_M(B_i, B_0)) + \min_2(\rho_M(B_i, B_0))},$$

где  $\min_1(\rho_M(B_i, B_0))$ ,  $\min_2(\rho_M(B_i, B_0))$  — значения двух наименьших ближайших значений  $\rho_M(B_i, B_0)$ , вычисленных в соответствии с (2). Таким образом, если результатом распознавания будут два примерно похожих эталона, это отразится на показателе  $\varepsilon$ . Из практических исследований целесообразным представляется взять значения  $\varepsilon > \varepsilon_0$  при  $\varepsilon_0 = 0,6$ .

Пусть  $L$  — база эталонов. Распознавание изображения на основе (2) реализуется путем минимизации на подмножестве  $L_0 \subset L$ ,  $L_0 = \{B_i\}$ , которые прошли пороговую проверку:

$$i^* = \arg \min_{i \in L_0} \rho_M(B, B_i), \rho_M(B, B_i) < \delta \quad (3)$$

либо

$$i^* = \arg \min_{i \in L} \rho_M(B, B_i), \varepsilon(B, B_i) > \varepsilon_0, \quad (4)$$

или

$$i^* = \arg \min_{i \in L_0} \rho_M(B, B_i), \rho_M(B, B_i) < \delta, \varepsilon(B, B_i) > \varepsilon_0. \quad (5)$$

Распознавание в соответствии с (3)–(5) основано на анализе только избранных изображений путем их отбора за счет применения соответствующих логических условий. Наряду с этим применимы векторные меры подобия, которые по аналогии с расстоянием Хэмминга подсчитывают число близких компонент векторов для наборов инвариантных признаков  $\{I_k(\theta)\}$  и  $\{I_k^0(\theta)\}$  при фиксированном  $k$ , которые описывают распознаваемое и эталонное изображения. Такая мера имеет вид:

$$\mu(I_k(\theta), I_k^0(\theta)) = \sum_{i=1}^M d_i, \quad d_i = \begin{cases} 1, & |I_i(\theta) - I_i^0(\theta)| < \beta, \\ 0 & \text{otherwise,} \end{cases}$$

где  $\beta$  — априорный порог близости (эквивалентности) для отдельных признаков;  $M$  — число используемых признаков для заданного  $k$ . Процедуру принятия решения для этой модели запишем в виде:

$$i^* = \arg \max_{i \in L} \mu(I_k^i(\theta), I_k^0(\theta)). \quad (6)$$

Комбинирование в одной процедуре мер (6) и ограничений вида (5) представляют возможность дальнейшего улучшения результатов распознавания для конкретных баз видеоинформации.

## Экспериментальные результаты

Процесс обучения в задаче распознавания лиц представляет собой слежение за конкретным объектом в течение интервала времени и индексацию его изображений пользователем вручную в базе данных с помощью текстового описания (имени). Одномерная обработка посредством проекций позволяет получить инвариантные признаки нескольких изображений лиц одного и того же объекта в режиме реального времени, что дает возможность заполнить базу данных даже при условии ограниченного входного видеопотока. Обучение большинства систем распознавания требует соблюдения неких стандартных условий, например, неподвижность головы, постоянство освещения и т.п. Предложенный подход к распознаванию характеризуется обучением на разноплановых изображениях с целью получения инвариантных признаков в широком спектре условий.

Нами экспериментально показано, что для обучения целесообразно использовать набор до 200 эталонов, поскольку это обеспечивает возможность реализации предложенных методов в реальном времени. В базу также занесены результаты ошибочной работы детектора лиц (либо лица плохого качества), соответствующие наборы инвариантных признаков маркированы описанием «Неопределенная персона». Пример видеобазы для одного человека приведен на рис. 1, на ее получение была выделена 1 минута наблюдения.

Обучение выполняется в условиях, приближенных к условиям реальной верификации. Это вызвано особенностями процедур локализации лица, которые допускают участие фоновых компонент в построении признаков.

