

метно затормаживающий реакции компьютера, этой цели не соответствует. Поэтому большинство методов, использующихся для решения задачи обнаружения лица, принадлежат к категории эмпирического распознавания, как предоставляющей оптимальный для данных задач компромисс между скоростью работы, простотой использования и качеством распознавания.

1.2. Обнаружения лица с помощью цветовой сегментации

Среди эмпирических методов обнаружения лица получили широкое распространение методы, использующие цветовую сегментацию областей кожи. Причина этой популярности состоит в том, что использование цвета кожи как признака для обнаружения лица сочетает в себе несколько важных преимуществ:

- Малая вычислительная сложность и, как следствие, высокая скорость обработки;
- Устойчивость к изменению ориентации и масштаба лица;
- Устойчивость к изменению освещения (за исключением цветного);
- Устойчивость к изменению выражения лица и частичного перекрытия лица другим объектом сцены.

К недостаткам этого метода можно отнести:

- Потенциальная возможность ложного обнаружения предметов с цветом, близким к цвету кожи;
- Зависимость от цветового баланса камеры и цвета освещения.

Вышеперечисленные достоинства обнаружения лиц по цвету делают использование характерного цвета кожи как признака для распознавания весьма привлекательным, даже несмотря на имеющиеся недостатки. Учитывая высокую скорость обработки цветовой информации, процедура обнаружения лиц с помощью цвета кожи часто используется как первый шаг в обнаружении лиц для сужения области поиска лиц на изображении, перед тем как применяется более сложный алгоритм анализа изображения. При таком использовании этого метода относительно высокая вероятность ложного обнаружения не так страшна.

Обнаружение лиц по цвету кожи производится обычно в два этапа:

1. выделение пикселей, близких по цвету к коже;
2. поиск лиц на изображении с использованием априорного знания о структуре человеческого лица.

От того, насколько точно будут определены на изображении области, относящиеся к коже, в значительной мере зависит точность конечного результата.

1.2.1. Цветовая сегментация областей кожи на изображении

Цветовая сегментация является распространенной техникой, применяемой в различных задачах анализа изображений. Сегментация является предварительным этапом функционирования любой системы обработки изображений, так как позволяет упростить последующий анализ однородных областей изображения, их яркостных, цветовых и геометрических характеристик. Сегментацию можно рассматривать как начальный этап



Рис. 1.1. Типичные примеры цветовой сегментации кожи.

построения формального описания сцены, качество выполнения которого во многом определяет успех решения задачи распознавания изображений, интерпретации и идентификации объектов сцены. Цветовая сегментация означает разделение изображения на области по цветовым признакам. Конкретная техника сегментации сильно различается от задачи к задаче - в некоторых случаях требуется разделить изображение на области однородного (в смысле некоторой метрики) цвета, в других же нужно найти области с определенным, известным заранее цветом (как при сегментации кожи).

Цвет, как известно, есть не физическое свойство объекта, а свойство человеческого восприятия, поэтому строгого математического определения

понятия “цвет кожи” не существует, что представляет собой некоторую трудность при построении систем по автоматического распознавания кожи. В компьютерном представлении изображения цвет пикселя задается в качестве координат в некотором цветовом пространстве. Психологи установили, что в глазе человека присутствует три типа светочувствительных клеток, настроенных на различные распределения энергии по видимому спектру. Исходя из этого факта любой цвет, воспринимаемый человеческим глазом, можно представить в виде трехмерного вектора. Подробнее о различных цветовых пространствах и о природе и свойствах цветового восприятия можно прочитать в /59/, /2/.

Результатом цветовой сегментации кожи должен быть набор областей изображения, близких по цвету к человеческой коже. Обнаружение таких областей дает дополнительный признак для локализации лица на изображении. Еще в ранних работах по автоматическому обнаружения областей кожи было замечено, что цвет кожи занимает достаточно компактную область в большинстве цветовых пространств, и что даже достаточно простым набором неравенств /34/ можно отделить область цвета кожи от остального пространства достаточно эффективно, чтобы использовать это разделение для обнаружения кожи на изображениях.

Поскольку, как уже было сказано, цвет объекта есть свойство восприятия человека, то единственным источником исходных данных для конструирования и обучения алгоритма автоматического обнаружения областей кожи по цвету является набор тренировочных изображений, на каждом из которых вручную выделяются области, относящиеся к коже.

