

4. *E. Ukkonen*. Approximate String Matching over Suffix-Trees. In Proceedings of the Fourth Annual Symposium on Combinatorial Pattern Matching, Padova, Italy, June. 1993. – P. 229-242.
5. *D.R. Morrison*. PATRICIA - practical algorithm to retrieve information coded in alphanumeric. Journal of the ACM, 15(4):514-534 (1968).
6. *Kahveci, T.M.* Proceedings of the 27th International Conference on Very Large Databases // T. Kahveci, Ambuj K. Singh // An Efficient Index Structure for String Databases. 2001. – P. 351-360.
7. *Resnikoff, H.L.* The Nature of Affixing in Written English. Part 1, in Mechanical Translation, 8, No. 3 (1965), Part 11 in Mechanical Translation 9, No. 2 (1966).
8. *Shang, H.G.* Tries for Approximate String Matching – H. Shang T.H. Merret – In IEEE Transactions on Knowledge and Data Engineering, volume 8(4). 1996. – P. 540 – 547,

Злыгостев Илья Сергеевич

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: workmaking@gmail.com.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 8(8634)371-606.

Кафедра высшей математики; аспирант.

Zlygostev Iliya Sergeevich

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: workmaking@gmail.com.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 8(8634)371-606.

The Department of Higher Mathematics; post-graduate student.

УДК 004.932.72

А.Ю. Антипова, В.В. Губарев

**ПРИМЕНЕНИЕ МЕТОДА МОМЕНТОВ В ЗАДАЧЕ ГЕОМЕТРИЧЕСКОГО
ВЫРАВНИВАНИЯ ЛИЦ НА ИЗОБРАЖЕНИЯХ***

Данная работа посвящена методу улучшения алгоритмов обработки цифровых изображений лиц. Для определения угла наклона лица в плоскости изображения предлагается модель, основанная на понятии эллипса рассеяния и вычислении моментов изображения. Проанализирована применимость этого метода для различных цветовых пространств и методов фильтрации. Приведены результаты вычислительных экспериментов.

Изображения; детекция лиц; метод моментов; геометрическое выравнивание; цветовые пространства; фильтрация.

A.Y. Antipova, V.V. Gubarev

**APPLICATION OF METHOD OF MOMENTS IN PROBLEM
OF GEOMETRICAL ALIGNMENT OF FACES IN IMAGES**

This work focuses on methods of improving algorithms of processing digital images of individuals. To determine the angle of inclination of face in image's plane a model based on the notion of concentration ellipse and computation of image moments is proposed. Application of this

* Работа выполнена при финансовой поддержке РФФИ, проект №08-07-00129, №07-07-00067.

method for different color spaces and methods of filtering is analyzed. The results of computational experiments are given.

Images; face detection; method of moments; geometrical alignment; color spaces; filtering.

Автоматическая обработка изображений является одной из важных и востребованных информационных технологий. Методы и алгоритмы автоматического обнаружения и распознавания лиц используются в широком спектре современных систем компьютерного зрения: биометрической идентификации, распознавании эмоций, зрении роботов, компьютерной анимации, видеоконференции. Область применения технологий распознавания изображений во многом определяет специфику разработанных алгоритмов.

Обнаружение лиц на изображении является одной из наиболее востребованных задач анализа изображений. Важно, чтобы система автоматической обработки фото- и видеопотока работала быстро и эффективно. Для уменьшения вычислительной нагрузки алгоритмов можно применять методы предварительной обработки изображений. К таким методам можно отнести: фильтрацию, сегментацию, переход в другое цветовое пространство и т.п.

Геометрическое выравнивание лиц в алгоритмах детекции

Большинство методов обнаружения лиц работают по следующей схеме [1]:

- ◆ изображение последовательно сканируется окном определенного размера (рис. 1);
- ◆ сканирование происходит на различных уровнях масштаба изображения;
- ◆ для каждого положения окна в изображении оценивается вероятность того, что содержимое данного окна является лицом.



Рис. 1. Детекция лиц

Такой алгоритм детекции эффективен лишь при фронтальном расположении лица на изображении. Существует огромное количество разнообразных методов детекции лиц. На сегодняшний день одним из лучших является детектор фронтальных лиц Виола/Джонса (Viola/Jones) [2]. Этот алгоритм состоит из каскада «легковесных» классификаторов. Каждый классификатор определяет некоторый признак рассматриваемого объекта. Данный детектор реализован в библиотеке OpenCV и является открытым [3].

