

# Модели и методы сопоставления изображений в задаче распознавания лиц

**Аннотация.** В статье выполнен анализ предметной области и обоснована актуальность проблемы распознавания лиц. Рассмотрены методы сопоставления изображений с применением линий положения, сверток и инвариантов к группе аффинных преобразований для 2D- и 3D-изображений. Корректное сопоставление является необходимым этапом решения задачи распознавания. Приведены примеры применения метода линий положения для приведения изображений лиц к нормализованному виду.

**Ключевые слова:** графический образ, инвариантные моменты, аффинные преобразования, сопоставление изображений, распознавание.

## Введение

Определение ориентации и распознавание графических образов являются актуальными задачами машинного зрения и искусственного интеллекта, решение которых требуется для биометрической идентификации, поиска и наблюдения за людьми [1-7]. Несмотря на создание новых алгоритмов, методов и подходов, до настоящего времени нет реально работающих систем компьютерного зрения, способных обнаруживать и распознавать человека в сложных условиях. В последнее время наблюдается повышение интереса к задаче распознавания лиц, сопровождаемое проведением российских и международных конкурсов. Так в 2015 г. был проведен Открытый конкурс Фонда перспективных исследований [8] на лучший демонстрационный образец технологии распознавания лиц людей. В 2016 г. прошло мировое соревнование по распознаванию лиц на данных Megaface Challenge, организованное Вашингтонским Университетом (University of Washington) [9]. Безусловно, интерес вызывают работы ученых, посвятивших свои исследования различным проблемам сравнения и распознавания изображений. Эксперименты показывают, что 2D-модели имеют ограниченное

применение, поскольку их затруднительно применять для распознавания лиц при различных ракурсах головы, естественной мимики, наличии гримас и других помех. В связи с этим все большее внимание уделяется 3D-моделям, полученным с использованием камер высокого разрешения, применение которых в ряде случаев позволяет увеличить точность и полноту распознавания [10-14].

Отметим рост числа работ по распознаванию лиц на основе различных инвариантов к яркостным и аффинным преобразованиям. Однако при ближайшем изучении оказывается, что многие из них содержат те или иные неточности, что требует фундаментальной проверки инвариантности 2D-моментов, выполненной, например, в работе [7]. К сожалению, еще больше неточностей наблюдается в работах, посвященных различным трактовкам 3D-инвариантов [15-18].

Процедура наилучшего сопоставления является важнейшей составной частью задачи определения близости изображений с целью их кластеризации или распознавания [19, 20]. В настоящей работе основное внимание уделено задаче корректного сопоставления и сравнения изображений лиц с применением таких инструментальных средств как свертки, линии положения и инвариантные моменты. Одной из

задач является экспериментальная проверка и формирование набора 3D-инвариантов для распознавания лиц.

## 1. Задачи корректного сопоставления и сравнения двух плоских изображений

### 1.1. Метод свертки

Для оценки точности сопоставления изображений, заданных функциями яркости  $F_1(x, y)$  и  $F_2(x, y)$ , используются различные меры. В простейшем случае для определения параметров рассогласования может быть применена, например, взаимная корреляционная функция [19]. Математически задача наилучшего сопоставления двух полутоновых изображений, ограниченных прямоугольными областями, сводится к расчету и максимизации двойного интеграла (свертки изображений) [5]:

$$Z(\Delta x, \Delta y, k, \gamma) = \int \int_{(s)} F_1(x, y) F_2(x, y) dx dy,$$

где функция яркости  $F_1(x, y)$  задана дискретно на равномерной сетке  $(M_1, M_2)$  с размером ячейки  $l$ , а соответствующая прямоугольная область  $S$  интерпретируется как окно наблюдения или маска. Для цифровых изображений задача формулируется следующим образом:

$$Z(\Delta x, \Delta y, k, \gamma) = \sum_x \sum_y F_1(x, y) F_2(x, y) \rightarrow \max.$$

Маска содержит участок эталонного изображения. Наблюдаемое изображение с функцией яркости  $F_2(x, y)$ , заданной дискретно на равномерной сетке с размером ячейки  $L = k \cdot l$ , располагается в прямоугольном окне  $N \times N$ , которое смещено относительно маски по осям  $Ox$ ,  $Oy$  на величины  $\Delta x$ ,  $\Delta y$  и повернуто на угол  $\gamma$ , как это показано на Рис. 1.

