

Р. В. Маслий

ИСПОЛЬЗОВАНИЕ ЛОКАЛЬНЫХ БИНАРНЫХ ШАБЛОНОВ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ НА ПОЛУТОНОВЫХ ИЗОБРАЖЕНИЯХ

В статье описаны методы распознавания лиц, которые используют локальные бинарные шаблоны. Предложен метод распознавания лиц, который использует локальный бинарный шаблон в форме эллипса. Для исследования метода использована база лиц ORL.

Ключевые слова: *распознавание лиц, локальные бинарные шаблоны, полутоновые изображения, расстояние хи-квадрат.*

Введение

Сегодня существует большое количество приложений, в которых используются методы распознавания лиц, таких как: идентификация личности, контроль удостоверений личности, использование интерфейса „человек-компьютер” и т. п. Главные сложности при распознавании лиц – множество изменений при представлении лиц, таких как: различное освещение, выражение лица, поворот головы, возрастные изменения и т. п. Поэтому проблема автоматического распознавания лиц полностью не решена до сих пор [1].

Процесс идентификации личности, который базируется на распознавании по изображению лица, состоит в том, что изображение лица неизвестной личности подается на вход системы распознавания, в которой сравнивается с изображениями лиц известных личностей, находящихся в базе изображений. Процесс идентификации можно разделить на три основных этапа:

- регистрация и нормализация изображения лица;
- выделение признаков;
- классификация.

В последнее время особенное внимание исследователей для выделения признаков из изображения лица привлекают методы, которые используют локальные бинарные шаблоны (ЛБШ). ЛБШ впервые были предложены в 1996 году для анализа текстур полутоновых изображений [2]. Методы распознавания лиц, которые используют для выделения признаков ЛБШ и их модификации, демонстрируют высокие результаты как по скорости, так и по точности распознавания [3-5]. Такие методы робастны при использовании изображений лиц с различной мимикой, различным освещением, поворотами головы.

Целью данной **работы** является модификация ЛБШ, которая бы позволила повысить точность распознавания изображений лиц по сравнению с использованием классических ЛБШ и их модификаций, а также классических методов распознавания лиц (анализ главных компонент, сравнение эластичных графов).

1. Локальные бинарные шаблоны

ЛБШ представляет собой описание окрестности пикселя изображения в двоичной форме. Оператор ЛБШ, который применяется к пикселю изображения, использует восемь пикселей окрестности, принимая центральный пиксель в качестве порога. Пиксели, которые имеют значения больше, чем центральный пиксель (или равное ему), принимают значения «1», те, которые меньше центрального, принимают значения «0». Таким образом получается восьмиразрядный бинарный код, который описывает окрестность пикселя. Пример работы оператора ЛБШ над полутоновым изображением показан на рис. 1.

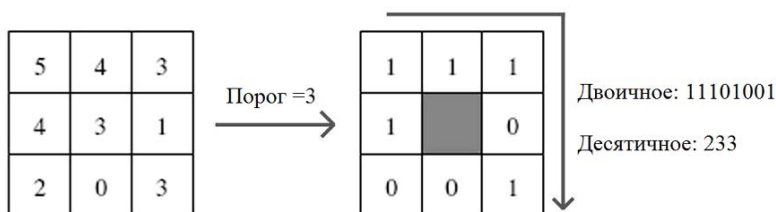


Рис. 1. Пример работы ЛБШ оператора

В работе [3] была описана модификация ЛБШ, в которой используют пиксели окрестности, находящиеся на некотором расстоянии от центрального пикселя. В этом случае пиксели окрестности лежат на круге с радиусом R . Количество точек этого круга может быть выбрано произвольно, обозначим их как P . Для вычисления значений в этих точках для разных радиуса R и количества точек P используется билинейная интерполяция. Для представления ЛБШ с радиусом R и расстоянием P будем использовать обозначение $LBP_{P,R}$. На рис. 2 представлены наборы пикселей окрестности для разных P и R .