В случае сильно повернутого лица поиск не даст положительных результатов. Один из способов улучшения такого алгоритма заключается в сканировании изображения не только для различных масштабов, но и для различных углов поворота рамки детектора. Недостатком такого подхода является высокая вычислительная сложность. Другим примером решения данной проблемы является алгоритм Роули (Rowley) [4]. Роули описывает работу двухслойной нейронной сети, первый слой которой получает часть изображения заданного размера, подвергает предваритель-

ной обработке, определяет угол наклона лица и выравнивает его. При этом точно неизвестно, на самом ли деле это лицо. Затем повернутое изображение передается на вход нейронной сети обычного детектора. Существенными недостатками детектора Роули являются затраты на обучение и организацию нейронной сети, а также высокая вычислительная сложность алгоритма, что является крайне важным при построении систем, работающих в реальном времени.

Данная работа посвящена методам геометрического выравнивания лиц на цифровых изображениях. Под геометрическим выравниванием будем понимать поворот лица (рис. 2), при котором лицо, представленное на изображении, достигает наиболее «естественного» положения с точки зрения человеческого восприятия.

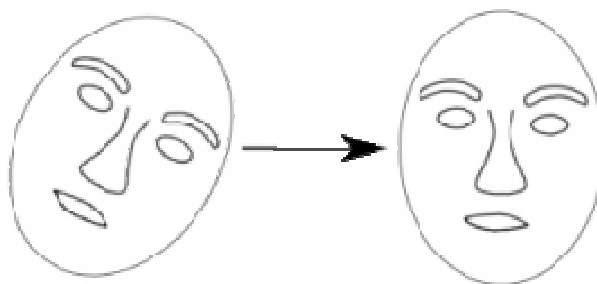


Рис. 2. Выравнивание лица

Рассмотренные в работе методы позволяют повысить точность и скорость работы алгоритмов детекции и распознавания лиц.

Геометрическое выравнивание лиц методом моментов

Метод моментов точечной оценки неизвестных параметров, заданного распределения, состоит в приравнивании теоретических моментов соответствующим эмпирическим моментам того же порядка. Для решения задачи геометрического выравнивания лиц на изображениях разработан алгоритм вычисления угла поворота лица, основанный на методе моментов.

Рассмотрим некоторую двумерную случайную величину (ξ, η) с дискретным распределением, задаваемым таблицей:

$$P(\xi = a_i, \eta = b_j) = p_{ij}, \quad i = 0, 1, \dots, \quad j = 0, 1, \dots$$

Двумерный начальный момент порядка $l + k$ данного распределения имеет вид

$$\alpha_{lk} = E(\xi^l \eta^k) = \sum_{i,j} \xi_i^l \eta_j^k p_{ij}.$$

В частности, начальные моменты нулевого и первого порядков определяются выражениями:

$$\alpha_{00} = E(\xi^0 \eta^0) = \sum_{i,j} p_{ij} = 1, \quad \alpha_{10} = E(\xi) = m_\xi, \quad \alpha_{01} = E(\eta) = m_\eta, \quad (1)$$

где m_ξ, m_η – математические ожидания одномерных частных распределений ξ и η соответственно.

Начальные моменты второго порядка имеют вид [5]:

$$\begin{aligned}
\alpha_{20} &= E(\xi^2) = \sum_i \xi_i^2 p_i = \sigma_\xi^2 + m_\xi^2, \\
\alpha_{02} &= E(\eta^2) = \sum_j \eta_j^2 p_j = \sigma_\eta^2 + m_\eta^2, \\
\alpha_{11} &= E(\xi\eta) = \sum_{i,j} \xi_i \eta_j p_{ij},
\end{aligned} \tag{2}$$

где $\sigma_\xi^2, \sigma_\eta^2$ – стандартные отклонения.

С физической точки зрения начальные моменты нулевого, первого и второго порядков можно интерпретировать как «массу», «центр масс» и «момент инерции» некоторой области.

