

УДК 004.93'12, 004.932.1, 004.042

Система распознавания эмоций человека на видеоданных

Видьманов Д. А., аспирант

*Россия, 105005, г. Москва, МГТУ им. Н. Э. Баумана,
кафедра «Информационные системы и телекоммуникации»*

*Научный руководитель: Петросян О. Г., к.т.н., доцент,
Россия, 105005, г. Москва, МГТУ им. Н. Э. Баумана
iu3@bmstu.ru*

Введение

Способность «распознавать» видимые объекты - одно из основных свойств всех живых организмов. Человек постоянно осуществляет акт распознавания. Можно утверждать, что человек представляет собой очень сложную информационную систему, и некоторые функции этой системы, такие как распознавание, в последнее время очень востребованы в современных системах машинного зрения, обработки данных и принятия решений [1].

Данные системы создаются для решения различных прикладных задач: обеспечение безопасности, автоматизация управленческого учёта, создание человеко-машинных интерфейсов. Основные цели данных систем – улучшить качество и производительность труда, а также максимально облегчить процесс взаимодействия человека и компьютера.

Формализация задачи

Распознавания эмоций человека – прикладная задача машинного зрения, которая относится к классу алгоритмов высокого уровня цифровой обработки изображений – модуль принятия решений (понимание изображений) [2, 3]. Решение задачи связано с использованием алгоритмов среднего уровня (сегментация изображений), а также алгоритмов нижнего уровня (фильтрация шумов, гистограммная обработка). Под сегментацией понимается выделение простых яркостно-геометрических структур и областей на изображении, сопоставление точек, фрагментация.

Процесс распознавания – это разметка сцены, которая представляет собой проекцию трёхмерного рабочего пространства на плоскость объектива регистрирующего устройства – цифровую камеру. В процессе разметки происходит выделение информации

относительно объектов, расположенных в поле зрения цифровой камеры или другого регистрирующего устройства.

Образ представляет собой описание распознаваемого объекта. Также под образом подразумевается некоторая упорядоченная совокупность дескрипторов, которые также называются признаками [4]. Класс образов – совокупность образов, обладающих некоторыми общими свойствами.

Для реализации распознавания и идентификации различных образов и явлений на видеоданных необходимо решить следующие задачи:

1. Сбор и обработка видеоданных в режиме реального времени или обработка видеоданных из архива.
2. Захват исследуемого объекта или явления в кадре видеоданных.
3. Создание базы данных примитивов исследуемого объекта или явления.
4. Управление базой данных примитивов во время работы системы.
5. Обучение и переобучение системы по базе данных примитивов.
6. Автоматическое и ручное управление процессом распознавания исследуемого объекта или явления.

Общая структура системы

Выбранная структура системы распознавания образов соответствует общепринятой [1] и показана на рис.1.

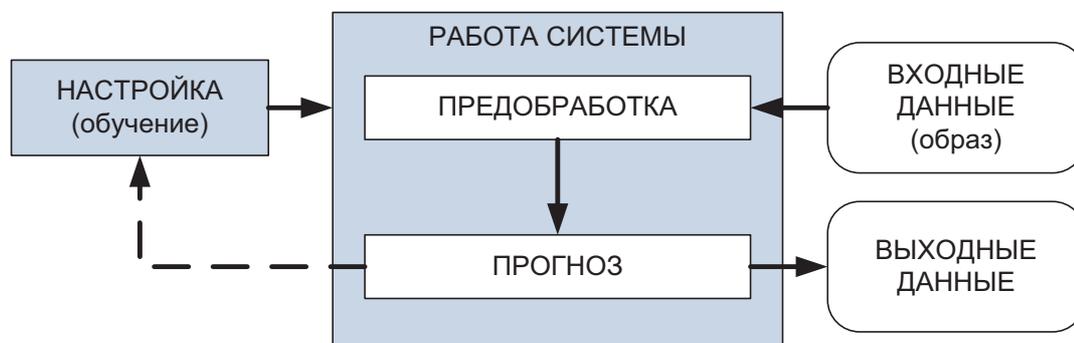


Рис. 1. Структура системы распознавания образов

Входные данные – исходное изображение, которое изменяется в модуле предобработки для устранения шума или кодировки (подготовки данных) под выбранный алгоритм распознавания. Далее система осуществляет прогнозирование результата.