Перебирая значения  $\Delta x$ ,  $\Delta y$  и  $\gamma$  на основе значений функционала определяют степень близости изображений. Наилучшему совмещению соответствует максимальное значение свертки. Для решения задачи вводятся две системы координат:  $A$  – с началом в левом нижнем углу малой области  $S$  (маски) и  $B$  – с началом в левом нижнем углу большой области (подложки). Задачу можно решать приближен-

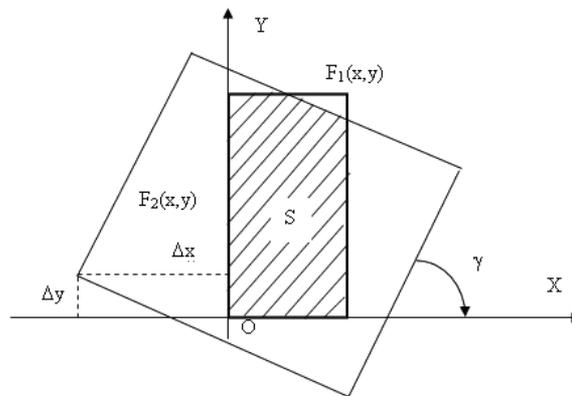


Рис. 1. Схема сопоставления маски с изображением

но различными методами. Суть одного из них в следующем:

- 1) выбор очередной точки в области  $S$  (точка соответствует одной ячейке);
- 2) смена системы координат  $A \rightarrow B$ ;
- 3) проверка - лежит ли выбранная точка в пределах подложки;
- 4) если да – выполняется произведение значений яркости ячеек, в которые попала точка, и результат добавляется в общую сумму;
- 5) возврат к п. 1.

Количество итераций алгоритма равно числу точек в маске. Данная схема может применяться для создания системы слежения за динамическим объектом.

Процесс подбора коэффициентов совмещения может быть ускорен за счет применения эффективного по быстродействию, но достаточно «грубого», метода определения ориентации полутоновых объектов на основе линий положения [6].

### 1.2. Сопоставление методом линий положений

Рассмотрим метод построения линий положения объекта, который представлен в виде бинарного или полутонового цифрового изображения [6]. Метод дает возможность определять ориентацию полутонового графического объекта без процедуры перебора или выделения пар связанных точек.

Линия положения  $y = a_0 + a_1 x$  проводится через центр тяжести изображения  $(\bar{x}, \bar{y})$  так, что

$$S = \sum_{i=1}^n s_i^2 f_i \rightarrow \min, \tag{1}$$

где  $s_i$  – расстояние от точки  $(x_i, y_i)$  с яркостью  $f_i$  до линии положения объекта,  $s_i = \frac{a_0 + a_1 x_i - y_i}{\sqrt{a_1^2 + 1}}$ . Пару значений  $(a_0, a_1)$ , удовлетворяющих условию, можно получить из уравнения:  $pa_1^2 + qa_1 - p = 0$ , где

$$p = \sum_{i=1}^n f_i \sum_{i=1}^n f_i x_i y_i - \sum_{i=1}^n f_i y_i \sum_{i=1}^n f_i x_i,$$

$$q = \sum_{i=1}^n f_i \left( \sum_{i=1}^n f_i x_i^2 - \sum_{i=1}^n f_i y_i^2 \right) + \left( \sum_{i=1}^n f_i y_i \right)^2 - \left( \sum_{i=1}^n f_i x_i \right)^2.$$

Решая уравнение, получаем пару значений  $(a_1^{(1)}, a_1^{(2)})$ . В соответствии с формулами

$$a_0 = \frac{\sum_{i=1}^n f_i y_i - a_1 \sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i}, \quad \bar{x} = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i},$$

$$\bar{y} = \frac{\sum_{i=1}^n f_i y_i}{\sum_{i=1}^n f_i}$$

получаем два значения другого коэффициента  $(a_0^{(1)}, a_0^{(2)})$ . Графическим решением являются две перпендикулярные линии, проходящие через центр тяжести  $(\bar{x}, \bar{y})$ , из которых главная соответствует минимальному значению критерия  $S$ . Рассмотренный алгоритм является составной частью технологии распознавания и применяется для нормализации положения объектов для их последующего сравнения.

Рассмотрим пример. Пусть даны три изображения лиц в различных ракурсах (Рис. 2). На Рис. 3 показан пример обработки алгоритма построения линий положения, на Рис. 4 - приведения исходных объектов к некоторому заранее установленному стандартному виду, в котором главная линия располагается перпендикулярно оси OX.