Эллипс рассеяния служит геометрической характеристикой разброса значений случайного вектора. Если средние значения и дисперсии двумерного случайного вектора (X, Y) равны $m_\xi, m_\eta, \sigma_\xi^2, \sigma_\eta^2$ соответственно, а коэффициент корреляции равен $\rho_{\xi\eta}$, то эллипс рассеяния можно записать в виде

$$\frac{(x - m_\xi)^2}{\sigma_\xi^2} - 2\rho_{\xi\eta} \frac{(x - m_\xi)(y - m_\eta)}{\sigma_\xi \sigma_\eta} + \frac{(y - m_\eta)^2}{\sigma_\eta^2} = 4(1 - \rho_{\xi\eta}^2).$$

Центр эллипса находится в точке (m_ξ, m_η) , а его оси образуют с осью Ox углы φ и $\pi/2 + \varphi$, где φ определяется из условия [6]:

$$\operatorname{tg} 2\varphi = \frac{2\rho_{\xi\eta} \sigma_\xi \sigma_\eta}{\sigma_\xi^2 - \sigma_\eta^2}. \tag{3}$$

Для вычисления угла наклона головы будем подсчитывать моменты нулевого, первого и второго порядков по формулам (1), (2) внутри некоторой области, ограничиваемой окружностью.

В соответствии с формулой (3), угол φ наклона оси эллипса рассеяния определяется следующей формулой [7]:

$$\varphi = \frac{1}{2} \operatorname{arctg} \left(\frac{2(\alpha_{11} - \alpha_{10}\alpha_{01})}{(\alpha_{20} - \alpha_{10}^2) - (\alpha_{02} - \alpha_{01}^2)} \right), \tag{4}$$

где α_{10}, α_{01} – координаты центра эллипса рассеяния. Таким образом, угол φ и является искомым углом наклона головы.

Модельная задача

В качестве объекта рассмотрим круг, состоящий из двух областей, каждая из которых отличается оттенком. Изображение объекта считается одноканальным. В соответствии с формулой (4) осуществим геометрическое выравнивание данного объекта (рис. 3).

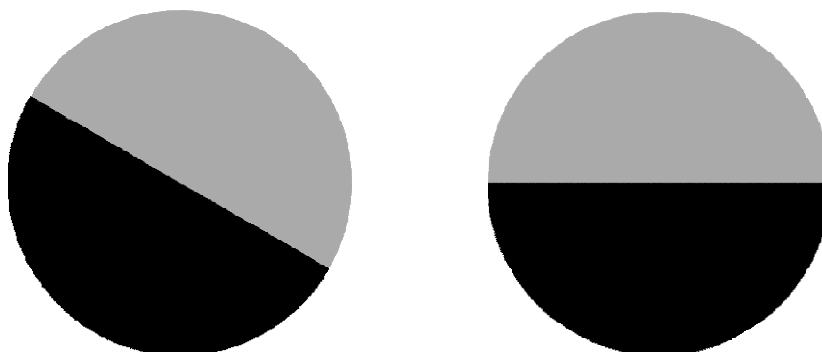


Рис. 3. Выравнивание неоднородного круга. Слева – исходное изображение, справа – выровненное изображение

Из рисунка видно, что для данной модельной задачи метод показал адекватный результат: объект принял «естественное» положение, при котором изображение стало симметричным относительно вертикальной оси.

Инвариантность к освещенности метода геометрического выравнивания лиц

Цветные изображения. Как отмечалось выше, в рамках данной задачи метод моментов следует применять к одноканальному изображению.

Цветное изображение состоит из нескольких каналов. Наиболее очевидным способом получения одноканального изображения является преобразование цветного изображения в полутоновое. Однако такой подход не является эффективным, так как левая и правая половины лица могут иметь различный уровень освещения. Как следствие, центр масс объекта, рассчитанный по формулам (1), будет смещенным, расчет угла – нестабильным. Таким образом, возникает вопрос о получении некоторого одноканального изображения, инвариантного относительно направления и уровня освещения.

Цветовое пространство предоставляет возможность описания цветов некоторым стандартным образом. По существу, цветовое пространство определяет некоторую систему координат, в котором каждый цвет представляется единственной точкой. Рассмотрим наиболее популярные цветовые пространства.