Выходные данные – выбранная классификация и выделенный объект в исходных данных, который относится к распознанному классу образов.

Система должна быть переобучаема. В частности, для неизвестного класса образов, полученного из входных данных, создается отдельный набор примитивов, который используется при переобучении.

Аппаратное и программное обеспечение

Персональный компьютер с подключенной цифровой USB видеокамерой - аппаратное обеспечение системы.

Операционная система Windows, пакет разработчика Microsoft Visual Studio 2010 и библиотека OpenCV 2.4.2 – программное обеспечение разрабатываемой системы. Используемый язык программирования – C++.

OpenCV является открытой библиотекой, и использование доступных готовых процедур и функций не нарушает авторских прав.

Для разработки системы использовался проект по созданию «Консольного приложения Win32».

Внешний вид интерфейса

Внешний вид интерфейса системы приведен на рис.2.

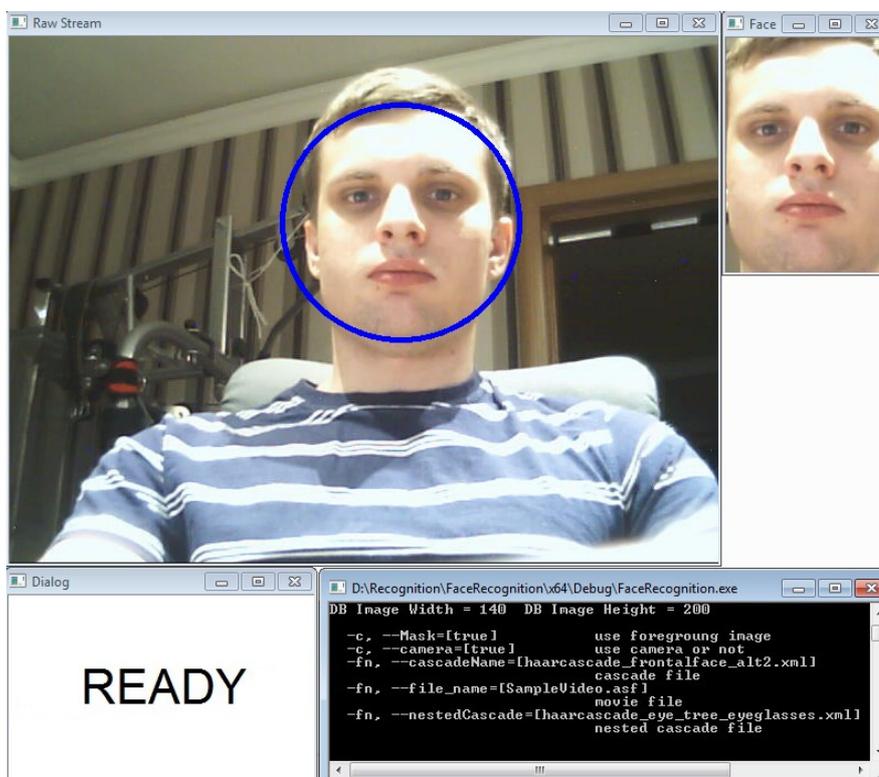


Рис. 2. Внешний вид интерфейса системы

Диалог вывода состоит из 3-х окон, созданных при помощи predefined процедур OpenCV:

- Окно «Raw Stream» - поток необработанных (исходных) видеоданных, полученный при помощи цифровой камеры или видеофайла.
- Окно «Face» - выделенный при помощи классификатора фрагмент кадра, содержащий лицо человека.
- Окно «Dialog» - основное диалоговое окно, на которое выводятся диалоговые сообщения системы (в том числе результат прогнозирования).