Фактически, можно сформулировать задачу обнаружения кожи по

цвету как задачу построения классификатора (для случая двух классов) в пространстве признаков, совпадающим с цветовым пространством. Данными для тренировки классификатора является набор векторов цветового пространства, для которых известна их принадлежность к одному из классов (кожа и “не-кожа”). В процессе построения цветового классификатора приходится решать следующие вопросы:

Выбор признакового пространства

Колориметрия, компьютерная графика и стандарты сжатия и передачи видеосигнала дали жизнь большому количеству различных цветовых пространств. Выбор признакового пространства может значительно сказаться на характеристиках использующего его классификатора. Значительная часть существующих цветовых пространств была применена к задаче распознавания кожи в надежде найти пространство, позволяющее получить классификатор с наилучшими характеристиками. Авторы экспериментировали с различными пространствами в надежде добиться максимальной компактности и связности области, соответствующей оттенку кожи в данном пространстве, и получить хорошую разделяемость цвета кожи и не-кожи. Нужно сказать, что часто эксперименты с выбором цветового пространства носили достаточно бессистемный характер, и лишь относительно небольшое число авторов приводило в своих работах строгие обоснования справедливости выбора используемого пространства (/87/, /86/, /54/, /68/).

Как показали дальнейшие исследования /69/, /15/, выбор цветового пространства является важным далеко не для всех классификаторов. Су-

существует ряд способов моделирования распределения кожи в цветовом пространстве, нечувствительных к преобразованиям в другие цветовые пространства.

Помимо улучшений условий распознавания, велся поиск пространства, которое снизило бы зависимость кожного классификатора от условий освещения. Частичную независимость дает использование пространств, явно разделяющих оттеночную (chrominance) и яркостную (luminance) компоненту цвета. Если отбросить яркостную компоненту, это даст независимость классификации кожи от яркости освещения сцены, но несколько ухудшит разделяемость цвета кожи и “не-кожи”. Подробнее о проблеме выбора цветового пространства можно прочитать в работе автора /84/, а также в /95/, /23/, /20/, /50/.

Выбор способа моделирования цвета кожи

Целью создания цветового классификатора кожи является построение решающего правила (или набора правил), разделяющего цвета на близкие к цвету коже и нет. Обычно, задача решается путем создания модели распределения цвета кожи в выбранном цветовом пространстве и проверки цвета обрабатываемого пикселя на удовлетворение полученной модели. Источником данных для построения модели является набор тренировочных изображений с выделенными вручную областями кожи.

Существующие методы моделирования оттенка кожи можно разделить на три семейства /84/:

1. Набор явно заданных условий (неравенств), налагаемых на координаты векторов в цветовом пространстве;

2. Методы непараметрического моделирования;
3. Параметрические методы моделирования.

На основании сравнения существующих методов цветовой сегментации кожи был выбран метод моделирования распределения с помощью Байесовской карты вероятностей. Он заключается в вычислении отношения $\frac{P(\text{skin}|c)}{P(\neg\text{skin}|c)}$ (соотношение вероятностей принадлежности цвета с координатами $[c]$ к оттенку кожи или “не-кожи”) для каждой точки дискретизированного цветового пространства по следующим формулам:

$$\frac{P(\text{skin}|c)}{P(\neg\text{skin}|c)} = \frac{P(c|\text{skin})P(\text{skin})}{P(c|\neg\text{skin})P(\neg\text{skin})} \quad (1.1)$$

и последующего сравнения $\frac{P(\text{skin}|c)}{P(\neg\text{skin}|c)}$ с некоторым выбранным порогом Θ . В (1.1) $P(c|\text{skin})$ и $P(c|\neg\text{skin})$ напрямую высчитываются из набора тренировочных изображений, с областями кожи, сегментированными вручную. Априорные вероятности $P(\text{skin})$ и $P(\neg\text{skin})$ также можно вычислить, исходя из тренировочного набора. Однако, как легко показать, их можно принять равными 0.5, поскольку изменение априорных вероятностей влияет лишь на величину порога, используемого для обнаружения кожи, не изменяя соотношения процента ложных обнаружений и ложных промахов. Данная модель обладает высокими показателями вероятности успешного распознавания, малой вероятностью ложного обнаружения, и является инвариантной по отношению к выбору цветового пространства, в котором проводится сегментация.

1.2.2. Обработка результатов цветовой сегментации для обнаружения лица

Результатом цветовой сегментации кожи является информация о степени близости цвета каждого пикселя к цвету кожи. На основе этой информации обычно формируется полутоновое изображение, где интенсивности пикселей установлены равными характеристике близости пикселя к цвету кожи. Изображение может быть бинарным, отмечающим принадлежания к цвету кожи пиксели, или отражать вероятность принадлежности каждого пикселя к коже (так называемая карта вероятности кожи, skin probability image).

