УДК 681.327.12.001.362

В.В. КРАСНОПРОШИН, Е.В. КОБЛОВ

ЭФФЕКТИВНЫЙ АЛГОРИТМ ПОИСКА ЛИЦЕПОДОБНЫХ ОБЛАСТЕЙ

In this paper a pattern-recognition based approach to solving face detection task is proposed. All main steps are revealed and algorithms for their solution are built. It is shown that these algorithms in the aggregate effectively solve the face detection task. The conditions when the exact solution can be obtained arc determined.

В настоящее время широкое распространение получили биометрические системы идентификации, которые основываются на уникальных биологических характеристиках человека [1]. **Одним из носителей таких характеристик выступает лицо. Поэтому при реализации биометрической технологии важной является задача поиска лиц на изображении.**

На данный момент эта задача далека от разрешения. Применяемые подходы [2-6] носят, как правило, универсальный характер, имеют высокую алгоритмическую сложность (порядка $0(n')$, где n - количество пикселей изображения) и не инвариантны к условиям съемки. Все это делает проблематичным их использование в системах реального времени.

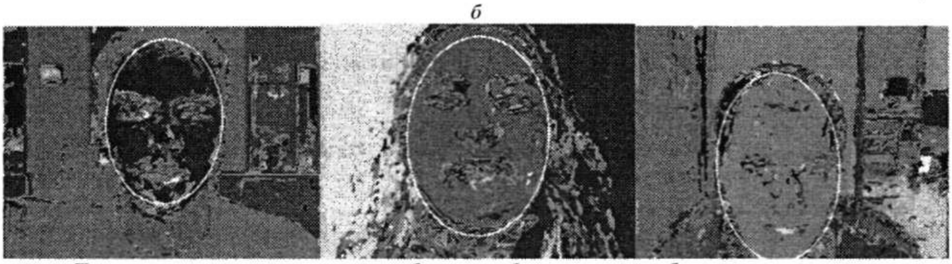
В данной работе предлагается подход, алгоритмическая сложность которого составляет $0(n)$, и определяются условия, при выполнении которых возможно построение точного решения.

Анализ проблемы и постановка задачи. **Одним из путей решения проблемы поиска лиц является ее сведение к задаче распознавания образов [7]. Для этого определяют систему признаков, описывающую множество допустимых объектов, задают его разбиение на классы и строят алгоритм распознавания. После того как алгоритм построен, поиск лица на произвольном изображении решается в три этапа (каждому этапу соответствует своя подзадача):**

- **исходное изображение разбивается на фрагменты;**
- **фрагменты описываются в терминах допустимых объектов;**
- **определяется принадлежность объектов к классам (рисунок).**

а

б



Примеры решения: *а* – исходные изображения, *б* – результаты работы алгоритма поиска

Задача распознавания в этом процессе является определяющей, но эффективность решения в целом зависит от каждой из указанных подзадач.

Для решения задачи поиска лиц (ЗПЛ) в принципе можно использовать любой из алгоритмов распознавания. Возникают естественные вопросы: какой из алгоритмов предпочтителен с точки зрения эффективности (сложности и точности), как на общую эффективность решения влияют остальные задачи и можно ли управлять эффективностью в рамках всего процесса решения?

Анализ существующих алгоритмов распознавания показывает, что сложность решения ЗПЛ определяется универсальностью применяемых схем - алгоритмов распознавания или преобразования Хафа. Это в конечном итоге приводит к необходимости полного перебора всевозможных фрагментов. В обоих случаях сложность составляет $O(n^3)$. Что касается точности решения, то она закладывается либо на этапе построения алгоритмов распознавания, либо на этапе выполнения преобразования Хафа и в дальнейшем не улучшается.

Рассмотрим подход, в котором предпринята попытка понижения сложности и повышения точности решения ЗПЛ. Оставаясь в рамках задачи распознавания образов, построим алгоритмы, которые позволят управлять эффективностью на всех этапах процесса решения.

