



УДК 004.932.2

КОНТЕКСТНО-ЗАВИСИМЫЙ АНАЛИЗ ПОРТРЕТНЫХ ФОТОГРАФИЙ

Алексеев А.В., Орлова Ю.А.

Волгоградский государственный технический университет, г. Волгоград, Россия

Alekseev.yeskela@gmail.com

В данной работе представлена методика контекстно-зависимого анализа портретных фотографий. Весь анализ автоматический, вмешательство пользователя не требуется. Разработанная методика позволяет локализовать лицо и найти контуры рта и глаз. В заключении показаны примеры работы на тестовых фотографиях.

Ключевые слова: обработка изображений; контекстно-зависимый анализ; анализ портретных изображений; лицо человека; автоматический поиск.

Введение

В настоящее время в связи со стремительным развитием цифровой фотографии и цифрового видео очень перспективной является задача распознавания образов на цифровых изображениях и лиц в частности.

Последние десять лет в области распознавания лиц и их элементов ведутся активные разработки и предложены различные методы распознавания, использующие такие подходы, как метод сверточных и обычных нейронных сетей [Розалиев и др., 2010], скрытый марковских моделей [Nefian, 1999], активные модели формы [Prabhu et al], активные модели внешнего вида [Mitchell et al, 2002], методы с использованием гистограмм, статистические методы и т. д.

Распознавание лиц и элементов лиц может применяться в следующих областях: системы автоматического учета числа посетителей; системы пропускного контроля в учреждениях, аэропортах и метро; автоматические системы предотвращения несчастных случаев; интеллектуальные интерфейсы "человек-компьютер" и др. Нахождение различных элементов лица позволяет вести учет дополнительной статистики и получать больше информации о человеке. Задача обнаружения лица человека (а также элементов лица) является сложной ввиду нескольких основных причин: высокая вариативность лиц человека, обусловленная анатомическими и фенотипическими особенностями индивидов; различные условия освещенности, определяющиеся типом, количеством и направлением источников света.

Целью данной работы является автоматический поиск лиц на изображении и выделение контура для каждого лица контуров следующих элементов: рот, глаза.

Предлагаемая методика

Мы предлагаем следующую методику контекстно-зависимого анализа лиц:

1. Детектирование лиц на основе метода Виолы-Джонса, хорошо зарекомендовавший себя метод с высоким качеством и скоростью работы.

2. Все искомые элементы будут находиться в зависимости от найденного прямоугольника лица, исходя из физиологии человека. Глаза симметричны и находятся в верхней части лица (на симметрию накладывается ограничения в случае незначительных наклонов головы, тогда глаза не будут симметричны относительно найденного прямоугольника). Рот находится посередине в нижней части лица.

3. Локализовать рот мы будем на основе эвристически подобранного соотношения относительно прямоугольника лица. Уточнение границ рта мы будем делать методом сегментации GrabCut, метод трактует все изображение как граф, и по заданным правилам в этом графе ищется минимальный разрез, в результате сегментации мы получаем две области: объект и не объект.

4. Локализовать глаза мы будем также с помощью метода Виолы-Джонса, только в этом раз мы будем искать глаза в пределах верхней части лица, также дополнительно будет учитываться то, что глаза должны быть не цвета кожи.

Детектирование лица

Для детектирования лиц мы используем метод Виолы-Джонса [Viola et al, 2001]. Метод был разработан в 2001 году Полом Виолой и Майклом Джонсом, на данный момент является одним из лучших по соотношению качества и скорости для детектирования объектов на изображении (лиц в частности). Основные принципы, на которых основан метод, таковы:

1) Используются изображения в интегральном представлении, что позволяет быстро вычислять необходимые объекты. Интегральное представление изображения – это матрица, размерность которой совпадает с размерностью исходного изображения и каждый элемент интегрального изображения содержит в себе сумму пикселей изображения в прямоугольнике от пиксела с координатами (0,0) до него самого (x,y).

2) Используются признаки Хаара, с помощью которых происходит поиск нужного объекта (лица) (Рисунок 1). Признаки Хаара – признаки цифрового изображения, используемые в распознавании образов. Своим названием они обязаны интуитивным сходством с вейвлетами Хаара. Признак Хаара состоит из смежных прямоугольных областей. Они позиционируются на изображении, далее суммируются интенсивности пикселей в областях, после чего вычисляется разность между суммами. Эта разность и будет значением признака, заданного размера, определенным образом спозиционированного на изображении.

