



ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И МАШИННОЕ ОБУЧЕНИЕ

А.А. Александров

РАСПОЗНАВАНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА ПО ИЗОБРАЖЕНИЮ ЛИЦА

(Казанский национальный исследовательский технический университет им.
А.Н. Туполева-КАИ)

Одним из важнейшим фактором общения между людьми является обмен эмоциональной информацией. Поэтому при создании современных систем по взаимодействию с человеком актуально применение методов автоматического распознавания эмоций.

О эмоциональном состоянии человека можно судить различным факторам, включавшим в себя рефлекторные движения, жестикуляцию и прочие рефлексы, но мимика является универсальным, не зависящим от расовой или культурной принадлежности источником, что было подтверждено еще 30-х годах прошлого века. Поэтому задача распознавания мимики является одной из важнейших для определения эмоций.

Для человека распознавание эмоций является тривиальной задачей. На результат автоматического распознавания могут повлиять положение лица и внешние условия. Поэтому, для повышения качества обработки используют первичную или предварительную обработку. Также, стоит отметить, для определения эмоций человек считывает соответствующие левые признаки.

Для решения задачи распознавания эмоций были выбраны: метод Виолы-Джонса, метод гистограмм локальных бинарных паттернов (LBPН) и дерева решений.

Основной алгоритм определения эмоционального состояния состоит из следующих этапов:

1. Получение изображения;
2. Предварительная обработка изображения;
3. Поиск лица на изображении;
4. Разбиение лица на области;
5. Выделение элементов лица;
6. Классификация элементов лица;
7. Классификация эмоций.

Под предварительной обработка изображения подразумевается удаление шумов, цветовые и геометрические преобразования. Удаление шумов производится посредством усредняющей фильтрации.



Для поиска лица на изображении наиболее эффективным показал себя метод Виолы-Джонса [1]. Данный метод обеспечивает хорошую скорость и высокую точность обнаружения. В методе применяется принцип сканирующего окна. В качестве исходного подается изображение размерностью $N \times M$ пикселей, в которой каждый пиксель характеризуется значением от 0 до 255. Скользящее окно проходит по всем пикселям изображения и в каждом его положении применяется классификатор, основанный на методе Бустинга – усиление слабых классификаторов [2].

Методом Виолы-Джонса производится обнаружение объектов на изображении в интегральной форме, что обеспечивает быстрое вычисление суммарной яркости прямоугольника любого размера на исходном изображении за постоянное время, независимо от размеров этого прямоугольника. В качестве признаков используются признаки Хаара, на основе вейвлетов Хаара (Рис. 9).

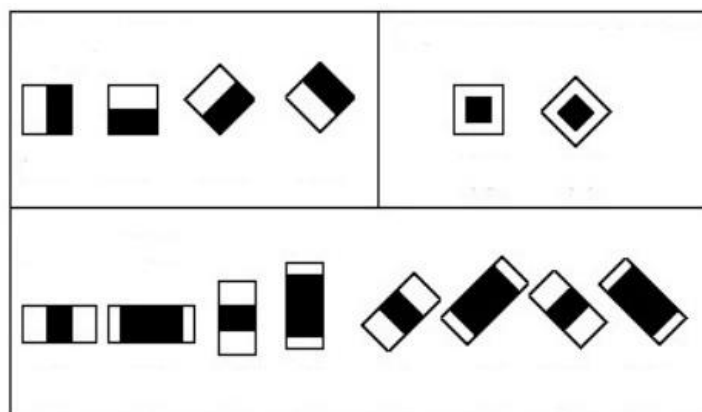


Рис. 9. Признаки Хаара

После того как лицо было выделено производим разбиения лица на три локальные области: область лба, область глаз и область рта.

Область лба не имеет отличительных признаков и определяется в зависимости от области глаз. Данная область может содержать в себе небольшие деформации кожи и помехи в виде волос.

Область глаз включает в себя глаза, брови и верхнюю часть носа. Наилучшим вариантом поиска областей будет начать с данной области, так как в нее входят достаточное количество информативных признаков, что повышает точность выделения, а также она находится по середине лица, что прощает выделение остальных областей.

Область рта начинается с нижней части носа и содержит в себе губы, подбородок и щеки. Наибольший интерес в данной области представляют губы.

Человек проявляет эмоции с помощью лицевых экспрессий. Для выделения элементов на лице также используется метод Виолы-Джонса. В качестве входных данных подается выделенное на предыдущем этапе изображение лица. Для ускорения распознавания и уменьшения числа ложных обнаружений, на лице определяются зоны (рот находится в нижней части лица, а глаза – в верхней).



Следующим шагом является классификация элементов лица. В рамках данной работы используется метод гистограмм локальных бинарных паттернов (LBPН) [3].

Локальный бинарный шаблон (LBP) - это оператор текстуры, который маркирует пиксели изображения путем установки порогового значения окрестности каждого пикселя и рассматривает результат как двоичное число. Используя LBP в сочетании с гистограммами, мы можем представить изображения простым вектором данных.