Задача распознавания. При решении задачи распознавания образов будем использовать следующую гипотезу [8]: «По форме лицо человека похоже на эллипс, у которого соотношение большой и малой осей близко к 75 %». Исходя из этого, допустимым объектом является любой фрагмент изображения, который имеет эллиптическую форму. Тогда его можно описывать в системе признаков, соответствующих параметрам эллипса, т. е. допустимый объект $F=(x_0,y_0,a,b,\theta)$, где (x_0, y_0) - центр эллипса, θ - угол поворота и *a, Ъ* - большая и малая полуоси.

Множество всех допустимых объектов разобьем на два класса: все объекты, у которых $\left|\frac{b}{a}-0,75\right|\le\varepsilon$, где ε - некоторое пороговое значение, и все остальные.

Введем предикаты

$$P_1(F)=\begin{cases}1, & \text{если } |b/a-0,75|\le\varepsilon,\\ 0 & \text{в противном случае;}\end{cases} \qquad P_2(F)=\begin{cases}1, & |b/a-0,75|>\varepsilon,\\ 0 & \text{в противном случае.}\end{cases}$$

Тогда алгоритм распознавания можно записать в виде [8]: $A=r\circ R(R_1,R_2)$, где $R_1=\langle F, P_1\rangle$, $R_2=\langle F, P_2\rangle$ - распознающие операторы, $r=\max(R_1, R_2)$ - решающее правило.

Нетрудно видеть, что сложность предлагаемого алгоритма определяется сложностью вычисления предикатов (и составляет *0(1)),* а точность решения - выбором соответствующего порога ε. Можно также отметить, что алгоритм является инвариантным относительно геометрических преобразований масштабирования и поворота (см. рисунок).

Задача описания. Пусть *U* - произвольный фрагмент изображения I , *E* - растровая развертка аппроксимирующего его эллипса.

Будем считать, что фрагмент *U* имеет *эллиптическую форму,* если существует *E,* для которого выполняется условие $\frac{\|c_E-c_{I_0}\|}{\|c_E\|}<\delta$, где δ - некоторое пороговое значение, c_E, c_U - характеристические функции фрагментов *E и U* соответственно, $\|\ \|$ - произвольная функциональная норма.

Теперь **для** того, чтобы фрагмент изображения перевести в допустимый объект задачи распознавания, необходимо:

- аппроксимировать фрагмент эллипсом;

- проверить аппроксимацию на допустимость (на степень ее соответствия границам фрагмента).

При решении первой из подзадач будем использовать следующие свойства [9]: центры эллипса и его массы совпадают, угол наклона определяется ориентацией, а полуоси равны половине высоты и ширины граничного прямоугольника. Так как лицо имеет эллиптическую форму, то, очевидно, описанные свойства можно использовать для вычисления параметров аппроксимирующего эллипса.

Алгоритм **1**

Шаг 1 *(определение центра эллипса* (x_0, y_0)*).*

Пусть фрагмент $U=\{(i,j)\}_{m\times n}$ задан характеристической функцией

$$c(i,j)=\begin{cases}1,(i,j)\in U,\\ 0,(i,j)\notin U.\end{cases}$$

Центр масс фрагмента (по определению) вычислим с помощью моментов первого порядка:

$$x_0=\frac{\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}kc(k,l)}{\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}c(k,l)},\quad y_0=\frac{\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}lc(k,l)}{\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}c(k,l)}.$$