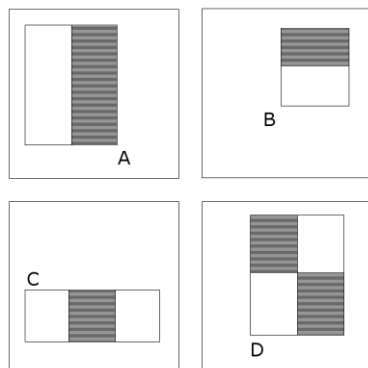


Рисунок 1 – Признаки Хаара

3) Используется бустинг для выбора наиболее подходящих признаков для искомого объекта на данной части изображения. Бустинг – комплекс методов, способствующих повышению точности аналитических моделей. Эффективная модель, допускающая мало ошибок классификации, называется «сильной». «Слабая» же, напротив, не позволяет надежно разделять классы или давать точные предсказания, делает в работе большое количество ошибок. Поэтому бустинг означает дословно «усиление» «слабых» моделей [Jan et al, 2010] – это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.

4) Все признаки поступают на вход классификатора, который даёт результат «истина» - объект есть, либо «ложь» - объекта нет.

5) Используются каскады классификаторов для быстрого отбрасывания окон, где не найдено лицо. Каскад представляет собой цепочку классификаторов, каждый из которых обучается на ошибках предыдущего. Объект считается найденным только тогда, когда все классификаторы каскада вернули «истину».

Детектирование рта

1.1. Локализация рта

Детектирование рта мы проводим в 2 этапа, первый – получение ограничивающего прямоугольника, это делается на основе физиологии лица, в результате экспериментов были получены границы, которые зависят от полученного ранее прямоугольника лица. Прямоугольник рта относительно лица определяется следующими соотношениями:

- Ширина рта = 50% от ширины лица;
- Высота рта = 30% от высоты лица;
- Относительно верхнего левого угла прямоугольника лица:
 - Координата x (ширина) рта = $0,25 * \text{ширину лица}$;
 - Координата y (высота) рта = $0,7 * \text{высоту лица}$;

На рисунке 2 представлен локализирующий рот прямоугольник (зеленый), красный прямоугольник – найденное лицо. Здесь и далее все изображения лиц взяты с базы MICT [Nicolls et al, 2010].



Рисунок 2 – Локализация рта

1.2. Нахождение контура рта

Границы рта уточняются на основе алгоритма сегментации GrabCut. Данный алгоритм основан на приближенном построении минимального разреза специального графа, который строится по пикселям изображения. Найденный контур рта представлен на рисунке 3.



Рисунок 3 – Контур рта

Детектирование глаз

1.3. Локализация глаз

Для детектирования глаз используется также метод Виолы-Джонса, специальный обученный каскад, как и каскад для детектирования лиц были взяты из открытой библиотеки компьютерного зрения OpenCV [OpenCV, 2014]. Кроме того, используется предположение о том, что глаза находятся в верхней части лица (остальные детектирования отбрасываются как ложные). На рисунке 4 представлен пример детектирования глаз с помощью каскада: красный прямоугольник – лицо, синие – примерная область нахождения глаз, зеленые – найденные глаза.

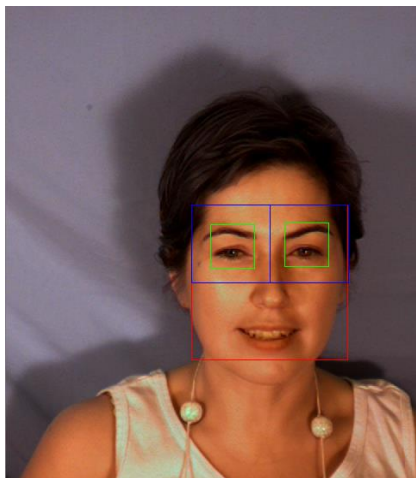


Рисунок 4 – Локализация глаз

1.4. Нахождения контура глаз

Контур глаз мы будем находить, ориентируясь на то, что глаза (непосредственно белки и зрачки) не телесного цвета. Цвет тела будем определять согласно [Vezhnevets et al, 2003] по формуле (1):

$$R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and } R > B \\ \text{and } R - G > 15 \text{ and } \max(R, G, B) - \min(R, G, B) > 15 \text{ and } (H < 25 \text{ or } H > 230) \quad (1)$$

где: R, G, B – соответственно значения красного, зеленого и синего каналов в цветовой модели RGB;

H – значение канала цветового тона в цветовой модели HSV.

Цвет пиксела считается цветом кожи тогда, когда значения соответствующих каналов удовлетворяют формуле (1).

Полученная маска кожи и выделенные контуры глаз с учетом локализации представлены на рисунке 5.



Рисунок 5 – Выделение контуров глаз

Результаты

В данной работе представлен метод контекстно-зависимого анализа портретных фотографий, который заключается в поиске лиц и локализации для каждого из найденных лиц глаз и рта, а также нахождения их контуров. Данная методика представляет собой комбинацию других методов, поэтому на нее накладываются ограничения, которые накладываются на каждый используемый метод в отдельности [Заболеева-Зотова и др., 2010]. А именно, метод детектирования Виолы-Джонса чувствителен к сильным поворотам лица, каскады для глаз чувствительны к очкам, эта проблема решается путем использования каскадов, специально обученных для детектирования глаз в очках или же лиц в профиль. Так же детектирование рта может работать некорректно в случае наличия обильной растительности на лице, GrabCut может посчитать за искомый объект именно усы, а не рот. Результаты работы предложенной методики изображены на рисунке 6, найденные рот и глаза выделены цветом.



