

ДЕТЕКТИРОВАНИЕ ЛИЦА НА ВИДЕОФРАГМЕНТЕ *А.С.Мигукин* *Санкт-Петербургский государственный политехнический университет*

В настоящее время для подтверждения личности используют различные пароли или специальные ключи (iButton, RFID, магнитные карты). Однако они не обеспечивают должного уровня информационной безопасности, так как могут быть потеряны или украдены. Для повышения информационной защиты используются биометрические системы аутентификации пользователя. Одним из методов, применяемых в таких системах, может быть отпечаток пальца. Хотя он и является уникальным, постоянным признаком, но, как и другие биометрические показатели, уязвим к подделкам (например, можно использовать жирный след отпечатка на сенсоре). Алгоритмы обработки изображения могут обеспечить высокую точность распознавания личности по отпечатку, инвариантность к повороту изображения, изменению масштаба и прочее (см. [1], сс. 86-88), но не помогут зафиксировать подлог. Поэтому для устранения этого недостатка удобно использовать комплексные системы, включающие в себя еще и систему видеонаблюдения.

Пусть камеры видеонаблюдения записывают данные в среднем качестве: изображение градации серого 320x240 точек (средний объем необходимого места на жестком диске на один кадр равен $C_1=15$ Кб/кадр). Ясно, что при записи с N ($N=100$) камер со стандартной телевизионной частотой F ($F=25$ кадр/с) за месяц ($t=2678400$ с) можно получить огромный объем информации ($C_{\Sigma}=t*N*C_1*F\approx 100,5$ Тб). Сократить поток данных, поступающий на запись в ПЗУ автоматической системы видеонаблюдения при неизменном числе видеокамер, можно двумя способами:

1. Уменьшить частоту съемки (основной используемый метод): для того чтобы объем входящей информации помещался на средний объем винчестера (100Гб) нужно снимать с частотой $F = C_{\Sigma}/(t*N*C_1) \approx 2...4$ (кадр/с).
2. Сохранять только полезную информацию. В этом случае удобно детектировать лица посетителей на видеофрагментах и записывать в базу данных только часть изображения.

Данная работа посвящена решению задачи детектирования лица на отдельном видеофрагменте. Это позволяет не только выделить интересующую нас информацию, но и значительно сократить частоту съемки. Даже при частоте F за время прохода коридора перед турникетом L ($L=5$ м) со скоростью V ($V=3...5$ км/ч) можно получить $L*F/V=7...24$ кадров, а небольшое увеличение частоты съемки только увеличивает количество кадров для детектирования, а значит и вероятность выделения лица посетителя.

Система детектирования лица на видеофрагменте моделировалась в программном пакете MATLAB (7 и 2006а). Рассматривается частный случай: детектирование лиц европеоидов. При разработке системы детектирования лица на видеофрагменте источниками изображения являлись данные с интернет-камеры и из баз данных компании BioID (www.bioid.com). В основе разработанного алгоритма лежат методы поиска по геометрическим характеристикам лица и сравнения найденных областей с эталонами.

Для снижения влияния шума, повышения точности детектирования и уменьшения размерности задачи производится начальная обработка в следующей последовательности:

1. Предварительная обработка, а именно: унификация изображения (обрабатывается только изображение градации серого, что позволяет одинаковым образом обрабатывать видеофрагменты с любых камер, в том числе данные полученные в полной темноте в инфракрасном спектре), гамма-коррекция и эквализация гистограммы;
2. Фильтрация изображения (гомоморфная фильтрация для выделения затемненных областей и высокочастотная фильтрация с помощью лапласиана гауссиана);
3. Сегментация (исходя из специфики цвета лица белого человека последовательное выделение сначала более светлых областей предполагаемого лица, затем на выделенных областях - более темные зоны черт лица таких, как глаза, рот, нос);
4. Бинаризация (перевод изображения градации серого в черно-белое с использованием оптимальной пороговой обработки Отсу, где белыми точками обозначаются предполагаемые области черт лица).

Затем производится выделение характерных опорных восьмисвязных (из каждой белой точки можно перейти в любую соседнюю) областей исходя из следующих эмпирически полученных ограничений:

1. Область должна быть достаточно большой (число точек $N>20$), но ее допустимый размер также ограничен сверху ($N<250$);
2. Область должна быть достаточно протяженная (область не должна вырождаться в тонкую линию, то есть минимальная A_{min} длина оси описанного вокруг области эллипса должна быть $A_{min}>3$, а эксцентриситет $E>0,25$);
3. Центр области должен лежать внутри выделенного объекта (нет жесткого требования выпуклости).

Определив характерные области, легко можно выделить предполагаемые области лица по разработанному правилу геометрических взаимосвязей. Поиск осуществляется с учетом геометрических фигур (треугольников), которые можно построить, взяв за вершины центры областей глаз (точки А и В) и носа или рта (точка С), как это показано на рисунке 1 ниже.

