

УДК 004.931

**АЛГОРИТМЫ ВЫДЕЛЕНИЯ ЛИЦ НА СТАТИЧЕСКИХ RGB ИЗОБРАЖЕНИЯХ И В ВИДЕОПОТОКЕ**

А.А. Друки

Томский политехнический университет

E-mail: [druki2008@yandex.ru](mailto:druki2008@yandex.ru)

*Разработаны алгоритмы выделения лиц на статических изображениях и в видеопотоке: алгоритм обнаружения лица с помощью цветовой сегментации, алгоритм выделения лиц на статических RGB изображениях с помощью деформируемых эллиптических моделей, метод статических моментов для выделения лиц в видеопотоке.*

**Ключевые слова:**

*Обработка изображений, выделение лиц, метод деформируемых эталонов, метод статических моментов.*

**Key words:**

*Image recognition, selection of person, method of deformable standards, method of static moments.*

**Введение**

Многие направления науки, техники и производства в значительной степени ориентируются на развитие систем, в которых информация носит характер поля (изображения). При обработке такой информации возникает ряд сложных научных, технических и технологических проблем. Одной из самых сложных задач является обработка и распознавание изображений. О важности этой проблемы говорит тот факт, что исследования по распознаванию образов, анализу изображений и речи включены в перечень приоритетных направлений развития науки и техники и критических технологий федерального уровня [1].

Распознавание изображений находит широкое применение в различных приложениях – это может быть контроль топологии печатных плат, текстуры ткани, контроль доступа к информации по идентификации личности (биометрическая идентификация), доступ к объектам ограниченного доступа, оперативный поиск в картотеке изображений, дактилоскопия и др. [2].

Широкое распространение получают биометрические системы идентификации человека. Традиционные системы идентификации требуют знания пароля, наличия ключа, идентификационной карточки, либо иного идентифицирующего предмета, который можно забыть или потерять. В отличие от них биометрические системы основываются на уникальных биологических характеристиках человека, которые трудно подделать и которые однозначно определяют конкретного человека. К таким характеристикам относятся отпечатки пальцев, форма ладони, узор радужной оболочки, изображение сетчатки глаза, индивидуальные характеристики лица [3].

Лицо человека является важным источником информации при общении между людьми. Выражение лица, мимика, артикуляция при разговоре, движения головой являются удобным, естественным и, что важно, необременительным способом передачи информации. Неспособность компьютера воспринять и распознать столь естественные для человека способы общения затрудняет передачу и восприятие информации при работе с ПЭВМ.

Для того, чтобы научить компьютер распознавать объекты и реагировать на их движения, требуются устойчивые алгоритмы анализа и распознавания изображения лица человека [4].

Цель работы является разработка алгоритмов обнаружения и выделения лица на статических и динамических изображениях низкого разрешения (порядка 320×240 пикселей) при наличии высокого уровня шума, характерного для бытовых фото- и видеокамер.

**Разработка алгоритма обнаружения лица с помощью цветовой сегментации**

Существующие алгоритмы обнаружения лица можно разбить на две категории: методы эмпирического распознавания и методы моделирования изображения лица.

Среди эмпирических методов обнаружения лица получили широкое распространение методы, использующие цветовую сегментацию областей кожи [5]. Причина этой популярности состоит в том, что использование цвета кожи как признака для обнаружения лица сочетает в себе несколько важных преимуществ:

- малая вычислительная сложность и, как следствие, высокая скорость обработки;
  - устойчивость к изменению ориентации и масштаба лица;
  - устойчивость к изменению освещения;
  - устойчивость к изменению выражения лица и частичного перекрытия лица другим объектом сцены.
- К недостаткам этого метода относятся:
- потенциальная возможность ложного обнаружения предметов с цветом, близким к цвету кожи;
  - зависимость результата от цветового баланса камеры и цвета освещения [6, 7].

Вышеперечисленные достоинства обнаружения лиц по цвету делают использование характерного цвета кожи как признака для распознавания весьма привлекательным, даже несмотря на имеющиеся недостатки.