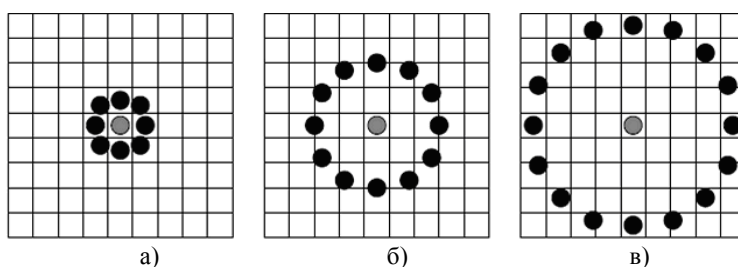


Рис. 2. Модифицированный ЛБШ для различных P и R :
 а) количество точек $P = 8$, радиус окрестности $R = 1.0$;
 б) количество точек $P = 12$, радиус окрестности $R = 2.5$;
 в) количество точек $P = 16$, радиус окрестности $R = 4$

В работе [2] было предложено для анализа изображения использовать не все ЛБШ, а только те, которым свойственно иметь не более двух переходов с «1» в «0» или наоборот. Такие шаблоны называют *uniform pattern*, обозначим их как $LBP_{P,R}^u$. Например, такие ЛБШ как 1111111 или 00001111 – это $LBP_{P,R}^u$, а 10101010 или 11001100 – это не $LBP_{P,R}^u$. Использование $LBP_{P,R}^u$ имеет два преимущества. Во-первых, происходит экономия памяти, потому что при использовании $LBP_{P,R}^u$ для анализа изображения нужно лишь $P(P-1)+2$ ЛБШ, тогда как для представления всех $LBP_{P,R}$ нужно использовать 2^P ЛБШ. Во-вторых, $LBP_{P,R}^u$ определяют только важные локальные текстуры, такие как концы линий, края, углы, пятна. Примеры таких $LBP_{P,R}^u$ приведены на рис. 3, где светлыми кружками обозначены пиксели, значения которых равно «0», темными кружками обозначены пиксели, значения которых равно «1».

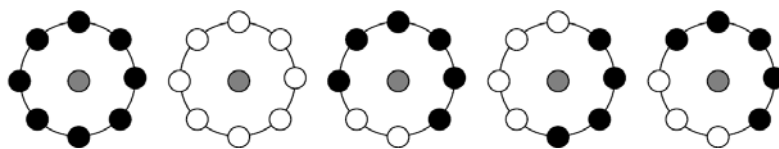


Рис. 3. Примеры $LBP_{P,R}^u$

В работе [5] авторы используют гибридный подход для описания изображения лица, в котором изображения описаны как локально, так и глобально. Сначала изображение лица Наукові праці ВНТУ, 2008, № 4

делится на $k*k$ областей. Потом в каждой области для каждого пикселя изображения вычисляется $LBP_{P,R}$ -код. При этом увеличивается на единицу значения столбика, который отвечает за этот код. Каждому $LBP_{P,R}^u$ -коду в гистограмме соответствует отдельный столбик, еще один столбик в гистограмме соответствует всем $LBP_{P,R}$ -кодам, которые не имеют свойства $LBP_{P,R}^u$. Таким образом, каждая гистограмма области состоит из $P(P-1)+3$ столбиков. После этого для представления глобального описания изображения лица все гистограммы областей объединяют в одну гистограмму. Эта гистограмма будет состоять из $k^2(P*(P-1)+3)$ столбиков и будет формировать вектор признаков лица. Пример деления изображения лица на области и соответствующие этим областям гистограммы изображены на рис. 4.

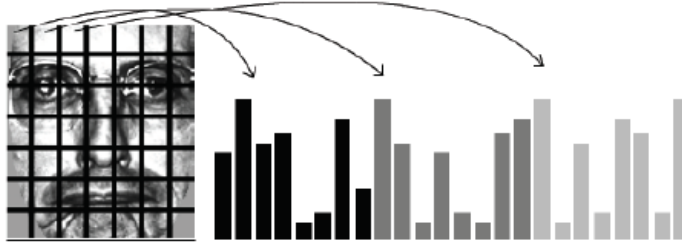


Рис. 4. Гистограммы из областей изображения

Для изображения размером $N*M$ вектор признаков лица строится вычислением кода ЛБШ для каждой точки (x_c, y_c) , где $x_c \in \{R+1, \dots, N-R\}$ $y_c \in \{R+1, \dots, N-R\}$. Если изображение разделено на $k*k$ областей, гистограмма области (k_x, k_y) , где $k_x \in \{1, \dots, k\}$ и $k_y \in \{1, \dots, k\}$ будет определена как:

$$H_i(k_x, k_y) = \sum_{x,y} I\{LBP_{P,R}(x, y) = L(i)\}, i = 1, \dots, P(P-1) + 3, \quad (1)$$