Диалог ввода организован при помощи консоли. Через консоль осуществляется ввод имён классов образов и выводится служебная информация.

Подключение классификатора

Классификатором называется модуль системы, который осуществляет захват лица человека в кадре видеоданных. Процедура выделения фрагмента изображения с лицом в кадре выполняется циклически в процессе вывода видеоданных в окно «Raw Stream». Выделенный фрагмент изображения транслируется в окно «Face». Именно изображение окна «Face» является исходными данными для модуля принятия решений, то есть классификатор осуществляет первоначальную подготовку данных.

В OpenCV имеется готовая процедура, использующая каскады Хаара для захвата и ведения объекта в кадре видеоданных. Объявление каскада Хаара в OpenCV:

```
CascadeClassifier cascade;
```

Для загрузки файла каскада используется процедура:

```
cascade.load( cascadeName );
```

Каскад Хаара – это набор примитивов, для которых считается их свёртка с изображением. Каскады из признаков – это база для построения систем выделения сложных объектов, таких как лица, руки, или другие предметы. Для создания каскадов используются простые примитивы, состоящие из прямоугольников и имеющие всего два уровня, +1 и -1. Под свёрткой подразумевается:

$$S = X - Y,$$

где Y — сумма элементов изображения в тёмной области, а X — сумма элементов изображения в светлой области (при разных масштабах используется свертка $S = X/Y$). Чем больше используется примитивов, тем точнее можно классифицировать объект.

Примитивы Хаара позволяют получить искомый образ на исходном изображении. Для этого необходимо иметь свёртки искомого образа с набором Хаар-признаков, которые сравниваются с набором свёрток тех же примитивов, рассчитанных на исходном изображении в областях, пропорциональных небольшому фрагменту искомого образа.

Таким образом, файл каскада Хаара, который содержится в архиве OpenCV и который мы подключаем к нашему объекту «CascadeClassifier», - это xml-файл, содержащий различные свёртки для определенной выборки примитивов.

Для нахождения лица в кадре в системе используется два каскада Хаара:

- haarcascade_frontalface_alt2.xml
- haarcascade_eye_tree_eyeglasses.xml

Первый файл – для определения лица в кадре, второй файл – для уточнения (фильтрации) лиц по глазам или очкам (сцепленный каскад).

Модуль принятия решений

В настоящее время определена следующая классификация методов распознавания образов [4]:

- 1) Методы теории принятия решений
 - сопоставление образов
 - статистические оптимальные классификаторы
 - нейронные сети
- 2) Структурные методы распознавания
 - сопоставление номеров фигур
 - сопоставление строк символов

В первом классе методы основываются на количественных характеристиках и по большей части игнорируют структурные связи, присущие форме объекта. Во втором классе методы нацелены на использование структурных связей образов (например, формы границ областей), принадлежащих одному классу.

В разработанной системе распознавания эмоций используется алгоритм из класса методов принятия решений через сопоставление образов под названием – Eigenfaces.

Eigenfaces - это название набора собственных векторов, которые используются в задаче компьютерного зрения распознавания человеческого лица [5]. Подход с использованием векторов Eigenfaces для распознавания объектов был разработан Sirovich L. и Kirby M. в 1987, и позже использовался Turk M. и Pentland A. для классификации лиц [6]. Собственные векторы являются производными от ковариационной матрицы

распределения вероятностей по векторам высокой размерности пространства изображений лица. Вектора Eigenfaces образуют базисный набор всех изображений, используемых для построения ковариационной матрицы. Это приводит к сокращению размерности, позволяя представить обучающие выборки изображений меньшим набором базисных изображений. Классификация достигается путем сравнения анализируемого изображения и представленного базиса.

Набор векторов Eigenfaces может быть синтезирован на наборе изображений различных лиц благодаря математическому преобразованию, именуемому как Метод Главных Компонент – один из основных способов уменьшить размерность данных, потеряв наименьшее количество информации.