RGB – это аппаратно-ориентированное цветовое пространство, которое описывает цвет как комбинацию трех цветных каналов (красный, зеленый, синий). Это одно из наиболее часто используемых цветовых пространств для обработки и хранения изображений. Однако высокая корреляция между каналами и смешивание цветовой характеристики объекта с его освещенностью делает это пространство не таким привлекательным для решения некоторых задач.

Существуют такие цветовые пространства, переход в которые позволяет отделить информацию об освещении от других компонент [8, 9].

HSV – (от англ. hue, saturation, value – тон, насыщенность, значение) – цветовое пространство, в котором компонентами цвета являются:

- ◆ *Hue* – тон. Варьируется в пределах 0–360, однако иногда приводится к диапазону 0–100 или 0–1.
- ◆ *Saturation* – насыщенность. Варьируется в пределах 0–100 или 0–1. Чем больше этот параметр, тем «чище» цвет, а чем ближе к нулю – тем ближе цвет к нейтральному серому.

- ♦ *Value* – значение цвета. Имеет смысл освещенности. Варьируется в пределах 0–100 или 0–1.

Переход из RGB в HSV осуществляется с помощью нелинейного преобразования:

$$H = \begin{cases} 0, & \text{если } MAX = MIN, \\ 60 \times \frac{G - B}{MAX - MIN}, & \text{если } MAX = R, G \geq B, \\ 60 \times \frac{G - B}{MAX - MIN} + 360, & \text{если } MAX = R, G < B, \\ 60 \times \frac{B - R}{MAX - MIN} + 120, & \text{если } MAX = G, \\ 60 \times \frac{R - G}{MAX - MIN} + 240, & \text{если } MAX = B, \end{cases}$$

$$S = \begin{cases} 0, & \text{если } MAX = 0, \\ 1 - \frac{MIN}{MAX}, & \text{если } MAX > 0, \end{cases}$$

$$V = MAX,$$

где

$$MAX = \max(R, G, B),$$

$$MIN = \min(R, G, B).$$

Интуитивная понятность компонент этого цветового пространства и невысокая степень корреляции между цветовым признаком и освещенностью сделали данное пространство достаточно популярным в задачах сегментации по цвету. Для решения поставленной задачи целесообразно использовать цветовые каналы, наименее коррелированные с уровнем освещения: Hue, Saturation.

TSL – (от англ. tint, saturation, lightness – тон, насыщенность, освещенность) – цветовое пространство, основанное на нелинейном переходе из пространства RGB [10]. Данное пространство содержательно достаточно близко к пространству HSV. Переход в TSL осуществляется по формулам:

$$T = \begin{cases} \arctg(r'/g')/2\pi + 1/4, & g' > 0, \\ \arctg(r'/g')/2\pi + 3/4, & g' < 0, \\ 0, & g' = 0, \end{cases}$$

$$S = \sqrt{\frac{9}{5}(r'^2 + g'^2)},$$

$$L = 0.299R + 0.587G + 0.114B.$$

где $r' = r - 1/3$, $g' = g - 1/3$.

Значения r , g , b и определяются формулами:

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B}.$$

Эти значения называют нормализованным пространством rgb.

В рамках поставленной задачи интерес представляют каналы Tint и Saturation, так как информация об уровне освещения содержится в компоненте Lightness.

YCbCr представляет собой цветное пространство, переход в которое из пространства RGB осуществляется следующим способом:

$$Y = 0.299R + 0.587G + 0.114B,$$

$$C_b = B - Y,$$

$$C_r = R - Y.$$

где Y – компонента пространства, отвечающая за освещенность, C_b и C_r – компоненты, полученные из синего и красного каналов соответственно, путем вычитания из них значения освещенности. Аналогично предыдущим цветовым пространствам, компонента Y не подходит для решения данной задачи.

Данное цветное пространство является основным рабочим пространством телевизионных студий в Европе, а также применяется в некоторых алгоритмах сжатия изображений [11].

Полутоновые изображения

Если исходное изображение является полутоновым, то переход в некоторое цветное пространство является невозможным. В этом случае можно применить различные методы фильтрации изображений с целью уменьшения влияния неоднородности освещения.