Шаг 2 *(нахождение угла наклона эллипса* θ*).*

Для вычисления θ используем моменты второго порядка:

$$\alpha=\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}(k-x)^2c(k,l),\quad \beta=2\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}(k-x)(l-y)c(k,l),$$

$$\gamma=\sum_{k=0}^{m-1}\sum_{l=0}^{n-1}(l-y)^2c(k,l).$$

Из соотношения $\operatorname{tg}2\theta=\beta/(\alpha-\gamma)$ [8], $\beta\neq 0$ или $\alpha\neq\gamma$, угол наклона (с точностью до $\pi/2$) вычислим по формуле

$$\theta=\frac{\operatorname{arc\,tg}(\beta/(\alpha-\gamma))}{2}.$$

Шаг 3 *(выделение граничного прямоугольника).*

Определим граничный прямоугольник фрагмента с учетом найденного угла наклона. Сначала повернем исходный фрагмент на угол θ:

$$\begin{cases}x^*=x_0+(x-x_0)\cos\theta-(y-y_0)\sin\theta,\\ y^*=y_0+(x-x_0)\sin\theta+(y-y_0)\cos\theta.\end{cases}$$

Теперь координаты диагональных точек прямоугольника определим по формулам

$$\text{left}=\min_{c_i(x,y)\neq 0}(x^*),\quad \text{top}=\min_{c_i(x,y)\neq 0}(y^*),$$
$$\text{right}=\max_{c_i(x,y)\neq 0}(x^*),\quad \text{bottom}=\max_{c_i(x,y)\neq 0}(y^*).$$

Шаг 4 *(уточнение угла наклона).*

С учетом определения полуосей эллипса уточним угол его наклона:

$$\theta=\begin{cases}\theta,\ \text{если } 2a=\text{right}-\text{left},\\ \theta+\dfrac{\pi}{2}\ \ \text{в противном случае}.\end{cases}$$

Таким образом, алгоритм строит эллипс, аппроксимирующий любой фрагмент изображения (см. рис. *б).* Последние не всегда имеют эллиптическую форму, поэтому аппроксимацию необходимо проверить на допустимость. Для решения этой подзадачи введем некоторые характеристики.

Пусть U_j - некоторый фрагмент и $V=(x_0, y_0, \theta, a, b)$ - аппроксимирующий его эллипс.

Обозначим через количество пикселей фрагмента, лежащих внутри и вне эллипса соответственно, а через C - общее количество пикселей эллипса.

Введем правило:

$$P_{\text{in}}/C>T_1 \text{ и } P_{\text{out}}/C<T_2, \tag{1}$$

где T_1 и T_2 - некоторые пороговые величины.

Покажем, что применение правила (1) эквивалентно вычислению $\dfrac{\|c_E-c_{I_0}\|}{\|c_E\|}<\delta$ по городской норме, т. е.

$$\frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)-c_{I_0}(i,j)\right|}{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)\right|}\le\delta,\ \text{где}\ \delta=1-T_1+T_2.$$

Действительно, по определению характеристической функции имеем $\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)\right| = C.$ Очевидно, что $\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)-c_{I_0}(i,j)\right| = P_{\text{out}}+(C-P_{\text{in}}).$

Поэтому $\dfrac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)-c_{I_0}(i,j)\right|}{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|c_E(i,j)\right|} = \dfrac{P_{\text{out}}+(C-P_{\text{in}})}{C} = 1-\dfrac{P_{\text{in}}}{C}+\dfrac{P_{\text{out}}}{C} \le 1-T_1+T_2$.

Правило (1) позволяет достаточно просто проверять построенные в результате аппроксимации объекты на допустимость, а заданием порогов T_1 и T_2 - варьировать уровень такой допустимости.

Таким образом, применение алгоритма в совокупности с правилом (1) дает возможность строить допустимые объекты для задачи распознавания. Нетрудно показать, что сложность такого построения линейна. Действительно, за один проход по изображению можно определить центры эллипсов и углы их поворота, за второй - найти граничный прямоугольник и полуоси эллипсов, а за третий - подсчитать величины P_{in}, P_{out} и С и проверить фрагменты на допустимость. Поскольку на обработку пикселя затрачивается время 0(1), а каждый из них обрабатывается не более трех раз, то сложность построения допустимого объекта равна *0(N),* где N - количество пикселей изображения.