Объявление модели для создания Eigenfaces векторов:

```
Ptr<FaceRecognizer> model = createFisherFaceRecognizer();
```

Для обучения модели (создания базиса по набору обучающих изображений) необходимо в качестве аргументов передать вектора, содержащие путь к примитивам (*images*) и соответствующие метки классов образов (*labels*):

```
model->train(images, labels);
```

Вектора *images* и *labels* загружаются из предварительно созданного файла содержимого базы данных. Пример файла представлен на рис.3.

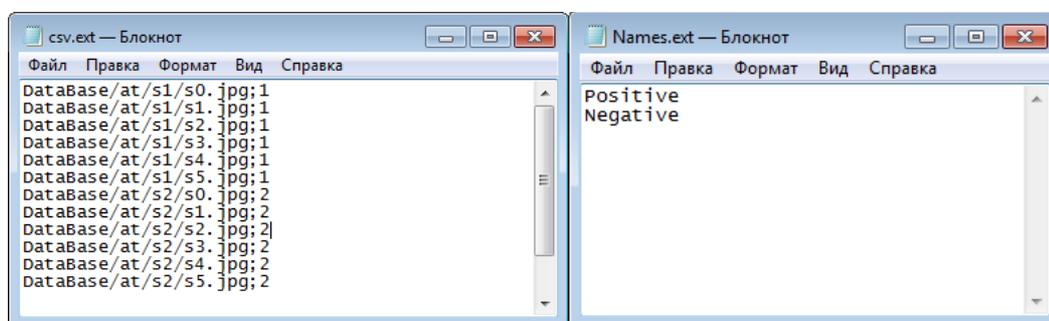


Рис. 3. Пример файла содержимого базы данных (слева). Имена классов (справа)

На рис.3 можно увидеть, что база данных содержит всего 2 класса образов, в каждом из которых 6 файлов примитивов.

Поскольку модель Eigenfaces для именованя классов может использовать только короткие числовые метки, используется еще один файл *names.ext*, который ставит в соответствие каждому классу образов некоторое имя.

Предсказание для анализируемого изображения выполняется командой:

```
int prediction = model->predict(face_resized);
```

Переменная prediction содержит метку класса образов, к которому, возможно, принадлежит анализируемый образ.

Запуск предсказания (распознавания) осуществляется по нажатию клавиши «г».

Управление базой данных примитивов

Управление базой данных примитивов осуществляется во время работы системы при помощи клавиатуры - нажатие клавиши «т».

На рис.4 представлен диалог управления и редактирования базы данных примитивов.

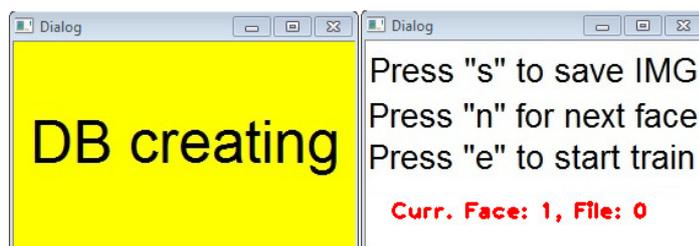


Рис. 4. Диалог управления и редактирования базы данных примитивов

По нажатию клавиши «s» сохраняются примитивы для одного класса.

По нажатию клавиши «n» запускается приглашение ввода имени сохраненного ранее класса образов и создается папка для нового класса, продолжается сохранение файлов для нового класса образов по нажатию клавиши «s».

По нажатию клавиши «e» осуществляется контроль целостности классов образов и происходит выход из режима редактирования базы данных примитивов, а затем автоматически запускается переобучение модели Eigenfaces.

Апробация распознавания эмоций

Предварительно создана база данных примитивов для 3-х классов образов:

- Положительные эмоции (позитивное лицо - Positive)
- Отсутствие эмоций (обычное лицо - Ordinary)
- Отрицательные эмоции (негативное лицо - Negative)

В каждом классе по 6 примитивов. Набор изображений для формирования базиса векторов Eigenfaces представлен на рис.5.