К сожалению, цветовая информация не может дать достаточно данных для устойчивого обнаружения лица. Неизбежные ошибки цветовой сегментации и области, близкие по оттенку к коже, не являющиеся лицами, создают помехи для распознавания. Для повышения устойчивости требуется дополнить набор признаков дополнительными свойствами изображения лица. Рассмотрим наиболее часто встречающиеся признаки, используемые в методах, основанных на цветовой сегментации кожи:

Форма лица

Лицо имеет характерную форму овала (эллипса). Можно подсчитать разумные ограничения на пропорции лица, справедливые для большинства лиц, и использовать это как критерий $/75/$, $/23/$, $/88/$. Обычно используются величины вблизи “золотого сечения”, которое задает, как считается, идеальные пропорции лица.

$$\frac{height}{width} \equiv \frac{1 + \sqrt{5}}{2} \approx 1,61803 \quad (1.2)$$

Движение в кадре

Часто, выделение кожи по цвету используется для решения задачи по отслеживанию перемещения лица между кадрами. При обработке видео в реальном времени особенно важна скорость методов распознавания, поэтому цветовая сегментация кожи находит здесь широкое применение. Обрабатывая последовательность кадров видеопотока, мы можем получить больше данных, чем путем анализа одного статического изображения. В частности, можно определить области изображения, изменяющиеся между кадрами (не являющиеся статичными), и использовать эту информацию, как дополнительный признак присутствия лица. Действительно, часто фоновая часть кадра в задачах отслеживания лица является слабо меняющейся, в то время как лицо подвижно и изменчиво /17/.

Наличие признаков характерных черт лица

Признаки, свидетельствующие о нахождении характерных черт лица внутри потенциальной области лица, могут также служить дополнительным признаком для обнаружения лица. Проверки на их наличие могут быть как очень простыми, например, проверка на наличие темных областей или пикселей с низкой кожной вероятностью, или более изощренными - моделирование распределения яркости изображения лица с помощью метода главных компонент - /75/, /88/, /22/, /57/.

В работе /57/ авторы используют информацию, предоставленную распределением цвета кожи для формирования черно-белого “изображения вероятности”. К такому изображению применяется метод главных компо-

нент для нахождения лица на нем.

Комбинированная классификация

В некоторых системах цветовая информация используется совместно с информацией, полученной от других источников. Так, например, в /17/, при отслеживании лица на последовательности кадров лицо находится, как пересечение области цвета кожи и движущегося поля. В /30/ цветовая сегментация используется совместно с нейросетевыми и стерео детекторами.

Шаблон изображения лица

В /22/ авторы используют модель цвета кожи и цвета волос. С помощью гистограмм определяется принадлежность пикселей разным классам - волосы, лицо, волосы/лицо, волосы/фон, и с помощью нечеткой логики производится разделение изображения на области, соответствующие коже и волосам. Затем используется модель изображения головы (5 моделей с различной ориентацией лица), в которой задано расположение областей кожи и волос для определения возможных лиц на изображении. Проверка лиц производится путем поиска внутри областей черт лица (с помощью обнаружения резких изменений яркости).

Форма лица как дополнительный признак распознавания

Из всех перечисленных ограничений, ограничения на форму используются чаще других, поскольку эта проверка, с одной стороны, достаточно проста в реализации, а с другой позволяет отместить большое количество областей, с высокой вероятностью не являющихся лицами, значительно уменьшив количество областей изображения, к которым стоит применять более изощренные способы проверки. В литературе используются, в ос-

новном, подходы, которые сначала группируют пиксели с цветом кожи в области путем применения различных алгоритмов обработки изображений (фильтров, операторов, наращивания областей). После чего найденные области проверяются на удовлетворение ограничениям по форме /75/, /23/, /88/ - потенциальными лицами считаются обычно эллиптические области с определенными пропорциями между осями эллипса. Иногда оценивается также ориентация главной оси области, когда ищутся лица, расположенные определенным образом. Наиболее распространенными способами объединения пикселей кожи в потенциальные области лиц являются метод интегральных проекций /45/ и метод выделения связных компонент.

Интегральные проекции используются для вычисления прямоугольников, ограничивающих области скопления пикселей с высокой кожной вероятностью. Вертикальная интегральная проекция столбца x изображения I задается следующим образом:

$$proj_{ver}(x) = \sum_{y=0}^N I_{skin}(x, y) \quad (1.3)$$

Аналогично задается горизонтальная проекция строки y :

$$proj_{hor}(y) = \sum_{x=0}^N I_{skin}(x, y) \quad (1.4)$$

Строки и столбцы, для которых величины интегральных проекций велики, с высокой вероятностью содержат скопление пикселей с высокой кожной вероятностью. Координаты, соответствующие областям резкого убывания интегральных проекций, относятся к границам областей пикселей кожи (см. рис 1.2). Результатом анализа интегральных проекций изображения кожной вероятности являются ограничивающие прямоугольники скопления пикселей с высокой кожной вероятностью.