Нормализация на основе линий положения позволяет в общем случае «грубо» совмещать два изображения для их последующего сравнения с целью классификации или идентификации. Таким образом, метод дает ключ к сравнению двух изображений. Изображения помещаются в центр тяжести и разворачиваются до совпадения линий положения, после чего можно проводить поточечное сравнение (на самом деле нужно провести разворот одного изображения относительно другого дважды, так как линия положения определяет расположение объекта с точностью до  $180^\circ$ .) Метод имеет ограничения, он не применим в случае, когда



Рис. 2. Исходные изображения лиц

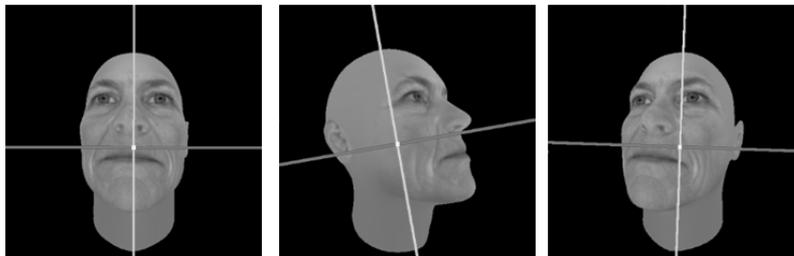


Рис. 3. Результат построения линий положения

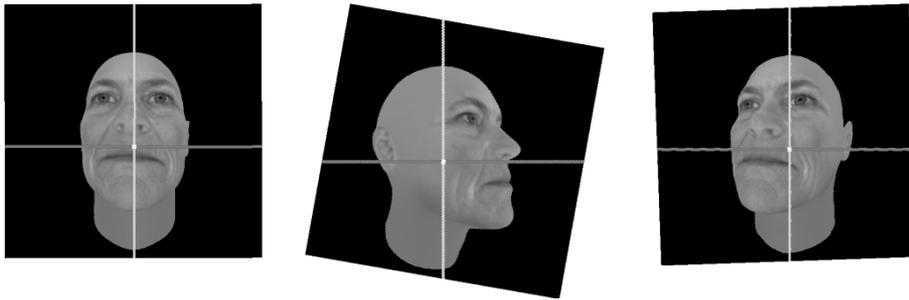


Рис. 4. Приведение изображений к нормализованному виду

квадратное уравнение не имеет решения. Развитие алгоритма на трехмерный случай выполнено в работе [9].

### 1.3. Сравнение на основе инвариантных моментов

Еще одним способом сравнения изображений является выделение дескрипторов в виде инвариантов к аффинным преобразованиям. Задача построения инвариантных моментов для плоского случая была рассмотрена в работе [1]. Среди всех геометрических инвариантов выделяются инварианты Ну, интерес к которым не ослабеваает и в настоящее время. Поскольку в первоисточниках отсутствуют доказательства утверждений об инвариантности, то заслуживают внимание теоретическая и экспериментальная проверки этого факта [2].

Инвариантные моменты позволяют проводить сравнение изображений при наличии воздействий на них аффинных преобразований. В Табл. 1 приведены значения трех первых инвариантных моментов Ну:

$$M_1 = m_{20} + m_{02}, M_2 = (m_{20} - m_{02})^2 + 4m_{11}^2, \\ M_3 = (m_{30} - 3m_{12})^2 + (3m_{21} + m_{03})^2,$$

где  $m_{pq} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q$ ;

$p+q \leq 3$ ,  $(x_i, y_i)$  – искомая точка изображения;  $(\bar{x}, \bar{y})$  – центр изображения;  $N$  – число точек изображения.

Метод обеспечивает инвариантность к поворотам, сдвигам и масштабированию только для идеальных (математических) объектов, как это показано в работе [2]. Для реальных же изображений с целочисленными значениями координат моментов, как это видно из Табл. 1, флуктуируют, так как операция поворота приводит к геометрическим искажениям. Другими словами, инвариантные моменты также являются достаточно грубым методом сопоставления дискретных изображений. Этот факт следует учитывать при формировании классов изображений в системах распознавания.

Табл. 1. Проверка инвариантности моментов к поворотам

	Варианты расположения 2D-изображений лиц			
$M_1$	0.33666	0.33631	0.30231	0.30245
$M_2$	0.03828	0.03805	0.01254	0.01261
$M_3$	5.68051	6.60609	21.43457	24.89683

## 2. Задачи корректного сопоставления и сравнения 3D-изображений

### 2.1. Сопоставление на основе линий положения

Как и в случае двумерных объектов, «грубое» сопоставление объектов возможно на основе определения ориентации тела в трехмерном пространстве. Примером подобного исследования, посвященного поиску специальных прямых линий на изображении, служит работа [20]. Задача состоит в нахождении проективного преобразования, которое переводит набор прямых на первом изображении в набор прямых на втором изображении. Отмечается, что сложность задачи состоит в следующем: соответствия между прямыми линиями на изображениях неизвестны; количество найденных на изображениях прямых линий может быть различно; прямые, найденные на первом изображении, могут не быть найдены на втором изображении и наоборот. Лучшим считается преобразование, для которого сумма квадратов интенсивностей пикселей разностного кадра минимальна. При вычислении разностного кадра учитывается только пересекающаяся часть обоих кадров.