Обнаружение лиц по цвету кожи производится в два этапа:

1. Выделение пикселей, близких по цвету к коже.
2. Поиск лиц на изображении с использованием априорного знания о структуре человеческого лица.

Для выделения пикселей с цветом кожи используется модель распределения оттенка кожи в цветовом пространстве. На основании сравнения существующих методов цветовой сегментации кожи оптимальным представляется метод моделирования распределения с помощью Байесовской карты вероятностей.

Этот метод заключается в вычислении отношения

$$\frac{P(\text{skin} / c)}{P(-\text{skin} / c)}$$

(соотношение вероятностей принадлежности цвета с координатами  $c$  к оттенку «кожи» или «не-кожи») для каждой точки дискретизированного цветового пространства по формуле

$$P_{out} = 2 \sum_{r \in \Omega_2} \frac{P_{skin}(x)}{S},$$

и последующего сравнения  $\frac{P(\text{skin} / c)}{P(-\text{skin} / c)}$  с некоторым выбранным порогом  $\theta$ .  $P(c/\text{skin})$  и  $P(c/-\text{skin})$  напрямую высчитываются из набора тренировочных изображений, с областями кожи, сегментированными вручную. Априорные вероятности  $P(\text{skin})$  и  $P(-\text{skin})$  также можно вычислить, исходя из тренировочного набора. Однако, как легко показать, их можно принять равными 0,5, поскольку изменение априорных вероятностей влияет лишь на величину порога, используемого для обнаружения кожи, не изменяя соотношения процента ложных обнаружений и ложных промахов. Данная модель обладает высокими показателями вероятности успешного распознавания, малой вероятностью ложного обнаружения, и является инвариантной по отношению к выбору цветового пространства, в котором проводится сегментация.

Результатом цветовой сегментации кожи является информация о степени близости цвета каждого пикселя к цвету кожи. На основе этой информации формируется полутоновое изображение, где интенсивности пикселей установлены равными характеристике близости пикселя к цвету кожи.

Результатом цветовой сегментации кожи является информация о степени близости цвета каждого пикселя к цвету кожи. На основе этой информации формируется полутоновое изображение, где интенсивности пикселей установлены равными характеристике близости пикселя к цвету кожи.

#### Разработка алгоритма выделения лиц с помощью деформируемых эллиптических моделей

К сожалению, цветовая информация не может дать достаточно данных для устойчивого обнаружения лица. Неизбежные ошибки цветовой сегментации и области, близкие по оттенку к коже, не являющиеся лицами, создают помехи для распознавания [8]. Поэтому для повышения устойчивости обнаружения лиц был разработан дополнительный метод. Метод заключается в выделении эллиптических областей максимального размера, внутри которых плотность количества пикселей цвета кожи максимальна, в то время как на границе количество пикселей минимально.

На рис. 1 представлен пример желаемого выделения эллиптической области из неидеальных исходных данных: 1) выделенная эллиптическая область; 2) область цвета кожи

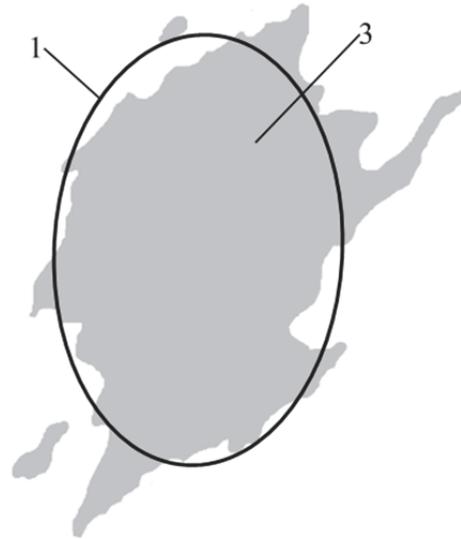


Рис. 1. Пример выделения эллиптической области из неидеальных исходных данных: 1) выделенная эллиптическая область; 2) область цвета кожи

Представляется естественным сформулировать эту проблему, как задачу оптимизации в пространстве параметров эллипса  $(x, y, \theta, a, b)$ . Чтобы избежать трудностей, связанных с поиском явного задания целевой функции и разработки метода оптимизации, не подверженного проблеме локальных экстремумов, было разработано два алгоритма, использующих последовательные приближения с помощью деформируемой эллиптической модели лица, жестко фиксирующие возможные формы областей лиц.