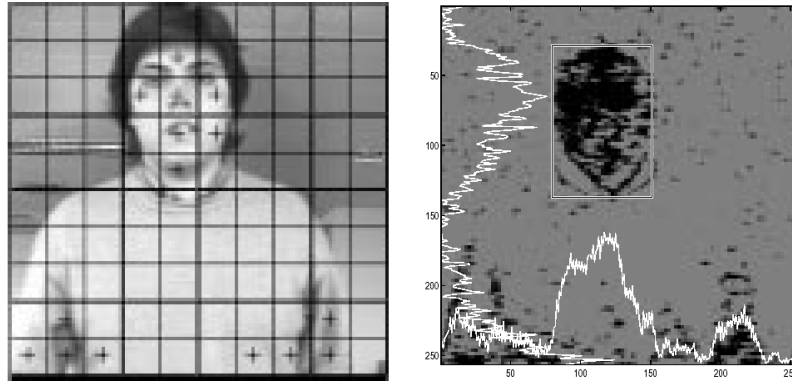


Рис. 1.2. Использование интегральных проекций для поиска лица по цвету (иллюстрация из работы /13/).

Получив бинарное изображение карты кожной вероятности (например, путем пороговой фильтрации), выделение связанных компонент пикселей кожи дает набор областей, часть из которых, как ожидается, является лицами. Конструктивным определением связанной компоненты является множество пикселей, в окрестности каждого из которых присутствует как минимум один пиксель, также принадлежащий данной области. От определение понятия окрестности пикселя зависит тип связанных областей. Используются два типа окрестностей: четырехсвязная и восьмисвязная (эвклидова):

0	1	0	1	2	3
2	x	3	4	x	5
0	4	0	6	7	8

Соседи пикселей

Обычно используются четырехсвязные компоненты. Выделение связанных компонент производится методом последовательного сканирования. Метод заключается в сканировании изображения слева направо и сверху вниз, и выполнения для каждого пикселя (x, y) следующих действий:

1. если пиксель (x, y) помечен единицей на исходном бинарном изображении:
2. если пиксели $(x - 1, y)$, $(x, y - 1)$ не отнесены ни к одной из
3. найденных областей:
4. прибавить счетчик областей $n = n + 1$;
5. приписать пиксель (x, y) к области n ;
6. если только один из пикселей $(x - 1, y)$, $(x, y - 1)$ отнесен к одной из
7. областей i :
8. приписать пиксель (x, y) к области i ;
9. если оба пикселя $(x - 1, y)$, $(x, y - 1)$ отнесены к найденной
10. области i :
11. приписать пиксель (x, y) к области i ;
12. если оба пикселя $(x - 1, y)$, $(x, y - 1)$ отнесены к разным
13. областям i, j : $i \neq j$:
14. приписать пиксель (x, y) к области i ;
15. зафиксировать эквивалентность областей i и j ;
16. перейти к следующему пикселю.

После проведения сканирования производится разрешение эквивалентностей областей и переразметка изображения с учетом этих эквивалентностей. После того, как на изображении обнаружены связные области пикселей, предположительно являющиеся лицами, производится проверка формы этих областей на удовлетворение определенным ограничениям: пропорции, соотношение периметра и площади, эксцентриситет и другие. Пример обнаружения лица по цвету с помощью анализа связных компонент представлен на рисунке 1.3.

