

ПРИМЕНЕНИЕ ДЕФОРМИРУЕМЫХ ЭЛЛИПТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ВЫДЕЛЕНИЯ ЛИЦ НА ИЗОБРАЖЕНИЯХ И В ВИДЕОПОТОКЕ

А. А. Друки

Институт кибернетики Национального исследовательского
Томского политехнического университета, 634034, Томск, Россия

УДК 004.931

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.

The face detection algorithms on static images and video stream are developed: the face detection algorithm with the help of a color segmentation, the face detection algorithm on static RGB images with the help of deformable elliptic model, the method of static moments for the selection of persons in the video stream.

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 ориентированы на развитие систем, в которых информация имеет характер поля (изображения). При обработке такой информации возникает ряд сложных научных, технических и технологических проблем. Одной из наиболее сложных задач является обработка и распознавание изображений [1].

Распознавание изображений находит применение в различных приложениях: 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. [1].

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 этих систем биометрические системы основываются на уникальных биологических характеристиках человека, которые трудно подделать и которые однозначно определяют конкретного человека. К числу таких характеристик относятся отпечатки пальцев, форма ладони, узор радужной оболочки, изображение сетчатки глаза, индивидуальные характеристики лица [2].

Лицо человека является важным источником информации при общении между людьми. Выражение лица, мимика, артикуляция при разговоре, движения головой являются удобным, естественным и, что важно, простым способом передачи информации. Неспособность

компьютера воспринять и распознать столь естественные для человека способы общения затрудняет передачу и восприятие информации при работе с ПЭВМ [3].

Для того чтобы научить компьютер распознавать объекты и реагировать на их движения, требуются устойчивые алгоритмы анализа и распознавания изображения лица человека [3].

Целью 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 разделить на две категории: методы эмпирического распознавания и методы моделирования изображения лица [4].

Среди эмпирических методов обнаружения лица широкое распространение получили методы, использующие цветовую сегментацию областей кожи. Это обусловлено тем, что использование цвета кожи как признака для обнаружения лица имеет ряд 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 числу недостатков данного метода можно отнести:

- потенциальную возможность ложного обнаружения предметов с цветом, близким к цвету кожи;
- зависимость от цветового баланса камеры и цвета освещения [4].

Перечисленные выше преимущества метода обнаружения лиц по цвету позволяют использовать характерный цвет кожи в качестве признака для распознавания (несмотря на имеющиеся недостатки).

Процесс обнаружения лиц по цвету кожи обычно включает 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. Результаты сравнения существующих методов цветовой сегментации кожи показывают, что оптимальным является метод моделирования распределения с помощью байесовской карты вероятностей [5, 6]. Суть этого метода заключается в вычислении отношения $P(\text{skin}/c)/P(-\text{skin}/c)$ (соотношения 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 сравнении значения $P(\text{skin}/c)/P(-\text{skin}/c)$ с некоторым выбранным порогом θ . Значения $P(c/\text{skin})$ и $P(c/-\text{skin})$ непосредственно вычисляются 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 эти вероятности можно принять равными 0,5, поскольку изменение априорных вероятностей влияет

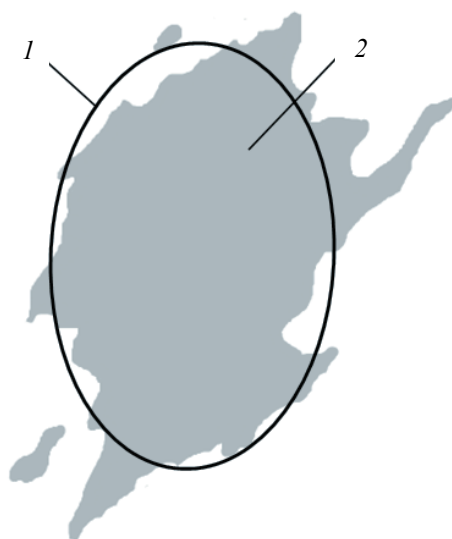


Рис. 1. Пример выделения эллиптической области из неидеальных исходных данных:
1 — выделенная эллиптическая область; 2 — область пикселей цвета кожи

лишь на величину порога, используемого для обнаружения кожи, и не приводит к изменению соотношения 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.

Разработка алгоритма выделения лиц с помощью деформируемых эллиптических моделей. Следует отметить, что цветовой информации недостаточно для устойчивого обнаружения лица. Неизбежные ошибки цветовой сегментации и наличие областей, близких по оттенку к коже, но не являющихся лицами, создают помехи для распознавания [6]. Для повышения устойчивости обнаружения лиц разработан дополнительный метод, заключающийся в выделении эллиптических областей максимального размера, внутри которых плотность пикселей цвета кожи максимальна, в то время как на границе она минимальна. На рис. 1 представлен пример желаемого выделения эллиптической области из неидеальных исходных данных.