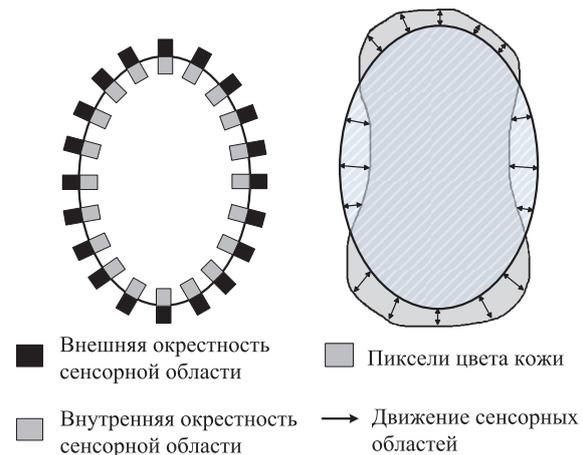


Рис. 2. Деформируемая эллиптическая модель для выделения области лица

Вдоль границы модели расположены сканируемые области, в которых происходит оценка плотности пикселей, близких по цвету к коже. Модель инициализируется вблизи ожидаемого положения лица на изображении (например, в центре масс связанной компоненты пикселей кожи). Деформация

модели производится по шагам. Для каждой сканируемой области производится расчёт плотности пикселей во внутренней и внешней окрестности границы модели. В зависимости от полученных плотностей сканируемая область перемещается перпендикулярно границе модели.

$$P_{in} = 2 \sum_{r \in \Omega 1} \frac{P_{skin}(x)}{S},$$

$$P_{out} = 2 \sum_{r \in \Omega 2} \frac{P_{skin}(x)}{S},$$

где  $\Omega 1$  и  $\Omega 2$  – внутренняя и внешняя части сканирующей области;  $S$  – площадь сканирующей области;  $P_{skin}(x)$  – вероятность принадлежности цвета пикселя  $x$  к цвету кожи. Исходя из рассчитанных величин, вычисляется вектор перемещения сканирующей области:

$$v_i = \begin{cases} -k_{in} \vec{n}_i, & \text{если } P_{in}^i < T_1, \\ k_{out} \vec{n}_i, & \text{если } P_{in}^i \geq T_1 \text{ и } P_{out} < T_2, \\ 0, & \text{иначе.} \end{cases}$$

где  $i$  – номер сканирующей области;  $\vec{n}_i$  – единичная нормаль к границе модели в точке центра сканирующей области, направленная от центра модели;  $P_{in}^i$  и  $P_{out}^i$  – плотности пикселей кожи во внутренней и внешней частях сканирующей области;  $T_1$  и  $T_2$  – пороговые величины плотности пикселей. Коэффициенты  $K_{in}$  и  $K_{out}$  ( $\geq 0$ ) управляют скоростью и характером перемещения сканирующих областей.