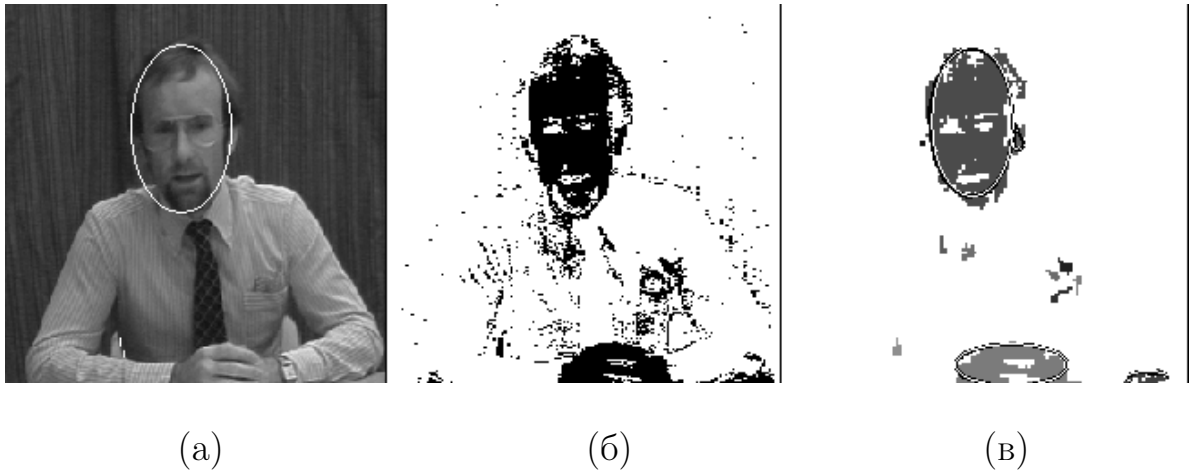


Рис. 1.3. Использование анализа связных компонент для поиска лица по цвету (иллюстрация из работы /75/). (а) - исходное изображение с отмеченным результатом обнаружения, (б) - карта пикселей цвета кожи, (в) - выделенные связные области цвета кожи с эллиптической формой

1.2.3. Несовершенство существующих методов

Важными свойствами перечисленных методов выделения областей кожи является простота реализации и малая вычислительная сложность. Поскольку быстрота - одно из важных преимуществ использования цвета для обнаружения лица, важно не потерять это преимущество из-за чересчур сложной постобработки цветовой информации. Основным (и очень серьезным) недостатком этих методов является то, что они способны обнаружить и сегментировать от окружения лишь хорошо локализованные, четко отделенные области. Очень легко привести пример, когда метод интегральных проекций и выделения связных областей не дает приемлемых результатов даже на простых изображениях. Примеры таких изображений приведены на рисунках 1.5, 1.4.

На рис 1.5 анализ интегральных проекций дает ошибочные ограни-



Рис. 1.4. Примеры ошибочной работы метода основанного на анализе связанных компонент

чивающие прямоугольник лица, на рис 1.4 приведен ошибочный результат применения метода выделения связанных компонент. Причина таких сложностей - примитивность признаков, по которым методы выделяют регионы потенциальных лиц. Для того, чтобы улучшить устойчивость методов, требуется сделать их более “интеллектуальными”, научить их выделять и отделять области, близкие по форме к возможной форме регионов лиц, сохранив, по возможности, простоту и скорость работы методов.

У самой идеи разделения шагов группировки пикселей в области и проверки областей на удовлетворение ограничениям есть существенный недостаток. Дело в том, что описанные методы, руководствуясь заложенными в них простейшими правилами, могут ошибочно сгруппировать в одну область пиксели, реально относящиеся к разным областям. В результате такой ошибочной группировки могут появиться области, не удовлетворяющие условиям постпроверки, однако являющиеся лицами. Исследователи