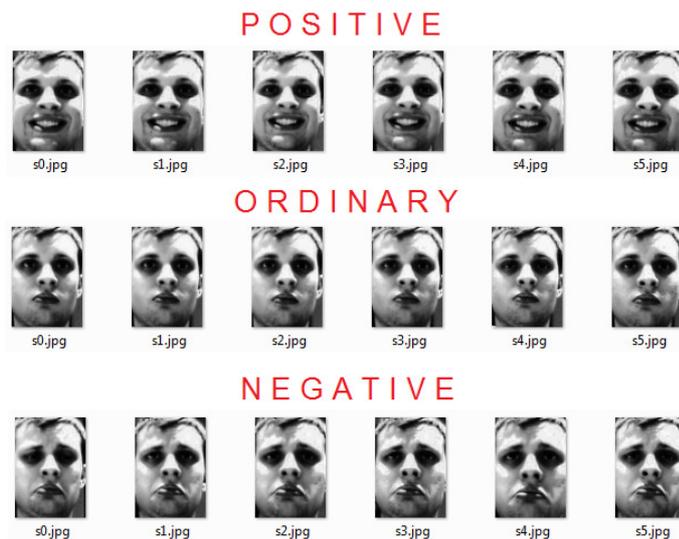


Рис. 5. База данных примитивов для распознавания эмоций

На рис.6 приведены результаты апробации распознавания 3-х классов эмоций, определённых примитивами, представленными на рис. 5.

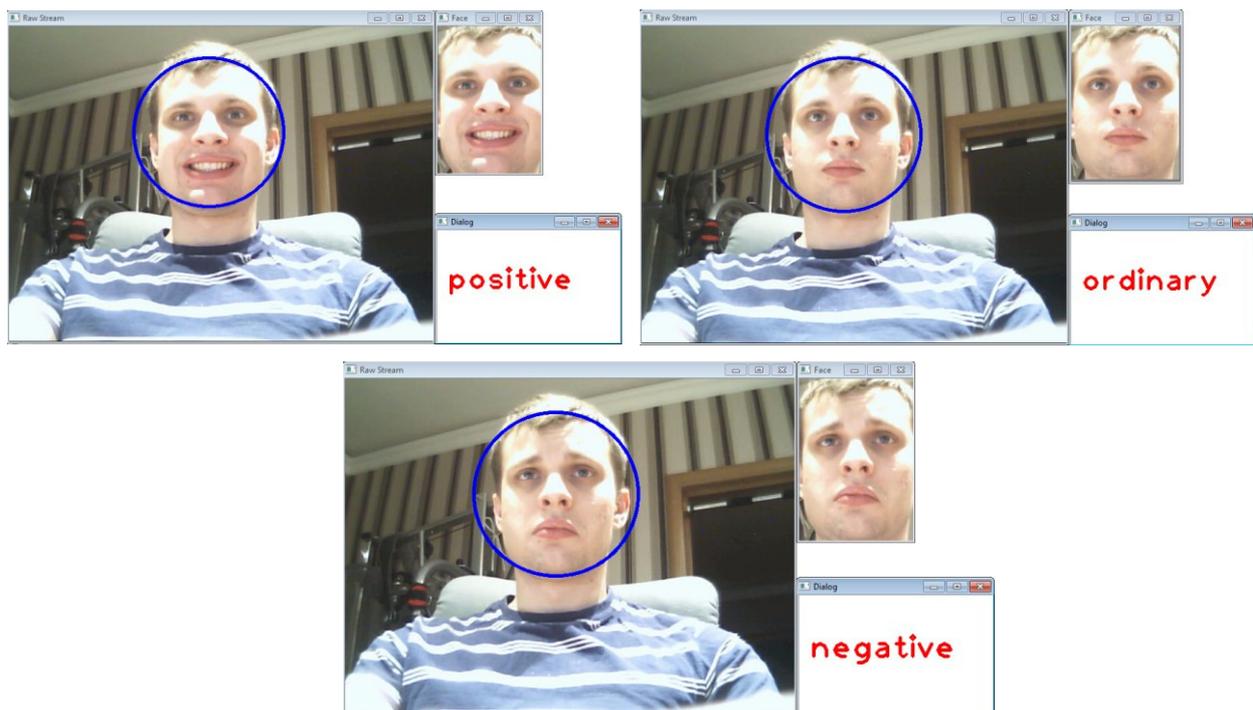
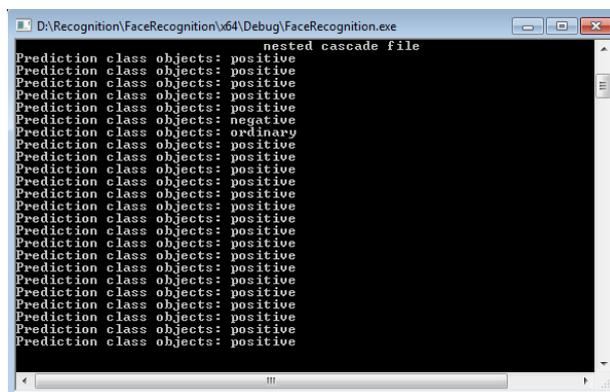