Рассмотрим метод определения ориентации тела по заданным точкам, расположенным на гранях трехмерной модели или на поверхности, в соответствии с положениями работ [3-4], которые снимают частично указанные проблемы. Известно, что три вектора однозначно задают положение и ориентацию несимметричного трехмерного объекта в пространстве. Как и в плоском случае, алгоритм основан на проведении специальных ортогональных линий положения, что позволяет измерять углы ориентации и выполнять нормализацию объекта для последующего распознавания. В работе [3] задача определения ориентации была поставлена для трехмерного случая и решалась численным методом.

#### 2.1.1. Постановка и пример решения задачи

Пусть графический объект задан перечислением координат  $(x_i, y_i, z_i)$  и яркостей  $f_i$  своих точек,  $i = 1, \dots, N$ . Необходимо провести пространственную линию положения

$$\frac{x - \bar{x}}{l} = \frac{y - \bar{y}}{m} = \frac{z - \bar{z}}{n},$$

так, чтобы выполнялось условие (1). Здесь  $(\bar{x}, \bar{y}, \bar{z})$  – координаты центра тяжести объекта:

$$(\bar{x}, \bar{y}, \bar{z}) = \left( \frac{\sum_{i=1}^N f_i x_i}{\sum_{i=1}^N f_i}, \frac{\sum_{i=1}^N f_i y_i}{\sum_{i=1}^N f_i}, \frac{\sum_{i=1}^N f_i z_i}{\sum_{i=1}^N f_i} \right);$$

$l, m, n$  – коэффициенты, определяющие положение объекта в пространстве (должны быть определены в результате решения задачи).

В работах [3, 4] показано, что решение задачи сводится к системе уравнений:

$$\begin{pmatrix} l \\ m \\ n \end{pmatrix} \times \begin{pmatrix} A & D & E \\ D & B & F \\ E & F & C \end{pmatrix} \cdot \begin{pmatrix} l \\ m \\ n \end{pmatrix} = 0 \quad (3)$$

где  $(\times), (\cdot)$  – знаки векторного и скалярного произведений,

$$\begin{aligned} A &= \sum_{i=1}^N (x_i - \bar{x})^2 f_i, & B &= \sum_{i=1}^N (y_i - \bar{y})^2 f_i, \\ C &= \sum_{i=1}^N (z_i - \bar{z})^2 f_i, & D &= \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) f_i, \\ E &= \sum_{i=1}^N (x_i - \bar{x})(z_i - \bar{z}) f_i, \\ F &= \sum_{i=1}^N (y_i - \bar{y})(z_i - \bar{z}) f_i. \end{aligned}$$

Система (3) может быть записана в виде:

$$w \times I \cdot w = 0, \quad |w| = 1, \quad (4)$$

где  $w = (l \ m \ n)^T$ ,  $I = \begin{pmatrix} A & D & E \\ D & B & F \\ E & F & C \end{pmatrix}$ .

*Утверждение 1.* Система (4) имеет решение тогда и только тогда, когда  $Iw = \lambda w$ , где  $\lambda$  – некоторое число. Доказательство приведено в работе [4].

Собственные векторы являются ортогональными и соответствуют направляющим векторам эллипсоида, определяющего положение тела в пространстве. Длины осей эллипсоида соответствуют собственным числам, причем максимальное собственное (характеристическое) число соответствует искомому направлению линии положения. Вектор, соответствующий второму по величине собственному значению, определяет направление, в котором обращен объект. Третий собственный вектор определяет поворот на угол вокруг главной оси.

Табл.2. Точки на поверхности конуса

$i$	$x_i$	$y_i$	$z_i$	$i$	$x_i$	$y_i$	$z_i$
1	2.014	0.3278	20.41	14	-2.832	1.993	34.63
2	-2.12	3.938	44.73	15	0.562	0.903	10.64
3	0.515	-0.155	5.373	16	0.029	-1.490	14.91
4	-0.01	-0.7949	7.950	17	0.644	4.772	48.15
5	-2.024	-1.110	23.08	18	-2.84	-2.56	38.20
6	0.690	-0.244	7.320	19	0.568	1.392	15.03
7	1.294	0.4052	13.56	20	-0.686	3.018	30.95
8	-1.597	3.828	41.48	21	-2.472	-1.93	31.34
9	0.1240	0.2372	2.676	22	2.735	-0.27	27.48
10	2.58	-1.85	31.79	23	-0.068	0.8384	8.411
11	3.136	3.103	44.12	24	0.041	2.328	23.29
12	-0.811	-1.820	19.92	25	1.22	-4.51	46.76
13	1.60	-1.54	22.18				

Рассмотрим применение подхода на простом примере из работы [4], где бинарный объект представлен 25 точками, расположенными на конусе  $x^2 + y^2 = z^2 / 100$  (Табл. 2).