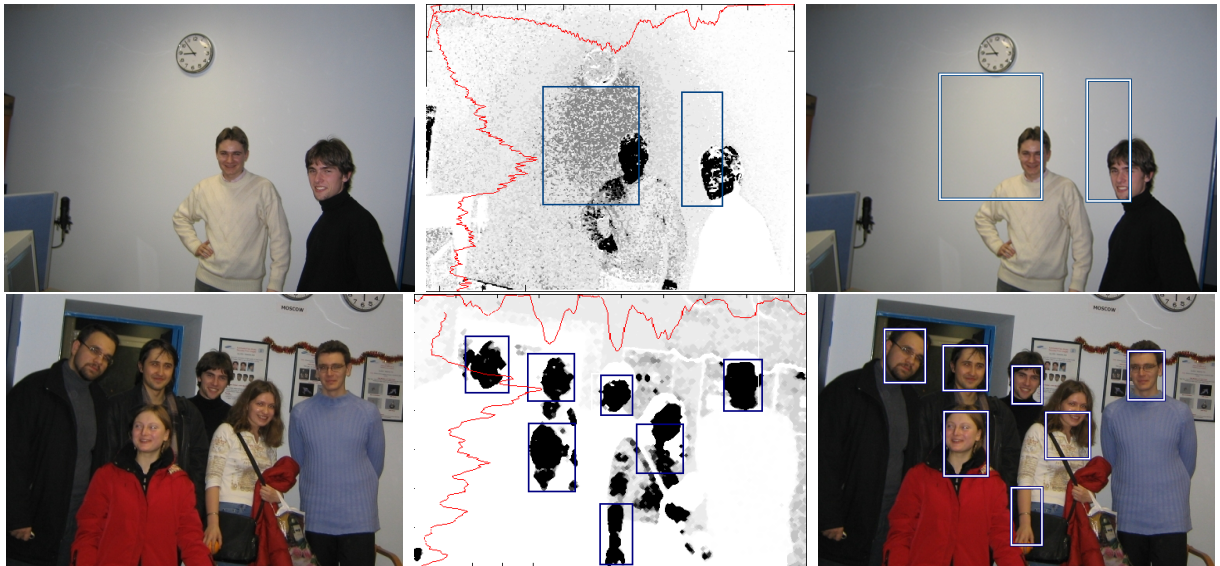


Рис. 1.5. Примеры ошибочной работы метода основанного на интегральных проекциях

осознают ограниченность перечисленных методов, было сделано несколько попыток преодолеть эти ограничения, предложив методы, более устойчивые к различным конфигурациям пикселей кожи.

в /87/ используется многомасштабная сегментация на однородные по цвету области, каждая из которых затем классифицируется как кожа/не кожа (если область более чем на 70% состоит из пикселей, классифицированных как кожа), после этого области объединяются, начиная с самого мелкого масштаба до самого крупного, пока не образуют области, близкие по форме к эллипсу. Такой подход менее чувствителен к ошибкам и неточностям цветовой сегментации, однако разделение на однородные области происходит отдельно и независимо от выделения кожи, что может сказаться отрицательно на финальной точности метода. Также, применяемая авторами процедура выделения однородных областей, достаточно сложна как вычислительно, так и в реализации.

В /76/, в задаче отслеживания перемещения лица в видеопотоке, используется предыдущее положение лица для расчета весов вклада кожных пикселей на текущем кадре при вычислении нового положения лица (Гауссово нормальное распределение). Этот алгоритм позволяет добиться значительно большей устойчивости, нежели использование выделения связных компонент, однако, может использоваться только при работе с видеопотоком при условии знания положения лица на предыдущем кадре и, в ряде случаев (когда в изображении присутствует большое количество пикселей цвета кожи, не относящихся к лицу), может по-прежнему выдать ошибочные результаты.

В /79/ автор попытался расширить метод выделения связных компонент, сделав его менее чувствительным к одиночным пикселям. Его метод выделения называется радиальным расширением (radial spanning) и использует простую деформируемую модель - набор движущихся “щупов” (probes), движение которых рассчитывается, исходя из действующих на них сил. Каждый щуп стремится в сторону от центра инициализации модели до тех пор, пока на его пути не встретится пиксель, классифицированный как не относящийся к коже. Для повышения устойчивости метода, а также для того, чтобы форма выделяемой области была по возможности близка к выпуклому многоугольнику, введена сила влияния щупа на своих соседей. Несмотря на большую, по сравнению с анализом связных компонент, устойчивость, радиальное расширение использует слишком слабую связь между щупами, что зачастую приводит к неудовлетворительным результатам.

Другая деформируемая модель была применена в работе /75/, для

решения задачи отслеживания перемещения лица. Использованная модель также представляла собой набор “щупов” (называемых “snake”, производное от “snake” и “element”), каждый из которых может двигаться перпендикулярно границе модели либо в сторону ее центра (если он не обнаруживает в своей окрестности достаточно пикселей цвета кожи), либо от центра (в противном случае). Взаимосвязь щупов реализуется стандартными силами взаимного влияния элементов деформируемой модели, называемой “змеями” (snakes) /46/. Результаты применения, к сожалению, достаточно скромны, и устойчивость модели и точность выделения оставляют желать лучшего. Основной причиной столь невысоких показателей является слабая взаимосвязь элементов модели и отсутствие ограничений на форму результирующих областей.

1.2.4. Предлагаемые методы обнаружения лиц на изображениях

В связи с популярностью использования цветовой сегментации кожи для обнаружения лица отсутствие методов эффективного использования результатов сегментации осложняет получение высоких результатов задачи обнаружения лиц на изображениях. Существует потребность в методе (наборе методов) выделения потенциальных областей лиц, обладающего следующими свойствами:

1. Устойчивость к неизбежным ошибкам и неточностям цветовой сегментации;
2. Точность выделения областей;
3. Высокая скорость работы;

4. Простота реализации;
5. Использование только результатов цветовой сегментации кожи, без собственного анализа изначального изображения (для использования с различными способами моделирования и выделения цвета кожи).

Задача состоит в выделении эллиптических областей максимального размера, внутри которых плотность содержания количество пикселей цвета кожи максимальна, в то время как на границе количество кожных пикселей минимально. На рисунке 1.6 представлен пример желаемого выделения эллиптической области из неидеальных исходных данных.