Полный цикл обнаружения потенциального региона лица модели задается следующим алгоритмом:

- 1) инициализация модели;
- 2) количество итераций  $k=0$ ;
- 3) пока  $k < n$ ;
- 4) для каждой сканирующей области  $i=1, n$ ;
- 5) подсчитать  $P_{in}^i, P_{out}^i$ ;
- 6) если  $P_{in}^i < T_{in}$  перемещение сканирующей области:

$$\mu_{ij} = \frac{\sum_{(x,y) \in S} (x - \mu_x)^i (y - \mu_y)^j I(x,y)}{\sum_{(x,y) \in SI} I(x,y)}, \text{ иначе:}$$

- 7) если  $P_{out}^i \geq T_{out}$  перемещение сканирующей области:  $\vec{v}_i = -\vec{n}_i v_i$ , иначе:
- 8)  $\vec{v}_i = \theta$ ;
- 9) пересчитать координаты центров сканирующих областей  $x_i^{k+1} = x_i^k + v_i$ ;
- 10) аппроксимировать центры сканирующих областей  $x_i^{k+1}, i=1, n$  эллипсом;
- 11) новые параметры формы модели = параметры полученного эллипса;
- 12)  $k=k+1$ , перейти на 2.

Для приближения центров сканирующих областей эллипсом был использован метод наименьших квадратов, обладающий высоким быстродействием и точностью.

Поведение модели можно варьировать и настраивать, модифицируя следующие параметры:

- количество, расположение и размер сканируемых областей;
- пороговые величины  $T_{in}$  и  $T_{out}$ ;
- коэффициенты длины векторов перемещения сканирующих областей  $k_{in}$  и  $k_{out}$ ;
- дополнительные ограничения на форму, размер, ориентацию и положение эллипса.

Путем модификации перечисленных параметров можно добиться тонкой настройки функционирования модели.

#### Модификация алгоритма

Для увеличения устойчивости метода в случае присутствия на изображении регионов кожных пикселей, значительно отличающихся из-за шума по форме от эллиптических, введены несколько усовершенствований базового метода:

- малое случайное перемещение сканирующих областей вдоль границы модели на каждом шаге;
- предобработка изображения кожи – заполнение дефектов в связанных компонентах кожи заданного меньшего размера, чем ожидаемый размер лица.

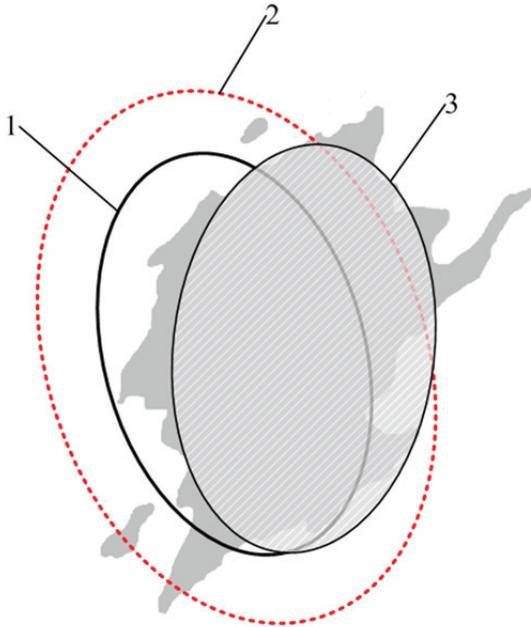
Для того, чтобы увеличить влияние сканирующих областей, находящихся на границе области, введена третья модификация метода. Пересчет параметров эллипса производится методом наименьших квадратов по набору координат центров сканирующих областей  $x_i^{k+1}, i=1, n$ . Особенность метода наименьших квадратов заключается в том, что ошибочное перемещение одной сканирующей области на большое расстояние может значительно сказаться на форме модели и привести к смещению модели с границы области кожи. Чтобы избежать такой ситуации и увеличить влияние стационарных сканирующих областей (в предположении, что они остановились на границе интересующей нас области), их центры участвуют несколько раз при пересчете параметров эллипса. Координаты центров стационарных сканирующих областей  $x_i$  дублируются  $l$  раз в массиве, передаваемом процедуре аппроксимации центром сканирующих областей эллипсом. Эксперименты показали, что  $l=5$  дает хороший компромисс между устойчивостью к несвязностям внутри областей и точностью выделения границ регионов.