Центр масс находится в точке (0.992, 0.352, 24.57). Собственные значения равны (4792.49, 121.27, 69.82). Они соответствуют собственным векторам (0.018, -0.069, -0.998), (0.043, -0.998, 0.04), (1.167, 0, 0). Полученные собственные значения и вектора позволяют приводить объекты к стандартному положению для последующего сравнения.

## 2.2. Сопоставление на основе 3D-инвариантов

3D-инвариантные моменты являются удобным средством для решения задач прямого сравнения и распознавания лиц, подвергаемых различным преобразованиям и искажениям. Так, например, в результате наблюдения сцен с помощью стереопары можно восстановить 3D-образ лица и сравнить его с соответствующими 3D-моделями, хранящимися в базе данных лиц на основе вычисленных инвариантов.

### 2.2.1. Постановка задачи

Пусть  $\rho(x, y, z)$  есть непрерывная функция, описывающая значения яркости точек изображения с координатами  $(x, y, z)$  в трехмерном пространстве. Для корректного сопоставления изображений требуется построить моменты, инвариантные к группе аффинных преобразований. В соответствии с работой [15] можно вычислить центральные моменты:

$$\mu_{lmn} = \sum_X \sum_Y \sum_Z (x - \bar{x})^l (y - \bar{y})^m (z - \bar{z})^n \rho(x, y, z),$$

где  $X, Y, Z$  – область определения координат пикселей изображения;  $(l + m + n)$  – порядок 3D-момента;  $(\bar{x}, \bar{y}, \bar{z})$  – координаты центра тяжести (центроид) 3D-объекта.

Требуется построить 3D-моменты, инвариантные к операциям поворота, сдвига и масштабирования. Для этого необходимо сгруппировать определенным образом центральные моменты, например, в соответствии с рекомендациями [15, 18, 21].

В соответствии с работой [15] выбраны следующие моменты:

$$I_1 = \mu_{200} + \mu_{020} + \mu_{002};$$

$$I_2 = \mu_{200}\mu_{020} + \mu_{200}\mu_{002} + \mu_{020}\mu_{002} - \mu_{401}^2 - \mu_{410}^2 - \mu_{011}^2.$$

К сожалению, другие предложения этой работы содержат различные неточности.

Выбраны два момента из работы [18]:

$$I_3 = \mu_{200}\mu_{020}\mu_{002} - \mu_{002}\mu_{410}^2 - \mu_{020}\mu_{401}^2 - \mu_{200}\mu_{011}^2 + 2\mu_{410}\mu_{401}\mu_{011};$$

$$F_1 = \mu_{003}^2 + 6\mu_{012}^2 + 6\mu_{021}^2 + 6\mu_{030}^2 + 6\mu_{402}^2 + 15\mu_{411}^2 - 3\mu_{402}\mu_{420} + 6\mu_{420}^2 - 3\mu_{021}\mu_{201} + 6\mu_{201}^2 - 3\mu_{003}(\mu_{021} + \mu_{201}) - 3\mu_{030}\mu_{210} + 6\mu_{210}^2 - 3\mu_{012}(\mu_{030} + \mu_{210}) - 3\mu_{402}\mu_{300} - 3\mu_{420}\mu_{300} + \mu_{300}^2.$$

В набор включены четыре момента из работы [21].

$$F_2 = \mu_{200}^2 + \mu_{020}^2 + \mu_{002}^2 + 2\mu_{110}^2 + 2\mu_{101}^2 + 2\mu_{011}^2;$$

$$F_3 = \mu_{200}^3 + 3\mu_{200}\mu_{110}^2 + 3\mu_{200}\mu_{101}^2 + 3\mu_{110}^2 + 3\mu_{101}^2\mu_{020} + 3\mu_{101}^2\mu_{002} + \mu_{020}^3 + 3\mu_{020}\mu_{011}^2 + 3\mu_{011}^2\mu_{002} + \mu_{002}^3 + 6\mu_{110}\mu_{101}\mu_{011};$$

$$F_4 = \mu_{300}^2 + \mu_{030}^2 + \mu_{003}^2 + 3\mu_{210}^2 + 3\mu_{201}^2 + 3\mu_{120}^2 + 3\mu_{102}^2 + 3\mu_{021}^2 + 3\mu_{012}^2 + 6\mu_{111}^2;$$

$$F_5 = \mu_{300}^3 + 2\mu_{300}\mu_{120} + 2\mu_{300}\mu_{102} + 2\mu_{210}\mu_{030} + 2\mu_{210}\mu_{030} + 2\mu_{201}\mu_{003} + \mu_{030}^2 + 2\mu_{030}\mu_{012} + 2\mu_{021}\mu_{003} + \mu_{003}^2 + \mu_{210}^2 + 2\mu_{210}\mu_{012} + 2\mu_{201}\mu_{021} + \mu_{120}^2 + 2\mu_{120}\mu_{102} + \mu_{102}^2 + \mu_{021}^2 + \mu_{012}^2.$$