Задача фрагментации. Для решения задачи распознавания необходимо уметь специальным образом переходить от исходного изображения к множеству его фрагментов, по которым затем строятся допустимые объекты. Иными словами, необходимо так решить задачу фрагментации, чтобы области изображения, соответствующие лицам, выделялись в отдельные фрагменты.

Пусть $I : P^2_{m\times n} \to C$ - растровое изображение над цветовым пространством С, здесь $P^2_{m\times n}$ - дискретная сетка размерности $m\times n$.

Пару пикселей $(i_1,j_1),(i_2,j_2)\in P^2_{m\times n}$ будем считать *соседними,* если для них выполняется соотношение

$$\max\ (|\, i_1-i_2\,|, |\, j_1-j_2\,|) = 1.$$

Два соседних пикселя (i_1,j_1) и (i_2,j_2) являются *близкими,* если

$$\left\|I(i_1,j_1)-I(i_2,j_2)\right\| < \nu,$$

где $\|\ \|$ - некоторая мера пространства С, ν - пороговая величина, определяющая детализацию изображения.

Основная идея предлагаемого алгоритма связана со специальной разметкой изображения, в результате которой пиксели одного фрагмента помечаются одинаковым идентификатором.

Алгоритм* 2 *(фрагментации изображения)

Шаг 1. Пусть L - некоторая очередь, $U=\{u_{ij}\}_{m\times n}$ - матрица размера $m\times n$, q - текущий идентификатор. Полагаем $L=\{\varnothing\}$, $U=\{0\}_{m\times n}$, q=1.

Шаг 2. Просматриваем изображение I слева направо, сверху вниз $i=\overline{0,m-1}$, $j=\overline{0,n-1}$ и ищем неразмеченные пиксели.

Если таковых нет, то алгоритм заканчивает работу. В противном случае находим первый из еще неразмеченных пикселей (i_0,j_0), помечаем его текущим идентификатором ($u_{i_0j_0}=q$) и добавляем в очередь ($L=\{(i_0,j_0)\}$).

Увеличиваем текущий идентификатор *(q=q+1)* и переходим к следующему шагу.

Шаг 3. Пока $L\neq\{\varnothing\}$, извлекаем из нее очередной пиксель (i',j') и сравниваем его с соседними еще неразмеченными пикселями.

Пусть (i_k, j_k) - очередной k-*й* соседний пиксель *(k = 1,c* , где *c* - количество соседей). Если $\|I(i',j') - I(i_k,j_k)\| < \nu$, то помечаем его тем же, что и (i',j'), идентификатором ($u_{i'j'} = u_{i_k j_k}$) и добавляем в очередь *L*. Если $L = \{\varnothing\}$, переходим на шаг 2.

В результате работы алгоритма будет построена матрица идентификаторов ГУ, которая задает разбиение изображения на множество фрагментов.

Определим условия, при которых область, соответствующая лицу, будет выделена в отдельный фрагмент.

Утверждение. Произвольная область изображения A выделяется алгоритмом в отдельный фрагмент, если для любой пары соседних пикселей $(i_1, j_1),(i_2, j_2)$ выполняются условия:

1) $\|I(i_1,j_1) - I(i_2,j_2)\| < \nu$, если оба они принадлежат/!;

2) $\|I(i_1,j_1) - I(i_2,j_2)\| \geq \nu$, если один из них не принадлежит ***A***.

Покажем, что это действительно так. Предположим противное, что условия 1 и 2 выполняются и выделен фрагмент $A' \neq A$ такой, что $A' \cap A \neq \varnothing$. Это означает, что существует пиксель $p_0 = (i_0, j_0)$ такой, что $p_0 \in A'/A$ или $p_0 \in A/A'$.

Пусть $p_0 \in A'/A$ и является соседним с $A \cap A'$. Так как $p_0 \notin A$, то согласно условию 2 для $\forall p_i \in A \cap A' \subseteq A$, соседнего с p_0, выполняется неравенство $\|I(p_0) - I(p_i\| > \nu$. Это также означает, что p_0 не может входить в один фрагмент с пикселями p_i . Но $p_i \in A'$, и тогда $p_0 \notin A'$, а значит, $p_0 \notin A'/A$. Следовательно, не существует пикселей, принадлежащих A'/A, т. е. $A'/A = \varnothing$.