Представляется естественным сформулировать рассматриваемую задачу как задачу оптимизации в пространстве параметров эллипса (x, y, θ, a, b) . Чтобы избежать трудностей, возникающих при поиске явного задания целевой функции и разработке метода оптимизации, не подверженного проблеме локальных экстремумов, созданы два алгоритма, использующие последовательные приближения с помощью деформируемой эллиптической модели лица, жестко фиксирующие возможные формы областей лиц. На рис. 2 представлена деформируемая эллиптическая модель для выделения области лица.

Вдоль границы модели расположены сканируемые области, в которых выполняется оценка плотности пикселей, близких по цвету к коже [7, 8]. Модель инициализируется вблизи

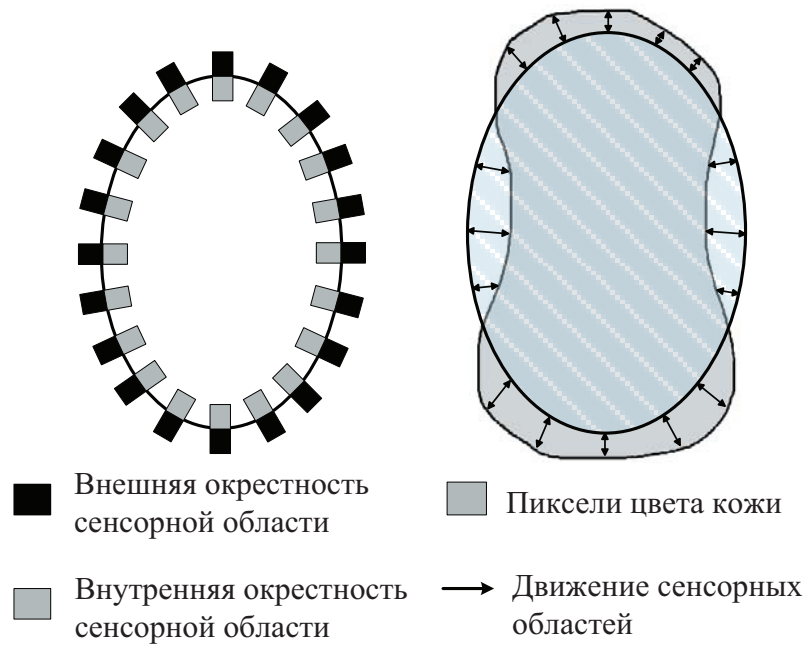


Рис. 2. Деформируемая эллиптическая модель для выделения области лица

ожидаемого положения лица на изображении (например, в центре масс связной компоненты пикселей кожи). Деформация модели производится по шагам. Для каждой сканируемой области проводится расчет плотности пикселей кожи во внутренней и внешней окрестностях границы модели. При определенных значениях плотностей сканируемая область перемещается перпендикулярно границе модели. Расчет плотности пикселей кожи во внутренней и внешней окрестностях границы модели проводится по формулам

$$P_{in} = 2 \sum_{r \in \Omega_1} \frac{P_{skin}(x)}{S}, \quad P_{out} = 2 \sum_{r \in \Omega_2} \frac{P_{skin}(x)}{S},$$

где Ω_1, Ω_2 — внутренняя и внешняя части сканируемой области; S — площадь сканируемой области; $P_{skin}(x)$ — вероятность принадлежности цвета пикселя к цвету кожи. С использованием рассчитанных величин вычисляется вектор перемещения сканируемой области

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

Здесь i — номер сканируемой области; \mathbf{n}_i — единичная нормаль к границе модели в центре сканируемой области, направленная от центра модели; P_{in}^i, P_{out}^i — плотности пикселей кожи во внутренней и внешней частях сканируемой области; T_1, T_2 — пороговые величины плотности пикселей. Коэффициенты K_{in} и K_{out} управляют скоростью и характером перемещения сканируемых областей.

Полный цикл обнаружения потенциальной области лица модели задается следующим алгоритмом:

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

2. Пока $k < n$:
3. для каждой сканируемой области $i = \overline{1, n}$:
4. подсчитать P_{in}^i, P_{out}^i ;
5. если $P_{in}^i < T_{in}$, перемещение сканируемой области $\mathbf{v}_i = -\mathbf{n}v_{in}$, иначе
6. если $P_{out}^i \geq T_{out}$, перемещение сканируемой области $\mathbf{v}_i = \mathbf{n}v_{out}$, иначе
7. $\mathbf{v}_i := \theta$;
8. вычислить координаты центров сканируемых областей $x_i^{k+1} = x_i^k + \mathbf{v}_i$;
9. аппроксимировать центры сканируемых областей $x_i^{k+1}, i = \overline{1, n}$ эллипсом;
10. новые параметры формы модели := параметры полученного эллипса;
11. $k = k + 1$, перейти на шаг 2.