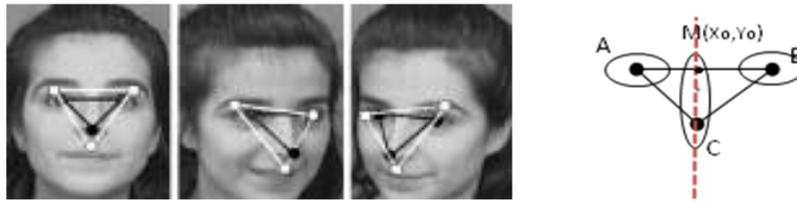


Рис. 1. Пример геометрических фигур, используемых для поиска областей черт лица

Координаты центра области поиска можно найти геометрически. Для любой точки $A(x_1, y_1)$ найдется пара (точка B) исходя из ограничения выбора характерных областей. Тогда точка C лежит на прямой, перпендикулярной к отрезку AB и проходящей через точку $M(x_0, y_0)$: $y_c = -1/k_1 * (x_c + x_0) - y_0$, где k_1 – тангенс угла наклона отрезка AB. Координаты точки C вычисляются как: $x_c = x_0 - d * tg(\alpha) * \cos(90 - k_1)$, при $k_1 \in (-90...0)^\circ$, $x_c = x_0 + d * tg(\alpha) * \cos(90 - k_1)$, при $k_1 \in (0...90)^\circ$, $y_c = y_0 - const$ $x_c = x_0 \pm |y_0 - y_1| * tg(\alpha)$, при $k_1 = \pm 90^\circ$, $x_c = x_0 - const$ $y_c = y_0 - |x_0 - x_1| * tg(\alpha)$, при $k_1 = 0^\circ$, где α – известный из геометрии лица угол BAC ($[30...72]^\circ$), d – расстояние между точками A и M.

В качестве области поиска берется прямоугольник, размер которого выбирается с запасом так, чтобы захватить на изображении рот, нос и брови. Для решения задачи детектирования области лица (кластеризации лиц и «нелиц») используется нейронная сеть двухслойная нейронная сеть прямого распространения с 5 нейронами внутреннего слоя (число определено эмпирически из контроля за «переобученностью» сети), обучаемая по процедуре обратного распространения ошибки по алгоритму Левенберга-Маквардта. В качестве вектор-признаков для обучения нейронной сети применялись неиспользуемые ранее для решения подобных задач инвариантные моменты или «кумулянты» (см. также [1], сс. 86-88).

Ниже, на рисунке 2, приведены некоторые примеры детектирования лиц на изображениях.

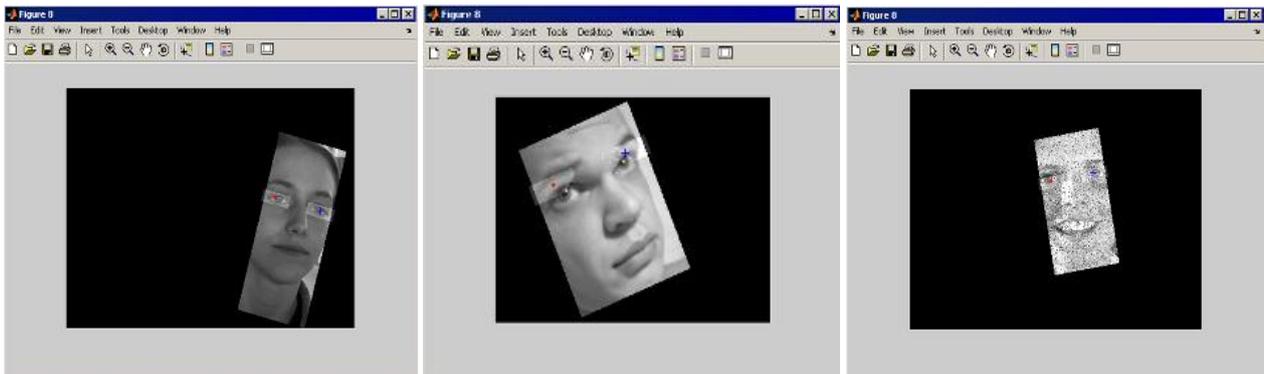


Рис. 2. Детектирование лиц, повернутых на некоторый радиальный и азимутальный угол, а также детектирование лица на изображении, зашумленном негауссовым («розовым») шумом

По результатам моделирования (в ходе работы использовалось 200 цифровых изображений по 100 для обучения и валидации нейронной сети), разработанная система детектирования лица на видеофрагменте с использованием кумулянтов показала:

- корректное детектирование лица для зашумленных (в определенных пределах) изображений;
- высокую робастность к гауссовому и негауссовому шуму;
- инвариантность к радиальному повороту лица в диапазоне $[-75...75]$ градусов;
- инвариантность к азимутальному повороту лица в диапазоне $[-40...40]$ градусов;
- высокая скорость вычисления (оценочная скорость работы алгоритма порядка 0.8...1 секунды);
- ошибка первого рода FRR (количество неправильно детектированных областей лица) составляет около 7% , а для видео потока 3 кадра в секунду – 5%;
- ошибка второго рода FAR (избыточность, ложные области лица) составляет около 10%.

По точности детектирования разработанная система в целом не хуже аналогичных систем наблюдения, а применение данного алгоритма дает экономию в пространстве на винчестере цифрового видеоманитора порядка 30...60% (в зависимости от расположения камер, качества съемки, размера изображения и прочее).

ЛИТЕРАТУРА

1. XXXV Неделя науки СПбГУ: Материалы Всероссийской межвузовской научно-технической конференции студентов и аспирантов. Ч. V. СПб.: Изд-во Политехн. ун-та, 2007. 207 с.