Фильтрация изображения представляет собой пример локальных методов улучшения изображений [12]. Объектом локального преобразования является не все изображение в целом, а его отдельная область, являющаяся окрестностью некоторого пикселя. При пространственной фильтрации кроме этой области используется некоторая матрица тех же размеров. Эту матрицу называют маской, шаблоном, ядром или фильтром. Элементы такой матрицы называют коэффициентами маски. В результате некоторых вычислений в каждой области меняется значение центрального пикселя – отклик.

Обработка осуществляется путем перемещения маски по всему изображению и вычислению отклика для всех пикселей. Операция вычисления отклика может быть как линейной, так и нелинейной. В линейном случае отклик $d_{i,j}$ в точке $S_{i,j}$ изображения S вычисляется по следующей формуле:

$$d_{i,j} = \sum_{p=1}^n \sum_{q=1}^n s_{i+p-\frac{n+1}{2}, j+q-\frac{n+1}{2}} w_{p,q},$$

где $w = (w_{pq})$ – маска размером $n \times n$.

Изображение, инвариантное относительно уровня и направления освещения, можно получить с помощью выделения границ. Маски фильтров методов выделения границ называются градиентными.

В данной работе были использованы следующие методы: фильтрация оператором Собеля, фильтрация Лапласианом, алгоритм выделения краев Кэнни [13].

Вычислительный эксперимент

В работе [14] предлагается подобный метод обнаружения овала лица с помощью моментов: определяется центр и направляющий вектор главной оси эллипса рассеяния в нормализованном пространстве rgb . В статье [7] локализация лица осуществляется в пространстве HSV , а конкретно – в канале Hue .

Целью вычислительного эксперимента является определение наилучшего цветового пространства (и соответственно некоторой его компоненты), а также наилучшего преобразования для одноканальных изображений, позволяющего наиболее эффективно и точно вычислять угол наклона головы и осуществлять детекцию лиц.

В реализованных вычислительных экспериментах в качестве выборки изображений была взята база лиц AR [15], состоящая из цветных фронтальных изображений людей и файлов разметки особых точек лица (рис. 4). Количество таких изображений в базе – 370.



Рис. 4. Примеры изображений из базы AR

На основе координат центров зрачков позиционируется квадратная рамка. Эта рамка соответствует результату работы стандартного детектора лиц (например детектор Виола/Джонса [2]). Кроме того, все изображения базы были искусственно повернуты на заданные углы: -30° , -20° , -10° , 0° , 10° , 20° , 30° (рис. 5). Поворот осуществлялся относительно центра лица. Таким образом, была сформирована выборка размером $370 \times 7 = 2590$ изображений.

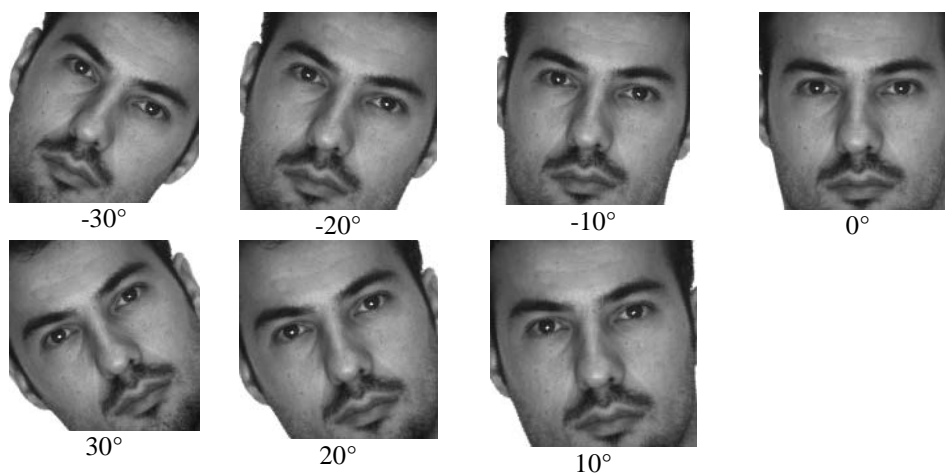


Рис. 5. Повернутые изображения

Для каждого изображения оценивался угол поворота лица с помощью описанного выше алгоритма. Оценка производилась для каждого варианта одноканального представления изображения (рис. 6).