#### Метод статических моментов

##### для отслеживания перемещения лица

Задача отслеживания перемещения лица в видеопотоке обладает определенными особенностями по сравнению с обнаружением лица на статическом изображении. Во-первых, мы можем отталкиваться от положения лица на предыдущем кадре, во-вторых мы можем рассчитывать на то, что изменения в размере и, главное, пропорциях лица между кадрами будут невелики. С другой стороны, требуется значительно более высокое быстродействие и сходимость метода. Также, частым отличием при обработке видеопотока является заметно меньшее разрешение и худшее качество изображения по сравнению со статикой [9, 10].

Учитывая перечисленные особенности, был разработан метод отслеживания перемещения лица, также основанный на методе последовательных приближений и использующий жесткое ограничение на возможную форму выделяемых областей, но использующий всю информацию внутри области, а не только на ее границе.



**Рис. 3.** Метод статистических моментов для выделения области лица: 1) текущее положение модели; 2) новое положение модели; 3) область для вычисления моментов

Метод основан на вычислении статистических моментов множества пикселей цвета кожи, попадающих в текущую обрабатываемую область:

$$\mu_x = \frac{\sum_{(x,y) \in SyI(x,y)} (x,y)}{\sum_{(x,y) \in SI(x,y)},}$$

$$\mu_y = \frac{\sum_{(x,y) \in SxI(x,y)} (x,y)}{\sum_{(x,y) \in SI(x,y)},}$$

$$\mu_{ij} = \frac{\sum_{(x,y) \in S(x-\mu_x)^i (y-\mu_y)^j I(x,y)} (x,y)}{\sum_{(x,y) \in SI(x,y)},}$$

где  $I(x,y)$  – вероятность принадлежности пикселя  $(x,y)$  к цвету кожи.

Инициализация модели производится либо положением лица на предыдущем кадре (в случае обработки видео), либо в центре масс связной компоненты пикселей кожи (как в случае с предыдущим методом). На каждом шаге деформации производится пересчет новых параметров эллипса. Пусть  $\vec{v}_k$  – вектор параметров эллипса на  $n$  шаге  $(x,y,\theta,a,b)$ . Здесь  $\theta$  – угол наклона главной оси эллипса;  $a, b$  – длины осей эллипса.

Полный цикл обнаружения области лица указанным методом описан в следующем алгоритме:

0. Инициализация модели.
1. Количество итераций  $k=0$ .

2. Пока  $k < n$ .
3. Рассмотреть эллипс увеличенного размера  $(x_k, y_k, \theta_k, \alpha a_k, \alpha b_k)$ ,  $\alpha > 1$ .
4. Подсчитать центр масс  $(\mu_x, \mu_y)$  и центральные моменты второго порядка  $\mu_{20}, \mu_{02}, \mu_{11}$  и множества пикселей, попавших в рассматриваемую область.
5. Подсчитать новые параметры эллипса по полученным статистикам

$$v_{k+1} = \left( \mu_x, \mu_y, a \tan \left( \frac{\mu_{11}}{\mu_{20} - \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \right), a_{k+1}, b_{k+1} \right).$$

6.  $k=k+1$ , перейти на 2.  
Новые длины осей эллипса  $a_{k+1}, b_{k+1}$  вычисляются как

$$a_{k+1} = \frac{1}{1+\beta} (a_k + d_k a_k \beta),$$

$$b_{k+1} = \frac{1}{1+\beta} (b_k + d_k b_k \beta),$$

где  $d_k$  – среднее расстояние Махаланобиса от центра эллипса лица до точек цвета кожи, находящихся внутри эллипса увеличенного размера;  $\beta$  – коэффициент, регулирующий скорость адаптации (по результатам экспериментов принятый равным 0,19).

Обычно двух-трех итераций пересчета параметров эллипса достаточно для вычисления нового положения лица, даже при достаточно больших перемещениях между кадрами.

Сравнение предложенных методов с существующими показало, что предлагаемые методы обладают заметно более высокой устойчивостью к зашумленным исходным данным и ошибкам цветовой сегментации, сохраняя простоту программной реализации и высокую скорость работы, позволяющую использовать разработанные алгоритмы в режимах реального времени.