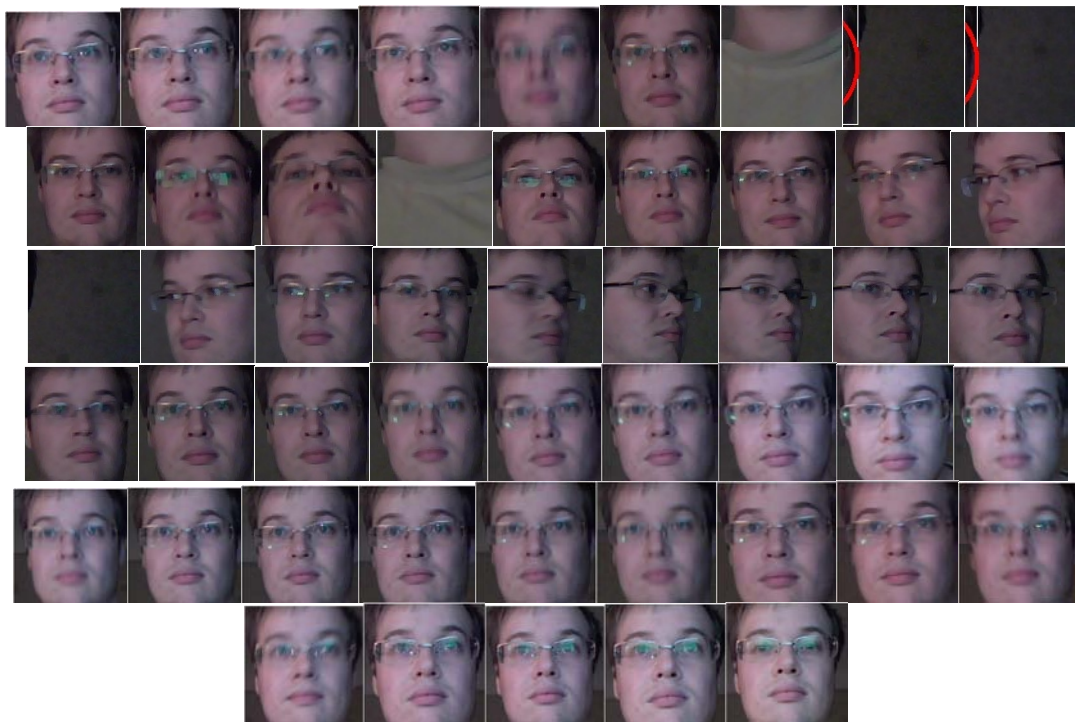


Рис. 1. Пример обучающего набора

Процесс верификации/распознавания включает в себя локализацию изображения лица, получение его инвариантных признаков и сопоставление их с внесенными ранее в базу данных. На рис. 2 приведены примеры правильно распознанных изображений лица под разными углами, с освещением, отличным от эталонного, и измененным масштабным коэффициентом. Конкретно использовано 40 коэффициентов для вычисления выражений (2).

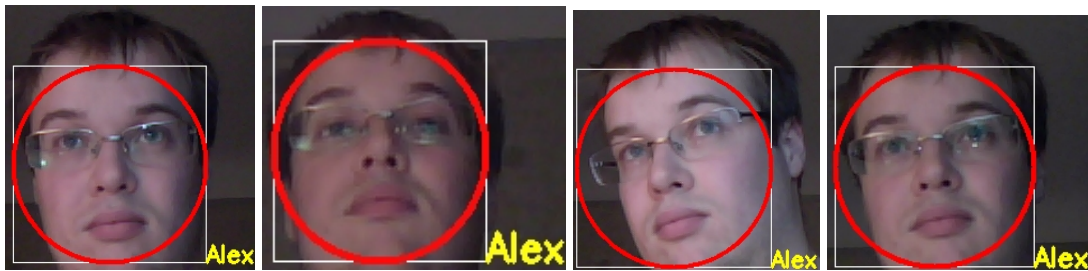


Рис. 2. Примеры правильного распознавания лица

Приведенная далее таблица содержит результаты отдельных экспериментов по верификации лица, которые дают возможность проанализировать общую картину. Последовательно указаны следующие параметры: длительность проведения эксперимента; количество лиц, найденных детектором за это время; реальное количество лиц (кроме ошибок определения и лиц плохого качества); число эталонных классов в базе данных (при наличии класса «неопределенная персона»); количество людей, которые участвовали в эксперименте / количество тех из них, которые занесены в эталонную базу; суммарное число эталонных инвариантных описаний в базе данных для всех классов; параметры распознавания (ограничения и модель); количество лиц, которые были идентифицированы; уровень ошибки, при котором лица были распознаны неверно.

По данным таблицы можно говорить о том, что предложенные модели характеризуются значительным уровнем ошибки FAR (эксперименты № 1, 3, 4), что объяснимо использованием инвариантных признаков, значительно больше отражающими общие признаки лица, чем его индивидуальные особенности, которые сглаживаются в результате интегрирования. Снижение уровня FAR может быть достигнуто путем использования более точных (но и более трудоемких) методов сравнения, основанных на характерных признаках изображений (биометрические соотношения, структура множества индивидуальных особенностей и т.п. [2]).

Наиболее удачным среди обсуждаемых можно назвать эксперимент № 2, который моделирует классическую систему верификации: незначительное количество эталонных классов, на вход подается набор признаков, близкий к одному из классов. Особенностью эксперимента является жесткий условный отбор в виде (5), который позволяет минимизировать ошибки. Вместе с тем, затруднена верификация даже эталонного изображения вследствие предъявленных требований, варьирование параметров ограничений дает возможность повысить уровень безопасности.