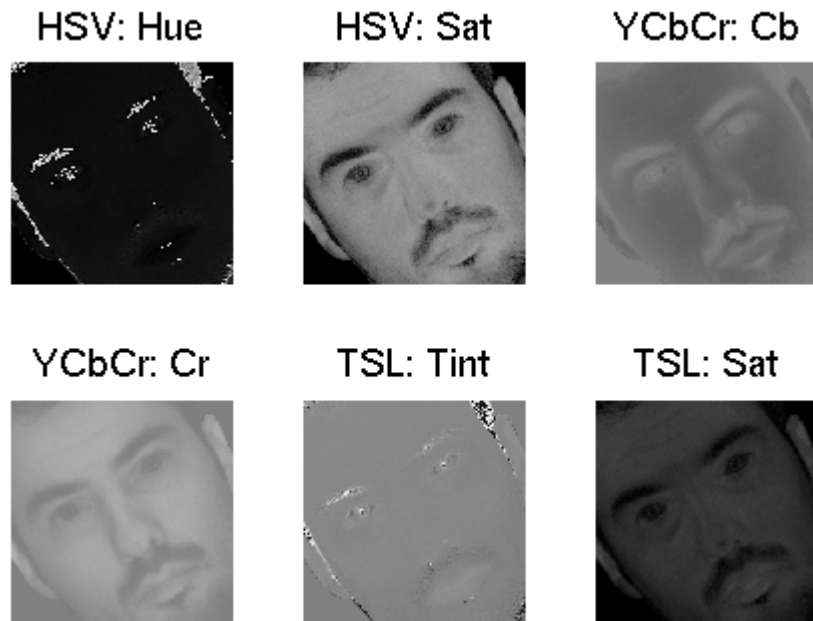


Рис. 6. Одноканальные представления изображения лица, повернутого на угол 30°

В результате работы программы вычислялся угол поворота лица. Рис. 7 демонстрирует результаты поворота лица на углы, рассчитанные по данным, представленным на рис. 6.

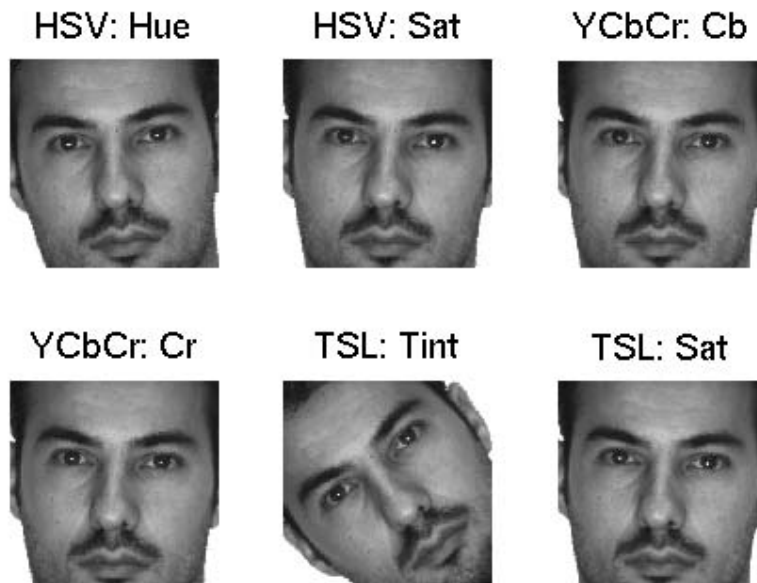


Рис. 7. Автоматическое выравнивание лица

Для оценки точности работы, предложенного алгоритма, результаты работы программы сравнивались с заданными на этапе формирования выборки значениями углов, которые представлены в табл. 1.

Таблица 1

Средняя ошибка

	HSV: Hue	HSV: Sat	YCbCr: Cb	YCbCr: Cr	TSL: Tint	TSL: Sat	So- bel	Lapla- cian	Canny
Средняя ошибка	9,13°	4,59°	5,26°	4,22°	17,74°	4,28°	9,52°	10,56°	11,80°

При этом под средней ошибкой понимается значение следующего выражения:

$$\varepsilon = \frac{1}{N} \sum_i |\varphi_i - \varphi'_i|,$$

где φ_i – истинное значение угла поворота лица на i -м изображении,

φ'_i – оценка угла предлагаемым методом,

N – размер выборки (количество изображений).