где

$$x \in \begin{cases} \{R+1, \dots, N/k\} & k_x = 1 \\ \{(k_x - 1)(N/k + 1, \dots, N - R)\} & k_x = k \\ \{(k_x - 1)(N/k + 1, \dots, k_x(N/k))\} & \text{иначе} \end{cases},$$

$$y \in \begin{cases} \{R+1, \dots, M/k\} & k_y = 1 \\ \{(k_y - 1)(M/k + 1, \dots, M - R)\} & k_y = k \\ \{(k_y - 1)(M/k + 1, \dots, k_y(M/k))\} & \text{иначе} \end{cases}.$$

и L название i -го столбика

$$L(i) = \begin{cases} 1, & i - \text{истина} \\ 0, & i - \text{ложь} \end{cases}.$$

При классификации изображений лиц для нахождения наименьшего расстояния между гистограммами, которые представляют лица, используется расстояние хи-квадрат. Расстояние между изображениями S и M определяется по формуле:

$$\chi^2(S, M) = \sum_{j=1}^{k^2} \left(\sum_{i=1}^{P(P-1)+3} \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}} \right). \quad (2)$$

В формуле (2) $S_{i,j}$ и $M_{i,j}$ – это размеры столбика i из области j (количество появлений ЛБШ $L(i)$ в области j).

Так как некоторые области лиц (например, область глаз) могут нести более важную

информацию, чем другие области, каждой области в зависимости от ее важности для распознавания могут быть присвоены веса.

При присвоении веса w_j области j формулу можно записать следующим образом:

$$\chi_w^2(S, M) = \sum_{j=1}^{k^2} w_j \left(\sum_{i=1}^{P(P-1)+3} \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}} \right). \quad (3)$$

2. Метод распознавания

В данной работе предложено использовать ЛБШ в форме эллипса (ЭЛБШ), то есть для формирования ЛБШ будут использоваться те пиксели окрестности, которые лежат на эллипсе относительно центрального пикселя. Обозначим больший радиус эллипса как A , меньший радиус как B , количество точек в ЛБШ обозначим как m . Тогда координаты g_{ix} и g_{iy} для каждого пикселя окрестности $g_i (i=1,2,\dots,m)$ будут определены по формулам (4) и (5).

$$R_i = \sqrt{\frac{A^2 B^2}{A^2 \sin^2 \theta_i + B^2 \cos^2 \theta_i}}. \quad (4)$$

$$g_{ix} = R_i \cdot \cos \theta_i, \quad g_{iy} = R_i \cdot \sin \theta_i, \quad (5)$$

где $\theta_i = \left(\frac{360}{m} * (i-1) \right)^\circ$.

На рис. 5 представлен пример ЭЛБШ с разными значениями A , B и m .

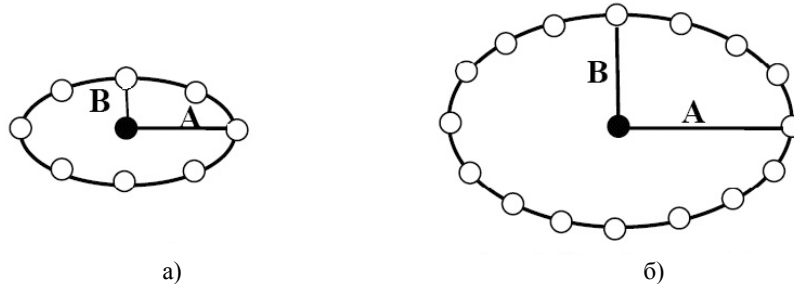


Рис. 5. ЭЛБШ:
а) $A=2, B=1, m=8$;
б) $A=3, B=2, m=16$

Координаты точек окрестности не всегда попадают точно в центры пикселей, поэтому для вычисления значений этих точек используем билинейную интерполяцию. Вектор признаков ЭЛБШ одной области изображения лица будет формироваться по формуле (1). Для классификации изображений лиц в данной работе будут использоваться формулы (2) и (3). Для использования формулы (3) нужно задать веса некоторым областям лица. Веса областей w подобранные эмпирическим путем изображены на рис. 6.