Результаты верификации/распознавания лиц на видеофрагментах							
№	1	2	3	4	5	6	7
Характеристика							
Время, мин	39	37	58	37	53	53	20
Подозрений на лица	1929	1732	3 315	1580	2942	2942	1107
Лиц	1742	1620	3018	1299	2453	2453	703
Классов в базе данных	3	3	3	3	1	3	1
Персон/из них в базе данных	3/2	1/1	2/2	2/2	1/1	1/2	2/1
Число эталонов	138	174	138	138	200	138	10
Параметры	(3), $\delta = 20$	(5), $\delta = 20,$ $\varepsilon_0 = 0,6$	(5), $\delta = 20,$ $\varepsilon_0 = 0,6$	(6), $\beta < 0,1$	(6), $\delta = 20,$ $\varepsilon_0 = 0,6$ $\beta < 0,1$	(6), $\beta < 0,005$	(6), $\beta < 0,15$ $\delta = 20$
Распознано лиц	883	91	614	82	351	2453	202
FAR	352 (40 %)	2 (2,2 %)	204 (33,2 %)	20 (24,3 %)	33 (9,5 %)	877 (30 %)	13 (6,4 %)

Эксперимент № 5 показывает, что даже в самой простой ситуации (сравнение типа «один к одному») возможны ошибки, которые заключаются в том, что лицами признаны те изображения, которые лица не содержат. Использование же только модели (6) (эксперимент № 6) без дополнительной фильтрации недостаточно эффективно вследствие высокой ошибки FAR.

Условия, близкие ко многим реализациям программного обеспечения по идентификации лица смоделированы в эксперименте № 7. Их особенностью является фиксированное положение лица (анфас), небольшое количество эталонных изображений, которые образуют единственный эталонный класс, одинаковая освещенность в рамках серии экспериментов. В данной ситуации предложенный подход показывает достаточно уверенные результаты, с вероятностью более 0,93 указан правильный эталонный класс либо отброшен ложный.

## Выводы

В результате исследований предложен метод верификации изображения лица человека с использованием инвариантных проекционных признаков. Пространственное интегрирование функции яркости изображения дает возможность быстро построить необходимое количество проекций, обеспечивающее нужную достоверность распознавания. Фоновые составляющие, воздействие освещения и других динамических характеристик может оказать влияние на результат построения признаков и распознавание.

Научная новизна состоит в разработке и реализации простого в вычислительном плане метода верификации изображения лица на основе инвариантных проекционных признаков, число которых может расширяться с целью достижения нужного качества распознавания.

Прикладные аспекты исследования связаны с возможностью применения данного метода в системах наблюдения, где ошибки FAR не являются критичными. Предложенный подход может быть реализован в качестве первичного этапа «грубой верификации» с последующим уточнением результата.

Работа содержит результаты исследования по верификации лиц на базе одной проекции, представляя изображение в виде одномерного образа. Преобразование Радона/Trace можно применить для анализа объекта под разными углами наблюдения, что потенциально приводит к перспективе использования накопления либо анализа результатов измерений по множеству проекций.

Моделирование предложенного подхода выполнено на C++ с использованием средств локализации лица из библиотеки OpenCV от Intel [6].

1. *Face Recognition Homepage* [Электронный ресурс]. — Режим доступа: <http://www.face-rec.org/algorithms/>
2. *Шапиро Л.* Компьютерное зрение / Л. Шапиро, Дж. Стокман.; [пер. с англ. А.А. Богуславского под ред. С.М. Соколова]; М.: БИНОМ. Лаборатория знаний, 2006. — 752 с.
3. *Гороховатский А.В.* Проекционные методы в задачах нормализации и распознавания изображений / А. В. Гороховатский // АСУ и приборы автоматики. — 2009. — № 147. — С. 64–70.
4. *Гороховатский А.В.* Анализ статистических особенностей набора инвариантных признаков / А.В. Гороховатский // Интелектуальні системи прийняття рішень і проблеми обчислювального інтелекту (ISDMCI'2010): зб. наук. праць міжн. наук. конф.: Євпаторія, 17–22 травня 2010. — Херсон: ХНТУ. — Т. 2. — С. 340–342.
5. *Gorokhovatsky A.V.* Image Invariant Recognition Methods in the Space of Projections / A.V. Gorokhovatsky // Прикладна радіоелектроніка. — 2010. — Т. 9, № 4. — С. 574–576.
6. *Open Computer Vision Library* [Электронный ресурс]. — Режим доступа: <http://opencvlibrary.sourceforge.net>
7. *Turán J.* Trace Transform Based Invariant Object Recognition System / J. Turán, D. Šiškovičová, P. Filo // Radioelektronika, 13-th International Czech–Slovak Scien. Conf. — Brno (Czech Republic), 2003. — P. 150–153. — [Электронный ресурс] — Режим доступа: [www.urel.feec.vutbr.cz/ra2007/archive/ra2003/papers/150.pdf](http://www.urel.feec.vutbr.cz/ra2007/archive/ra2003/papers/150.pdf)
8. *Petrou M.* Affine Invariant Features from the Trace Transform / M. Petrou, A. Kadyrov // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2004. — Vol. 26, N 1. — P. 30–44.

Поступила в редакцию 15.10.2011