Рис. 1.6. Пример выделения эллиптической области из неидеальных исходных данных

Представляется естественным сформулировать эту проблему, как задачу оптимизации в пространстве параметров эллипса (x_c, y_c, θ, a, b) . Однако, получение формулы целевой функции в явном виде не является тривиальной задачей. В работах /79/ и /75/ целевая функция задается неявно, через действующие на элементы деформируемых моделей силы. Попытка формулировки целевой функции в явном виде была предпринята в работе /48/ уже после опубликования автором предлагаемого метода. В

/48/ выделение эллиптических областей оттенка кожи осуществлялось путем максимизации взаимной информации между распределением вероятностей кожи для пикселей изображения и эллиптической функцией плотности пикселей. Для нахождения оптимальных параметров использовался метод градиентного спуска. Пока авторам удалось добиться выделения эллипсов, оси которых совпадают с осями системы координат изображения, снижая таким образом пространство параметров оптимизации. К сожалению, информации об устойчивости и скорости работы метода в статье не предоставлено.

Чтобы избежать трудностей, связанных с поиском явного задания целевой функции и разработки метода оптимизации, не подверженного проблеме локальных экстремумов, было разработано два алгоритма, использующих последовательные приближения с помощью деформируемой эллиптической модели, жестко фиксирующей возможные формы областей лиц.

Деформируемые модели являются мощным средством распознавания и анализа изображений, поскольку предоставляют возможность гибкой формулировки условий и ограничений на форму распознаваемых структур и объектов, а также применения широкого класса методов поиска оптимальной конфигурации модели. Деформируемая модель определяется двумя аспектами - способом задания формы модели и правилами деформации модели. Способ задания формы зависит от ожидаемой структуры объекта распознавания и известных ограничений на возможную изменчивость его формы, правила деформации определяются характеристиками представления объекта на изображении. Часто, однако не всегда, правила

деформации модели задаются путем введения понятия *энергии* модели, являющейся фактически целевой функцией задачи оптимизации параметров модели и задания способа поиска экстремума энергии.

В предлагаемых методах форма модели задается эллипсом произвольного положения, размеров и ориентации (x_c, y_c, θ, a, b) . Энергия модели не задается в явном виде, так как мы хотим избежать сложностей и потенциальных ошибок в формулировке целевой функции, поведение модели регулируется набором простых правил ее деформации. Поступая таким образом, также удастся сделать поведение модели “прозрачным” и легко управляемым путем изменения интуитивно понятных параметров и дать возможность легкого добавления дополнительных ограничений на обнаруживаемые области. Классические деформируемые модели подобного не допускают, поскольку характер их поведения контролируется исключительно выбором формулировки энергии и способа ее минимизации, что может сделать процедуру модификации модели для изменения ее поведения определенным образом достаточно нетривиальной задачей.

Метод эллиптической деформации для обнаружения лица

Вдоль границы модели расположены области (по аналогии называемые щупам), в которых происходит оценка плотности пикселей с высокой кожной вероятностью. В зависимости от этой плотности щупы деформируют модель таким образом, чтобы включить внутрь модели как можно больше кожных пикселей, избегая пикселей с малой кожной вероятностью.

Модель инициализируется вблизи ожидаемого положения лица на изображении (например, в центре масс связной компоненты пикселей кожи

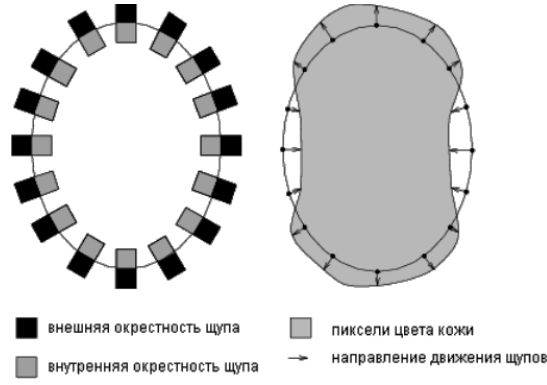


Рис. 1.7. Деформируемая эллиптическая модель для выделения области лица

подходящего размера). Деформация модели производится по шагам. Для каждого щупа производится расчет плотности пикселей кожи во внутренней и внешней окрестности границы модели (см. рис 1.8):

$$P_{in} = 2 \cdot \sum_{x \in \Omega_1} \frac{P_{skin}(x)}{S}; P_{out} = 2 \cdot \sum_{x \in \Omega_2} \frac{P_{skin}(x)}{S}; \quad (1.5)$$