Установленный набор создает определенные возможности для распознавания лиц и может быть пополнен за счет моментов более высокого порядка. Покажем, что моменты  $I_1, \dots, F_5$  действительно являются инвариантами к преобразованиям поворота и сдвига. Заметим, что вопросы инвариантности 3D-моментов к масштабированию в настоящей статье не рассматриваются и являются предметом отдельного изучения. Дополнительно оценим их чувствительность к линейным преобразованиям, вызванным помехами.

### 2.2.2. Проверка инвариантности и чувствительности моментов

Ранее в работе [2] была выполнена теоретическая проверка инвариантности и чувствительности группы 2D-инвариантов путем прямого воздействия аффинными преобразованиями на переменные в формулах для моментов.

*Утверждение 2.* Моменты  $I_1 - F_5$  являются 3D-инвариантами к операциям поворота и сдвига.

В Табл. 3 в качестве примера приведены значения моментов для трех различных положений выбранной модели 3D-поверхности лица. Исследования подтверждают правильность Утверждения 2, что позволяет практически использовать моменты  $I_1, \dots, F_5$  для решения задач сопоставления изображений.

Практика распознавания реальных графических образов свидетельствует о том, что признаки в виде инвариантных моментов могут иметь различную информативность применительно к конкретным задачам, что требует их дополнительного изучения.

Оценим чувствительность моментов к флуктуациям данных. Положим, что  $\delta x_i = \delta y_i = \delta z_i$  – возможные отклонения измеряемых значений  $x_i, y_i, z_i$ ,  $0 < \delta \leq 1$ , вызванные различными причинами. Тогда, для центрального момента  $\mu_{111}$ :

Табл.3. Пример применения 3D-инвариантов

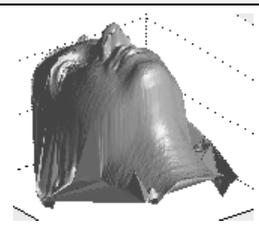
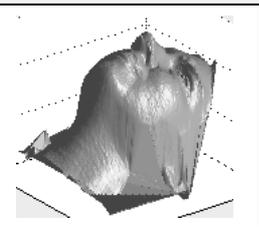
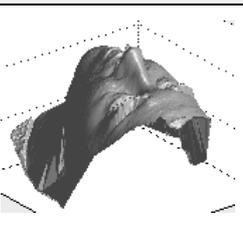
	Варианты расположения 3D-изображения лица		
			
$I_1$	1.01998e+08	1.01998e+08	1.01998e+08
$I_2$	2.31425e+15	2.31425e+15	2.31425e+15
$I_3$	1.36162e+22	1.36162e+22	1.36162e+22
$F_1$	6.63739e+18	6.63739e+18	6.63739e+18
$F_2$	5.77502e+15	5.77502e+15	5.77502e+15
$F_3$	3.93839e+23	3.93839e+23	3.93839e+23
$F_4$	5.8625e+18	5.8625e+18	5.8625e+18
$F_5$	5.34591e+18	5.34591e+18	5.34591e+18

Табл. 4. Чувствительность инвариантных моментов к флуктуации данных

Моменты	$I_1$	$I_2$	$I_3$	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$
Чувствительность	$\delta^2$	$\delta^4$	$\delta^6$	$\delta^6$	$\delta^4$	$\delta^6$	$\delta^6$	$\delta^6$

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})(\delta y_i - \delta \bar{y})(\delta z_i - \delta \bar{z}) = \\ & = \delta^3 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})(z_i - \bar{z}) = \delta^3 \mu_{111}. \end{aligned}$$

Отсюда следует, что чувствительность момента  $\mu_{111}$  к изменениям входных данных пропорциональна или равна  $\delta^3$ . Аналогично можно получить значения чувствительности других центральных моментов, а затем и инвариантов  $I_1, \dots, F_5$  без учета масштабирования (Табл. 4).

Полученные оценки могут быть учтены при отборе моментов для решения практических задач распознавания.

### 2.2.3. Распознавание на основе инвариантных моментов и нейронных сетей

Для решения задачи распознавания лиц может оказаться целесообразным объединение метода инвариантных моментов с искусственными нейронными сетями (ИНС) [1]. В случае потоковой обработки изображений не исключено и привлечение высокопроизводительных вычислителей [22].