Метод основан на вычислении статических моментов множества пикселей цвета кожи, попадающих в текущую сканируемую область:

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

Здесь $I(x, y)$ — вероятность принадлежности пикселя (x, y) к цвету кожи.

Для интерполяции центров сканируемых областей эллипсом использовался метод наименьших квадратов, обладающий высокими 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 сканируемых областей k_{in} и k_{out} ;
- дополнительные ограничения на форму, размер, ориентацию и положение эллипса.

Модифицируя перечисленные параметры, можно добиться тонкой настройки функционирования модели.

Модификация алгоритма. Для увеличения устойчивости метода в случае наличия на изображении областей кожных пикселей, вследствие шума значительно отличающихся по форме от эллиптических, введено несколько усовершенствований базового метода:

- малое случайное перемещение сканируемых областей вдоль границы модели на каждом шаге;
- предварительная обработка изображения кожи — заполнение дефектов в связанных компонентах кожи заведомо меньшего размера по сравнению с ожидаемым размером лица.

Для увеличения влияния сканируемых областей, находящихся на границе области, введена третья модификация метода. Пересчет параметров эллипса проводится методом наименьших квадратов по набору координат центров сканируемых областей $x_i^{k+1}, i = \overline{1, n}$. Особенность метода наименьших квадратов заключается в том, что ошибочное перемещение одной сканируемой области на большое расстояние может значительно повлиять на форму модели и привести к ее смещению с границы области кожи. Для того чтобы избежать такой ситуации и увеличить влияние стационарных сканируемых областей (в предположении, что они остановились на границе рассматриваемой области), их центры используются несколько раз при пересчете параметров эллипса. Координаты центров стационарных сканируемых областей x_i дублируются l раз в массиве, представляющем собой входные данные для процедуры аппроксимации центров сканируемых областей эллипсом. Эксперименты показали, что

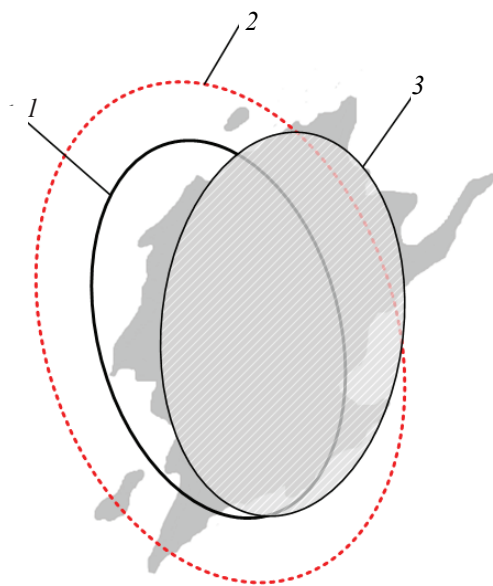


Рис. 3. Метод статистических моментов для выделения области лица:
 1 — текущее положение модели; 2 — новое положение модели;
 3 — область, в которой вычисляются моменты

при значении $l = 5$ достигается 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 имеет ряд особенностей, отличающихся от особенностей задачи обнаружения лица на статическом изображении. Во-первых, в качестве основы можно использовать положение лица на предыдущем кадре, во-вторых, можно предположить, что изменения размера, а главное, пропорций лица между кадрами будут невелики. В то же время требуются значительно более высокие быстродействие и сходимость метода. Еще одной особенностью обработки видеопотока является значительно меньшее разрешение и худшее качество изображения по сравнению со случаем статичного положения объекта [9, 10].

С учетом указанных особенностей предлагается метод отслеживания перемещения лица, также основанный на методе последовательных приближений с жестким ограничением на возможную форму выделяемых областей, но использующий всю информацию внутри области, а не только на ее границе [11].

Метод основан на вычислении статистических моментов множества пикселей цвета кожи, попадающих в текущую обрабатываемую область. Инициализация модели проводится либо изменением положения лица на следующем кадре (в случае обработки видео), либо в центре масс связной компоненты пикселей кожи. На рис. 3 показан пример выделения и отслеживания эллиптической области в видеопотоке. На каждом шаге деформации выполняется пересчет новых параметров эллипса [12]. Пусть \mathbf{v}_k — вектор параметров эллипса на n -м шаге (x, y, θ, a, b) . Здесь θ — 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. вычислить центр масс (μ_x, μ_y) и центральные моменты второго порядка $\mu_{20}, \mu_{02}, \mu_{11}$ и множества пикселей, попавших в рассматриваемую область;

5. вычислить новые параметры эллипса по полученным статистикам

$$\mathbf{v}_{k+1} = \left(\mu_x, \mu_y, a \operatorname{tg} \left(\frac{\mu_{11}}{\mu_{20} - \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \right), a_{k+1}, b_{k+1} \right);$$

6. $k = k + 1$, перейти на шаг 2.