Процесс классификации лица методом LBPН состоит из следующих шагов:

1. Первым вычислительным этапом LBPН является создание промежуточного изображения (Рис. 10), которое лучше описывает исходное изображение. Для этого алгоритм использует концепцию скользящего окна, основанную на параметрах радиуса и соседей.
 - 1.1. Выделение части изображения в виде окна размером $N \times N$ и представление его в виде матрицы $N \times N$, содержащей интенсивность каждого пикселя (0 ~ 255).
 - 1.2. Установление порогового значения равным значению центрального элемента матрицы.
 - 1.3. Бинаризация по установленному порогу.
 - 1.4. Теперь матрица будет содержать только двоичные значения (без учета центрального значения). Нужно объединить каждое двоичное значение из каждой позиции из матрицы строка за строкой в новое двоичное значение (например, 10001101).
 - 1.5. Затем следует преобразование этого двоичного значения в десятичное значение и установление его в центральное значение матрицы, которое фактически является пикселем от исходного изображения.
 - 1.6. В конце этой процедуры (процедура LBP) имеем новое изображение, которое лучше отражает характеристики исходного изображения.

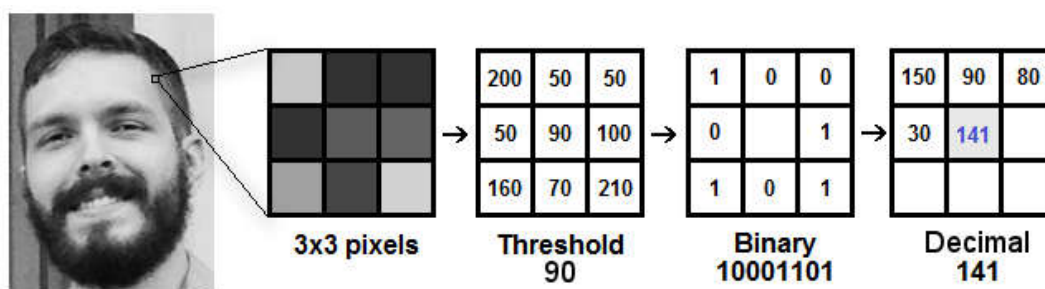


Рис. 10. Создание промежуточного изображения

2. Извлечение гистограмм. Теперь, используя изображение, сгенерированное на последнем шаге, используем параметры Grid X и Grid Y, чтобы разделить изображение на несколько сеток (Рис. 11). Поскольку у нас есть изображение в оттенках серого, каждая гистограмма (из каждой сетки) будет содержать только 256 позиций (0 ~ 255), представляющих вхождения каждой ин-



тенсивности пикселей. Затем нам нужно объединить каждую гистограмму, чтобы создать новую и большую гистограмму. Предположим, что у нас есть сетки 8×8 , мы будем иметь $8 \times 8 \times 256 = 16,384$ позиции в финальной гистограмме. Итоговая гистограмма представляет характеристики изображения исходного изображения.

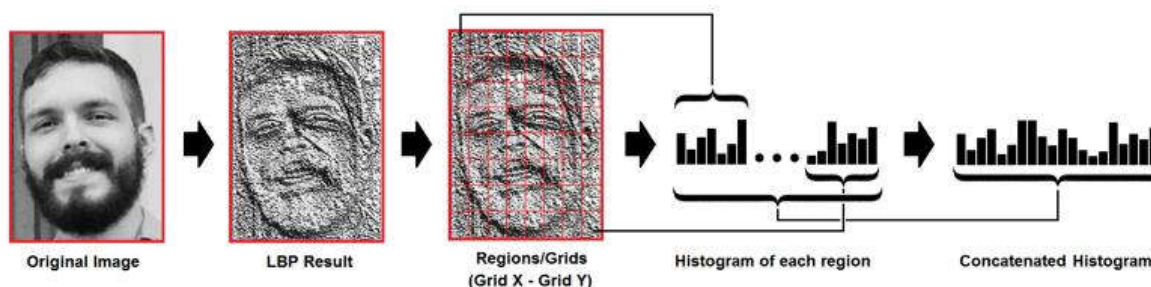


Рис. 11. Извлечение гистограмм

3. Выполнение распознавания.

3.1. Полученная гистограмма сравнивается с гистограммами исходного набора данных. Используются различные подходы для сравнения гистограмм (рассчитать расстояние между двумя гистограммами), например: евклидово расстояние, хи-квадрат, абсолютное значение и т.д.

3.2. Изображению присваивается тип, соответствующий наиболее схожему примеру из набора. Алгоритм также возвращает рассчитанное расстояние, которое можно использовать как «достоверное» измерение.

После классификация элементов лица можно выполнить классификацию эмоции. Для классификации эмоций применяется деревья решений. Дерево решений - математическая модель, которая задаёт процесс принятия решений так, что будут отображены каждое возможное решение, предшествующие и последующие этим решениям события или другие решения и последствия каждого конечного решения.

Литература

1. Ян Си, Автоматическое распознавание эмоций пользователя для организации интеллектуального интерфейсах / Ян Си // Электронный журнал «Молодежный научно-технический вестник», 2013. – 9 с.

2. Кирпичников А.П Контекстный поиск изображений / Кирпичников А.П., Ляшева С.А., Шлеймович М.П. // Вестник Казанского технологического университета, 2014. – С. 244-251.

3. Kelvin Salton do Prado, Face Recognition: Understanding LBPH Algorithm — Режим доступа: URL: <https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b> (12.03.2019)