Предположим теперь, что $p_0 \in A/A'$ и имеет соседние пиксели $p_i \in A \cap A'$ (такой пиксель существует, так как $A' \cap A \neq \varnothing$). Поскольку для всех пикселей фрагмента ***A*** выполняется условие 1, то $\|I(p_0) - I(p_i)\| \leq \nu$. Это означает, что p_0 и p_i входят в один фрагмент, т. е. $p_0 \in A \cap A' \Rightarrow p_0 \in A' \Rightarrow p_0 \notin A/A'$. Следовательно, не существует пикселей, принадлежащих A/A' , т. е. $A/A' = \varnothing$

Получим $A' \cap A \neq \varnothing$, $A/A' = \varnothing$, $A'/A = \varnothing$. Следовательно, $A = A'$. Таким образом, утверждение доказано.

Нетрудно показать, что описанный алгоритм имеет линейную сложность. Действительно, каждый пиксель изображения может быть помещен в очередь только один раз. Так как пикселей всего TV, то алгоритм фрагментации имеет сложность *O(N),* где TV - количество пикселей изображения. Управлять работой алгоритма можно путем выбора различного порога и нормы, использующейся для сравнения пикселей.

***

В описанном подходе решение задачи поиска лиц получаем в результате последовательного решения подзадач фрагментации, описания и распознавания. Построен алгоритм, который имеет линейную сложность, инвариантен относительно масштаба и поворотов лица и применим для изображений с произвольным фоном (все это позволяет использовать его в системах реального времени). Получены достаточные условия, при которых возможно построение точного решения.

Выделение подзадач в совокупности с полученными теоретическими результатами позволяет управлять точностью и сложностью решения на всех этапах его получения.

Критической с практической точки зрения является подзадача фрагментации. В случае, когда выделение эллипсоидного фрагмента по каким-то причинам невозможно (например, не видна часть лица или оно находится не во фронтальном положении и, как следствие, не выполняются условия разделимости), требуется дополнительная предобработка, направленная на устранение этих ограничений. Этого можно достигнуть, например, путем сегментации кожи, направленной фильтрации, подчеркивания границ, отсечения узких мест и т. д. При этом необходимо контро- лировать уровень сложности используемых алгоритмов.

1. Брилюк Д. Распознавание человека по изображению лица. Препринт (http:// neurofacc.narod.ru).

2. Hjelmas E., Low B. K.//Journal of Computer Vision and Image Understanding. 2001. Vol. 83. P. 236.

3. Yang M.-H., Kricgman D. J., Ahuja N. Il IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. Vol. 24. № 1. P. 34.

4. Craw 1., Ellis H., Lishman J. Il Pattern Recognition Letters. 1987. Vol. 5. P. 183.

5. Rowley H. A., BaIuja S., Kanadc T. Il IEEE Transactions on Pattern Analysis and Ma chine Intelligence. 2001. Vol. 20.

6. Vezhnevets V.// Proceedings of the Third International Conference on Digital Information Processing and Control in Extreme Situations. Minsk, 2002. P. 293.

7. Адерихо К., Киселевский Jl., Косюкевич С, Краснопрошин В. Физиче-ские основы дистанционного зондирования. Мн., 1991. С. 293.

8. Kukharcv G., Masic P., Masicz P.// Pattern Recognition and Information Processing. 2003. Vol. 2. P. 195.

9. Хори Б.К.П. Зрение роботов. М., 1989.

Поступила в редакцию 28.06.06.

Виктор Владимирович Краснопрошин **- кандидат физико-математических наук, доцент, заведующий кафедрой математического обеспечения автоматизированных систем управления.**

Евгений Владимирович Коблов **- аспирант. Научный руководитель - В.В. Краснопрошин.**