Пусть дано обучающее множество изображений лиц  $M$ , каждое из которых формализуется вектором значений признаков  $X_i = (x_{i1}, \dots, x_{in})$ ;  $i = 1, \dots, m$ ;  $n$  – количество признаков. Заметим, что в нашем случае признаками являются инвариантные моменты. Векторы отнесены экспертами к некоторым классам  $\Omega_l$ ,  $l = 1, \dots, k$ ,  $M = \bigcup_{l=1}^k \Omega_l$ . Вся выборка

подразделена на два непересекающихся подмножества: обучающее и тестовое. После обучения ИНС качество распознавания проверяется на тестовой части.

Для распознавания графических образов может применяться ИНС прямого распространения с обучением методом обратного распространения ошибки [1]. Нейроны скрытого слоя

имеют функцию активации типа «сигмоид», а выходного слоя – «линейный скачок». Альтернативой служит вероятностная ИНС, которая учитывает плотности распределения вероятностей значений признаков для каждого класса [23]. В обоих случаях при подаче на вход вектора признаков неизвестного образца будет выбран класс, которому соответствует максимальное значение выхода.

## Заключение

В работе рассмотрены некоторые инструментальные средства, необходимые для корректного сопоставления и сравнения 2D- и 3D-моделей объектов. Инструментарий предназначен для решения задач распознавания графических образов и, в частности, лиц. Задачи являются актуальными, но пока далеки от своего окончательного решения в силу достаточно высокой сложности. Ни один из рассмотренных подходов, к сожалению, не обеспечивает получения решения с гарантированной точностью. Из недостатков отметим возможность ошибки распознавания, ограниченность применения методов в условиях помех. Однако рассмотренные подходы вполне годятся для «грубого» решения задачи сопоставления объектов и их последующего сравнения, в том числе на основе инвариантов. Предполагается, что в дальнейшем на основе яркостных и геометрических инвариантов будут практически решены задачи распознавания лиц человека.

## Литература

1. Хачумов М.В., Нгуен З.Т. Задача распознавания лиц по фотографиям на основе инвариантных моментов // Современные проблемы науки и образования, №2, 2015. URL: <http://www.science-education.ru/pdf/2015/2-2/855.pdf> (дата обращения: 10.10.2016).
2. Нгуен З.Т. Инварианты в задачах распознавания графических образов // Вестник Российского университета дружбы народов. Серия: Математика, Информатика, Физика. №1, 2016, с.76-85.
3. Алибеков А.Г., Лагиева М.М., Хачумов В.М. Определение ориентации трехмерных графических объек-