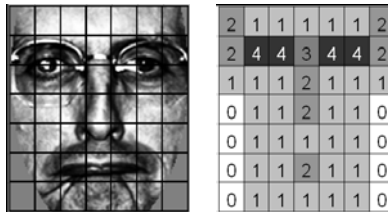


Рис. 6. Веса областей изображения лица

3. Результаты экспериментов

В данной работе для исследования предложенного метода используется база лиц Наукові праці ВНТУ, 2008, № 4

исследовательской лаборатории Оливетти (ORL-Olivetti research laboratory). База состоит из 400 изображений лиц 40 человек (по 10 разных изображений лица одного человека), у которых разные выражения лица (например, с открытыми/закрытыми глазами, с улыбкой/без улыбки) и также наличие/отсутствие очков. Все изображения лиц находятся на темном фоне, в фронтальной позиции с наклонами и поворотами головы до 20 градусов и изменением масштаба до 10%. Все изображения полутоновые и имеют размер 92*112 пикселей. На рис. 7 приведен пример изображений лица особы с базы ORL.



Рис. 7. Изображения лица человека с базы ORL

В экспериментах были использованы ЛБШ в форме эллипса с такими параметрами A , B и m : $A_1 = 1$, $B_1 = 1$, $m = 8$ (ЭЛБШ₁); $A_2 = 3$, $B_2 = 1$, $m = 16$ (ЭЛБШ₂); $A_3 = 3$, $B_2 = 2$, $m = 16$ (ЭЛБШ₃).

В таблице 1 приведены точность распознавания классических методов распознавания (анализ главных компонент (АГК), сравнения эластичных графов (СЭГ)), методов распознавания, использующих ЛБШ, а также точность распознавания разработанных методов, использующих модифицированные ЛБШ.

Таблица 1

Метод распознавания	Точность распознавания
АГК [6]	85%
СЭГ [7]	87.5%
ЛБШ [5]	94%
ЭЛБШ ₃	92%
ЭЛБШ ₁ + веса областей	93.5%
ЭЛБШ ₂ + веса областей	94.5%
ЭЛБШ ₃ + веса областей	95%

Из таблицы 1 видно, что разработанные методы распознавания, использующие модифицированные ЛБШ, при использовании базы ORL демонстрируют лучшую точность распознавания по сравнению с классическими методами (АГК, СЭГ). По сравнению с наилучшим методом распознавания, который использует ЛБШ в работе [5], разработанные методы демонстрируют близкие результаты.

Выводы

Экспериментально исследовано целесообразность использования ЭЛБШ для распознавания лиц. Совместно с присваиванием веса областям лица использование ЭЛБШ позволило увеличить точность распознавания до 95%, используя базу лиц ORL. Для повышения точности распознавания в дальнейшем планируется усовершенствовать методы предварительной обработки изображения лица, усовершенствовать способ деления лица на области, найти оптимальные веса областей, используя для этого алгоритм AdaBoost.

СПИСОК ЛИТЕРАТУРИ

1. W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4): 399 – 458, 2003.
2. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on Наукові праці ВНТУ, 2008, № 4

feature distributions. *Pattern Recognition* 29(1), 51–59 (1996).

3. Ojala, T., Pietikainen, M., Maenpaa, M.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002).

4. Zhang, G., Huang, X., Li, S.Z., Wang, Y., Wu, X.: Boosting local binary pattern (LBP)-based face recognition. In: Li, S.Z., Lai, J.-H., Tan, T., Feng, G.-C., Wang, Y. (eds.) *SINOBIOMETRICS 2004*. LNCS, vol. 3338, pp. 180–187. Springer, Heidelberg (2004).

5. Ahonen T., Hadid A., Pietikainen M.: Face description with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(12): 2037-2041. (2006).

6. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* 3, 71–86 (1991)

7. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE PAMI* 19(7), 711–720 (1997)

Маслій Роман Васильевич – ассистент кафедры автоматизации и информационно-измерительной техники.

Винницкий национальный технический университет.