Простые схемы алгоритмов гарантируют возможность получения быстрой программной реализации. Несмотря на простоту, они обладают высокой устойчивостью к шумным и неточным исходным картам вероятности кожи. Предложенные методы также обладают важным свойством устойчивости к выбору начального приближения, что выгодно отличает их от многих методов распознавания, основанных на деформируемых моделях. Описанная схема обладает большой гибкостью и позволяет с легкостью встраивать дополнительные условия и ограничения на возможную форму выделяемых областей (пропорции, положение, ориентации эллипса).

Учитывая устойчивость алгоритмов к точности начального приближения на основе описанной схемы, строится следующий порядок обнаружения лиц на изображениях:

- 1) провести сегментацию кожи, создать карту вероятности кожи;

- 2) выделить связные компоненты пикселей с высокой кожной вероятностью или подсчитать интегральные проекции карты кожи;
- 3) инициализировать модель в центрах масс связанных компонент или пиках интегральных проекций;
- 4) произвести обнаружение эллиптических областей пикселей кожи;
- 5) оценить их размеры, пропорции, форму и зафиксировать их как потенциальные лица, если проверка дает положительный результат;
- 6) убрать обнаруженные области из карты вероятностей;
- 7) повторять процедуру с шага 2, пока обнаруживаются новые области.

#### СПИСОК ЛИТЕРАТУРЫ

1. Визильтер Ю.В., Желтов С.Ю., Ососков М.В. Система распознавания и визуализации характерных черт человеческого лица в реальном времени на персональной ЭВМ с использованием web-камеры // Графикон: Труды конф. – Н. Новгород, 2002. – С. 251–254.
2. Rabiner L.R. A Tutorial on hidden Markov models and selected applications in speech recognition // IEEE scientific works compilation. – 1989. – V. 77. – № 2. – P. 257–285.
3. Jordao L., Perrone M., Costeira L., Santos L. Active face and feature tracking // Proc. Inter. Conf. on lineage analysis and processing. – Lisbon, 1999. – V. 9. – P. 572–576.
4. Albiol A., Torres L., Delp E.J. Optimum color spaces for skin detection // Proc. Inter. Conf. on image processing. – Valencia, 2001. – V. 1. – P. 122–124.
5. Baskan S., Bulut M.A., Atalay V. Projection based method for segmentation of human face and evaluation // Pattern Recognition Letters. – 2002. – V. 20. – P. 1623–1629.

#### Выводы

Предложены алгоритмы выделения лиц на статических RGB изображениях и в видеопотоке, работающие в реальном времени и удовлетворяющие следующим критериям:

- работа с изображениями низкого разрешения (порядка 320×240 пикселей) и высокого уровня шума, характерного для бытовых фото- и видеокamer;
- устойчивость к ошибкам и неточностям цветовой сегментации;
- быстрая работа методов;
- возможность гибкой настройки и задания ограничений на форму эллиптических моделей.

6. Brand J., Mason J. A comparative assessment of three approaches to pixellevel human skin – detection // Proc. Inter. Conf. on pattern recognition. – Swansea, 2002. – V. 1. – P. 1056–1059.
7. Craw I., Ellis H., Ashman J. Automatic extraction of face features // Pattern recognition letters. – 1987. – V. 5. – P. 183–187.
8. Edwards G., Taylor C., Cootes T. Interpreting face images using active appearance models // Proc. Conf. on automatic face and gesture recognition. – Manchester, 1998. – V. 7. – P. 300–305.
9. Hansen D.W., Hansen J.P., Nielsen M. Eye typing using Markov and active appearance models // Workshop on applications of computer vision. – 2002. – V. 12. – P. 132–136.
10. Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks // Transactions on pattern analysis and machine intelligence. – 2001. – V. 34. – P. 42–53.

*Поступила 10.12.2011 г.*