Новые длины осей эллипса a_{k+1}, b_{k+1} вычисляются по формулам

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

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

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

Сравнение предложенных методов с существующими показало, что предлагаемые методы обладают существенно 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 получение быстрой программной реализации. Несмотря на простоту, предложенные методы обладают высокой устойчивостью к шумным и неточным исходным картам вероятности кожи, а также обладают важным свойством устойчивости к выбору начального приближения, что выгодно отличает их от других методов распознавания, основанных на деформируемых моделях. Описанная схема обладает большой гибкостью и позволяет без труда встраивать дополнительные условия и ограничения на возможную форму выделяемых областей (пропорции, положение, ориентации эллипса).

С учетом устойчивости алгоритмов к точности начального приближения на основе описанной схемы строится следующий метод обнаружения лиц на изображениях:

1. Провести сегментацию кожи, создать карту вероятности кожи.
2. Выделить связные компоненты пикселей с высокой кожной вероятностью или вычислить интегральные проекции карты кожи.
3. Инициализировать модель в центрах масс связных 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. Удалить обнаруженные области из карты вероятностей.
7. Повторять процедуру с шага 2, до тех пор пока обнаруживаются новые области.

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 видеокамер;
- устойчивость к ошибкам и неточностям цветовой сегментации;
- быстрая работа методов;
- возможность гибкой настройки и задания ограничений на форму эллиптических моделей.

Список литературы

1. ВАИНИК В. Н., ЧЕРВОНЕНКИС А. Я. Теория распознавания образов // Стохастические проблемы обучения. М.: Наука, 1974. С. 144–157.
2. САМАЛЬ Д. И., СТАРОВОЙТОВ В. В. Обнаружение антропометрических точек лица с помощью фильтров Габора // Цифровая обработка изображений. М.: Наука, 2001. С. 141–150.
3. 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 // Тр. Междунар. конф. по компьютерной графике и машинному зрению “Графикон’2002”, Нижний Новгород, 2002. Н. Новгород: Графикон, 2002. С. 251–254.
4. RABINER L. R. A tutorial on hidden Markov models and selected applications in speech recognition // IEEE Sci. Works Compilat. 1989. V. 77, N 2. P. 257–285.
5. JORDAO L., PERRONE M., COSTEIRA L., SANTOS L. Active face and feature tracking // Proc. of the Intern. conf. on lineage analysis and processing. Lisbon: S. n., 1999. V. 9. P. 572–576.
6. ALBIOL A., TORRES L., DELP E. J. Optimum color spaces for skin detection // Proc. of the IEEE Intern. conf. on image processing, Thessaloniki, Oct. 8–10 2001. IEEE Conf. Publ., 2001. V. 1. P. 122–124.
7. BASKAN S., BULUT M. A., ATALAY V. Projection based method for segmentation of human face and evaluation // Pattern Recognition Lett. 2002. V. 20. P. 1623–1629.
8. BRAND J., MASON J. A comparative assessment of three approaches to pixellevel human skin-detection // Proc. of the 15th Intern. conf. on pattern recognition, Barcelona, 3–8 Sept. 2000. IEEE Computer Soc., 2000. V. 1. P. 1056–1059.
9. CRAW I., ELLIS H., ASHMAN J. Automatic extraction of face features // Pattern Recognition Lett. 1987. V. 5. P. 183–187.
10. EDWARDS G., TAYLOR C., COOTES T. Interpreting face images using active appearance models // Proc. of the 3rd IEEE conf. on automatic face and gesture recognition, Nara (Japan), 14–16 Apr. 1998. S. n., 1998. V. 7. P. 300–305.
11. HANSEN D. W., HANSEN J. P., NIELSEN M. Eye typing using Markov and active appearance models // 6th IEEE Workshop on applications of computer vision, Copenhagen (Denmark), June 2, 2002. S. n., V. 12. P. 132–136.
12. FERAUD R., BERNIER O., VIALLET J., COLLOBERT M. A fast and accurate face detector based on neural networks // Trans. Pattern Anal. Machine Intelligence. 2001. V. 34. P. 42–53.

*Друки Алексей Алексеевич — асп. Института кибернетики Национального
исследовательского Томского политехнического университета;
тел.: (382-2) 22-45-08; e-mail: druki2008@yandex.ru*

Дата поступления — 10.12.11 г.