Из таблицы видно, что наилучшие результаты работы программы получены при использовании канала Cr из цветового пространства YCbCr и канала Saturation из пространства TSL. Для полутоновых изображений наименьшая средняя ошибка наблюдается при фильтрации методом Собеля. Для полученных таким образом выровненных изображений можно применять стандартные детекторы лиц.

Выводы

Метод вычисления угла наклона лица на изображении, предложенный в данной работе, был проверен экспериментально и продемонстрировал приемлемые результаты. Определены наилучшие цветовые представления изображения и метод фильтрации, при которых предлагаемый алгоритм достаточно устойчив. Погрешность, с которой работает данный метод, позволяет эффективно применять его в качестве метода предобработки при детекции и распознавании лиц.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Hjelmas, E. Face detection: A survey* / E. Hjelmas, B. Low // *Computer Vision and Image Understanding*. 2001. – P. 236-274.
2. *Viola, P. Rapid object detection using a boosted cascade of simple features* / P. Viola, M. Jones // *Computer Vision and Pattern Recognition Conference*.
3. *Bradski G. The OpenCV Library* // *Dr Dobb's Journal of Software Tools*. – 2000.
4. *Rowley, H. Rotation Invariant Neural Network-Based Face Detection* / H. Rowley, S. Baluja, T. Kanade // *CMU CS Technical Report*. 1997.
5. *Крамер Г. Математические методы статистики: Пер. с англ.* – М.: Мир, 1975. – 648 с.
6. *Кремер Н.Ш. Теория вероятностей и математическая статистика: учеб. для вузов.* – М.: Юнити-Дана, 2003. – 543 с.
7. *Bradski G. Computer Vision Face Tracking as a Component of a Perceptual User Interface* // *In Proc. of the IEEE Workshop on Applications of Comp. Vision*. 1998. – P. 214-219.
8. *Vezhnevets, V.A survey on pixel-based skin color detection techniques* / V. Vezhnevets, V. Sazonov, A. Andreeva // *Proc. Graphicon 2003*. – September, 2003. – P. 85-92.
9. *Губарев В.В. Применение методов сегментации изображения по цвету кожи в алгоритмах детекции лиц* // *Молодежь и современные информационные технологии: тезисы докл. Всерос. конф. (Томск, 26-28 февр., 2008 г.)*. – Томск, 2008. – С. 231-232.

10. *Tomaz, F.* Improved automatic skin detection in color images / *F. Tomaz, T. Candeias, H. Shahbazkia* // Proc. VIIth Digital Image Computing: Techniques and Applications. 2003. – P. 419-427.
11. *Zuo F.* Embedded Face Recognition Using Cascaded Structures // Technische Universiteit Eindhoven. 2006. – P. 213.
12. *Gonzalez, R.* Digital Image Processing / *R. Gonzalez, R. Woods* // Prentice Hall. 2008. – 954 p.
13. *Canny J.A* Computational Approach To Edge Detection // IEEE Trans. Pattern Analysis and Machine Intelligence. – 1986. Vol.8. – P. 679-714.
14. *Vezhnevets V.* Face and Facial Feature Tracking for Natural Human-Computer Interface. – Режим доступа: <http://graphics.cs.msu.su/en/publications/text/gc2002vvezh.ps.gz> (дата обращения 08.02.2009).
15. *Martinez, A.M.* The AR face database / *A.M. Martinez, R. Benavente* // CVC Technical Report №24. – 1998.

Антипова Агата Юрьевна

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: agatha.antipova@gmail.com.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 8(8634)371-606.

Кафедра высшей математики; аспирантка.

Antipova Agatha Yurievna

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: agatha.antipova@gmail.com.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 8(8634)371-606.

The Department of Higher Mathematics; post-graduate student.

Губарев Владислав Вячеславович

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: vlad.gubarev@gmail.com.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 8(8634)371-606.

Лаборатория математических методов искусственного интеллекта; инженер.

Gubarev Vladislav Vyacheslavovich

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: vlad.gubarev@gmail.com.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 8(8634)371-606.

Laboratory of mathematical methods of artificial intelligence; engineer.