- тов//Известия вузов. Приборостроение. Т.38, №3-4, 1995, с.35-37.
4. Трушков В.В., Хачумов В.М. Определение ориентации объектов в трехмерном пространстве//Автометрия. 2008, №3, с.75-79.
  5. Недев М.Д., Талалаев А.А., Тищенко И.П., Хачумов В.М. Задачи распознавания, географической привязки и наблюдения объектов на основе анализа полутонных снимков//Авиакосмическое приборостроение. №12, 2009, с.19-24.
  6. Лагиева М.М., Хачумов В.М., Шабалов Д.В. Метод построения линий положения для идентификации полутонных изображений//Автометрия. №6, 1991, с.7-12.
  7. Нгуен З.Т. Анализ инвариантных моментов в задачах масштабирования и вращения изображений/Тезисы докладов Всероссийской конференции (с международным участием) «Информационно-телекоммуникационные технологии математическое моделирование высокотехнологичных систем» (ИТТММ-2016, РУДН). М.: РУДН, 2016, с.156-159.
  8. Открытый конкурс на лучший демонстрационный образец технологии распознавания лиц людей. URL: <http://fpi.gov.ru/activities/ideas/face> (дата обращения: 10.10.2016).
  9. Программисты из России повторили успех коллег в первенстве на распознавание лиц. URL: <https://lenta.ru/news/2016/09/06/ntechlab1/> (дата обращения: 10.10.2016).
  10. Naser Zaeri, Faris Baker. Thermal Face Recognition Using Moments Invariants // International Journal of Signal Processing Systems. Vol. 3, No. 2, December 2015, pp. 94-99.
  11. Akihiro Hayasaka, Takuma Shibahara, Koichi Ito, Takafumi Aoki, Hiroshi Nakajima and Koji Kobayashi. A 3D Face Recognition System Using Passive Stereo Vision and Its Performance Evaluation // Intelligent Signal Processing and Communications, 2006. ISPACS '06. International Symposium on Intelligent Signal Processing and Communications, pp.379-382.
  12. Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Expression-Invariant 3D Face Recognition // J. Kittler and M.S. Nixon (Eds.): AVBPA 2003, LNCS 2688, pp.62-70.
  13. Манолов А.И., Соколов А.Ю., Степаненко О.В., Тумачек А.С., Тяхт А.В., Цискаридзе А.К., Заварикин Д.Н., Кадейшвили А.А. Некооперативная биометрическая идентификация по 3D-моделям лица с использованием видеокамер высокого разрешения. - [http://www.graphicon.ru/html/2009/conference/se8/139/139\\_Paper.pdf](http://www.graphicon.ru/html/2009/conference/se8/139/139_Paper.pdf) (дата обращения: 10.10.2016).
  14. Gilles Burel, Hugues Henocq. 3D invariants and their application to object recognition // Signal Processing, Vol. 45, No. 1, July 1995, pp.1-22.
  15. Nita M. Thakare, V.M. Thakare. A Supervised Hybrid Methodology for Pose and Illumination Invariant 3D Face Recognition. – International Journal of Computer Applications (0975 – 8887) Volume 47, № 25, June 2012, 24. URL: <http://research.ijcaonline.org/volume47/number25/pxc3880474%20.pdf> (дата обращения: 10.10.2016).
  16. Dong Xu, Hua Li. Geometric moment invariants // Pattern Recognition 41. 2008, pp.240-249.
  17. Homogeneous invariants. URL:<http://zoi.utia.cas.cz/files/rot3dinvs8web.pdf> (дата обращения: 10.10.2016).
  18. Ingolf Sommer, Oliver Müller, Francisco S. Domingues, Oliver Sander, Joachim Weickert and Thomas Lengauer. Moment invariants as shape recognition technique for comparing protein binding sites // Bioinformatics, Vol. 23 no. 23 2007, pp.3139-3146.
  19. Бессмельцев В.П., Булушев Е.Д. Быстрый алгоритм совмещения изображений для контроля качества лазерной микрообработки//Компьютерная оптика. №2, 2014, с.343-350.
  20. Волегов Д.Б., Юрин Д.В. Предварительное грубое совмещение изображений по найденным на них прямым линиям для построения мозаик, сверхразрешения и восстановления трехмерных сцен// Программирование. №5, 2008, с.47-66.
  21. Suk T., Flusser J. 3D rotation invariants // Department of Image, 2012. URL: <http://zoi.utia.cas.cz/3DRotationInvariants> (дата обращения: 10.10.2016).
  22. Хачумов В.М., Фраленко В.П. Высокопроизводительная обработка изображений на кластерных устройствах//Нейрокомпьютеры: разработка, применение. №6, 2012, с.38-46.
  23. Хачумов В.М. Графические образы и нейронные сети/Тезисы докладов 4-й международной конференции «Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта CAD/CAM/PDM–2004. М.: ИПУ РАН, 2004, с.35.

**Нгуен Зуй Тхань.** Аспирант Российского университета дружбы народов (РУДН). Окончил Томский политехнический университет в 2008 году. Автор 6 печатных работ. Область научных интересов: распознавание образов, искусственный интеллект, информационные технологии и системы. E-mail: [nguyenduythanh1410@gmail.com](mailto:nguyenduythanh1410@gmail.com)

**Хачумов Вячеслав Михайлович.** Заведующий лабораторией ФИЦ ИУ РАН, профессор кафедры информационных технологий РУДН. Окончил Ленинградский кораблестроительный институт в 1971 году. Доктор технических наук. Автор более 200 печатных работ. Область научных интересов: разрядно-параллельные и конвейерные вычислительные системы, распознавание образов, искусственный интеллект. E-mail: [vmh48@mail.ru](mailto:vmh48@mail.ru)

## **Models and methods for matching images in the problem of face recognition**

Nguyen Duy Thanh, V.M. Khachumov

**Abstract.** In the article the analysis of the subject area is given and the relevance of the problem of face recognition is proved. Image matching methods with the use of position lines, convolution and invariants under the group of affine transformations for 2D and 3D images are considered. Correct comparison is a necessary stage of the recognition problem solution. Examples of position lines method application for normalizing face images.

**Keywords:** graphic image, invariant moments, affine transformation, image comparison, recognition.

**Nguyen Duy Thanh**, graduated from Tomsk Polytechnic University in 2008, specialty "Information systems and technologies", postgraduate student of the Peoples' Friendship University of Russia (RUDN), author of 6 publications. Research interests: pattern recognition, artificial intelligence, information technologies and systems. E-mail: nguyenduythanh1410@gmail.com

**Khachumov Vyacheslav Mikhailovich**, Ph. D., head of the laboratory "Intelligent control" ISA FRC CSC RAS, Professor of the Department of information technologies at the Peoples' Friendship University of Russia (RUDN), author of over 200 publications. Research interests: bit-parallel and conveyor computing systems, pattern recognition, artificial intelligence. E-mail: vmh48@mail.ru