Рис. 6. Результаты апробации распознавания эмоций

Погрешность распознавания

Для проведения исследований погрешности распознавания эмоций выведем в консоль историю распознавания после каждого нажатия клавиши «г». В процессе

проведения опыта осуществлялись повороты и наклоны головы в пределах +/- 30 градусов относительно нормального положения, также изменялась освещенность. История распознавания только положительных эмоций приведена на рис. 7.



```
D:\Recognition\FaceRecognition\vs64\Debug\FaceRecognition.exe
nested cascade file
Prediction class objects: positive
Prediction class objects: negative
Prediction class objects: ordinary
Prediction class objects: positive
```

Рис. 7. Множественное распознавание положительных эмоций

Можно утверждать, что система показывает положительный результат с высокой степенью вероятности. Ошибки распознавания возникают в случаях сильного поворота или наклона головы (более 30 градусов). При значительном изменении освещенности помещения, в котором происходит съемка, эффективность системы снижается.

Выводы

В статье проведен анализ возможностей библиотеки OpenCV и описаны основные концепции реализации системы распознавания эмоций человека на видеоданных. Для захвата и ведения лица человека в кадре были использованы каскады Хаара, приведён и разобран алгоритм свёртки искомого образа с набором Хаар-признаков. Для распознавания эмоций использован алгоритм из класса методов принятия решений через сопоставление образов – модель Eigenfaces. Основное преимущество выбранного метода распознавания образов – низкие требования к производительности системы. Предсказание класса объектов для анализируемых данных происходит за тысячные доли секунды.

Выявленные проблемы:

- 1) Погрешность распознавания возрастает при значительном наклоне или повороте головы – более 30 градусов относительно нормального положения.
- 2) Погрешность также сильно возрастает при значительном изменении освещенности помещения, в котором происходит съемка.

Перспективные задачи исследования в будущем:

1) Исследование возможности реализации полностью автоматического обучения системы и распознавания образов.

2) Исследование применимости выбранного подхода при распознавании личности и аномального поведения.

Список литературы

1. Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978. 411 с.
2. Ян С. Автоматическое распознавание эмоций пользователя для организации интеллектуального интерфейса // Молодежный научно-технический вестник. 2013. №9. Режим доступа: <http://sntbul.bmstu.ru/doc/616498.html> (дата обращения 20.02.2014).
3. Визильтер Ю.В., Желтов С.Ю. Обработка и анализ цифровых изображений с примерами на LabVIEW. М.: ДМК Пресс, 2007. 1104 с.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2012. 616 с.
5. Sirovich L., Kirby M. Low-dimensional procedure for the characterization of human faces // Journal of the Optical Society of America A. 1987. № 4 (3). Pp. 519–524.
6. Turk M., Pentland A. Face recognition using eigenfaces // Proc. IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, 1991. Pp. 586–591.