Рисунок 6 – Детектирование рта и глаз

Заключение

В данной работе была представлена методика автоматического поиска лиц и выделения на них таких элементов как рот и глаза. Представлены отдельные этапы данной методики, указаны ограничения на применение, как отдельных этапов, так и всей методики в целом.

Работа частично поддержана Российским фондом фундаментальных исследований (проекты 13-07-00459, 13-07-97042, 14-07-97016, 15-07-06322).

Библиографический список

- [Алексеев и др., 2012] Алексеев, А.В. Автоматизация определения шрифтов по изображению / Алексеев А.В., Розалиев В.Л. // Труды конгресса по интеллектуальным системам и информационным технологиям "IS&IT'12" (Дивноморское, Краснодарский край, 2-9 сент. 2012 г.). В 4 т. Т. 1 : тр. конф. "Интеллектуальные системы '12" и "Интеллектуальные САПР – 2012" / ЮФУ [и др.]. - М., 2012. - С. 292-293.
- [Заболеева-Зотова и др., 2010] Автоматизация начальных этапов проектирования программного обеспечения / Заболеева-Зотова А.В., Орлова Ю.А. // Изв. ВолгГТУ. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах»: межвуз. сб. науч. ст. - Волгоград, ВолгГТУ. 2010. - Вып. 8, № 6. - С. 121-124.
- [Розалиев и др., 2010] В.Л. Розалиев, А.С. Бобков, О.С. Федоров Применение нейронных сетей и грануляции при построении автоматизированной системы определения эмоциональных реакций человека / Изв. ВолгГТУ. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах»: межвуз. сб. науч. ст. / ВолгГТУ. - Волгоград, 2010. - Вып. 9, № 11. - С. 63-68.
- [Jan et al, 2010] Jan Šochman, Jiří Matas, «AdaBoost», Center for Machine Perception, Czech Technical University, Prague, 2010
- [Mitchell et al] S. C. Mitchell, J. G. Bosch, B. P. F. Lelieveldt, R. J. van der Geest, J. H. C. Reiber, and M. Sonka. 3-d active appearance models: Segmentation of cardiac MR and ultrasound images. IEEE Trans. Med. Imaging, 21(9):1167–1178, 2002
- [Nefian, 1999] Nefian, A.V. Hidden Markov Model-Based Approach for Face Detection and Recognition / Ara Nefian. 1999.
- [Nicolls et al, 2010] Nicolls S.M. and J.M. and F. The MUCT Landmarked Face Database // Pattern Recognit. Assoc. South Africa. 2010.
- [OpenCV, 2014] OpenCV – Open computer vision library, Itseez, Retrieved from <http://opencv.org>, 2014
- [Prabhu et al] Prabhu U., Seshadri K. Facial Recognition Using Active Shape Models, Local Patches and Support Vector Machines // contrib.andrew.cmu.edu. P. 1–8.
- [Vezhnevets et al, 2003] Vezhnevets V., Sazonov V., Andreeva A. A survey on pixel-based skin color detection techniques // Proc. Graph. Moscow, Russia, 2003. - P. 85–92.
- [Viola et al, 2001] P. Viola and M.J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001

CONTEXT-SENSITIVE ANALYSIS OF PORTRAIT IMAGES

Alekseev A.V., Orlova Yu.A.

Volgograd State Technical University

Alekseev.yeskela@gmail.com

This paper presents a method of context-sensitive analysis of the portrait photos. The entire analysis automatically, no user interaction is required. The presented method allows to locate the person and find the contours of the mouth and eyes. Finally, examples of work on the test photos.

Introduction

Currently, due to the rapid progressing of digital photography and digital video pattern recognition in digital images and faces in particular is very perspective task.

The aim of this work is the automatic search of faces on the image and contour selection for each person outlines the following elements: mouth, eyes.

Main Part

The proposed method consists in the following:

1. Detection of persons on the basis of the viola-Jones;
2. All required elements will be depending on the found rectangle face, based on human physiology;
3. To localize the mouth we will be on the basis of heuristically selected ratio relative to the rectangle face. The delimitation of the mouth are we going to do with the method of the GrabCut segmentation;
4. To localize the eyes we will also use the methods of the viola-Jones, we clarify the boundaries on the basis of skin color

Conclusion

This work was presented a technique to automatically search faces and allocation of items such as mouth and eyes. This paper presents the individual steps of this method, constraint specified on the application as the individual steps and the overall methodology.