где Ω_1 - внутренняя область щупа, Ω_2 - внешняя область щупа, S - площадь щупа, $P_{skin}(x)$ - вероятность принадлежности цвета пикселя x к цвету кожи. Исходя из рассчитанных величин, вычисляется вектор перемещения щупа:

$$v_i = \begin{cases} -k_{in} \cdot \vec{n}_i, & \text{if } P_{in}^i < T_1, \\ k_{out} \cdot \vec{n}_i, & \text{if } P_{in}^i \geq T_1 \text{ and } P_{out}^i < T_2, \\ 0, & \text{otherwise.} \end{cases} \quad (1.6)$$

В (1.6), i задает номер щупа, \vec{n}_i - единичная нормаль к границе модели в точке центра щупа, направленная от центра модели, P_{in}^i и P_{out}^i плотности пикселей кожи во внутренней и внешней областях щупа, а T_1 и T_2 пороговые величины. Коэффициенты k_{in} и k_{out} (≥ 0) управляют скоро-

стью и характером перемещения щупов. Полный цикл обнаружения потенциального региона лица модели задается следующим алгоритмом:

0. Инициализация модели;
1. Количество итераций $k = 0$;
2. Пока изменение параметров модели $> \varepsilon$ и $k < N$:
3. для каждого щупа $i = \overline{1, n}$:
4. подсчитать P_{in}^i, P_{out}^i ;
5. если $P_{in}^i < T_{in}$ перемещение щупа $\vec{v}_i = -\vec{n} * v_{in}$, иначе:
6. если $P_{out}^i \geq T_{out}$ перемещение щупа $\vec{v}_i = \vec{n} * v_{out}$, иначе:
7. $\vec{v}_i = \theta$;
8. пересчитать координаты центров щупов $x_i^{k+1} = x_i^k + v_i$;
9. аппроксимировать центры щупов $x_i^{k+1}, i = \overline{1, n}$ эллипсом;
10. новые параметры формы модели = параметры полученного эллипса;
11. $k = k + 1$, перейти на 2;

Приближение центров щупов эллипсом может осуществляться различными методами, автором был использован метод наименьших квадратов /33/, обладающий высоким быстродействием и точностью.

Поведение модели можно варьировать и настраивать, модифицируя следующие параметры:

- количество, расположение и размер щупов;
- пороговые величины T_{in} и T_{out} ;
- коэффициенты длины векторов перемещения щупов k_{in} и k_{out} ;

- дополнительные ограничения на форму, размер, ориентацию и положение эллипса.

Путем модификации перечисленных параметров можно добиться тонкой настройки функционирования модели.

Модификации метода

Для увеличения устойчивости метода в случае присутствия на изображении регионов кожных пикселей, значительно отличающихся из-за шума по форме от эллиптических, введены несколько усовершенствований базового метода:

1. Малое случайное перемещение щупов вдоль границы модели на каждом шаге;
2. Предобработка изображения кожи - заполнение “дыр” в связных компонентах кожи заведомо меньшего размера, чем ожидаемый размер лица.

Перечисленные модификации позволяют успешно выделять регионы, сильно “попорченные” ошибками сегментации, которые приводят к появлению “дырок” внутри связных областей кожи.

Для того, чтобы увеличить влияние щупов, находящихся на границе области, введена третья модификация метода. Пересчет параметров эллипса производится методом наименьших квадратов по набору координат центров щупов $x_i^{k+1}, i = \overline{1, n}$. Особенность метода наименьших квадратов заключается в том, что ошибочное перемещение одного щупа на большое расстояние может значительно сказаться на форме модели и привести к

“срыву” модели с границы области кожи. Чтобы избежать такой ситуации и увеличить влияние стационарных щупов (в предположении, что они остановились на границе интересующей нас области), их центры участвуют несколько раз при пересчете параметров эллипса. Координаты центров стационарных щупов x_i дублируются l раз в массиве, передаваемом процедуре аппроксимации центром щупов эллипсом. Эксперименты показали, что $l = 5$ дает хороший компромисс между устойчивостью к несвязностям внутри областей и точностью выделения границ регионов.

Метод статистических моментов для обнаружения отслеживания перемещения лица

Задача отслеживания перемещения лица в видеопотоке обладает определенными особенностями по сравнению с обнаружением лица на статическом изображении. Во-первых, мы можем отталкиваться от положения лица на предыдущем кадре, во-вторых мы можем рассчитывать на то, что изменения в размере и, главное, пропорциях лица между кадрами будут невелики. С другой стороны, требуется значительно более высокое быстродействие и сходимость метода. Также, частым отличием при обработке видеопотока является заметно меньшее разрешение и худшее качество изображения по сравнению со статикой.

Учитывая перечисленные особенности, автором предлагается метод отслеживания перемещения лица, также основанный на методе последовательных приближений и использующий жесткое ограничение на возможную форму выделяемых областей, но использующий всю информацию внутри области, а не только на ее границе.

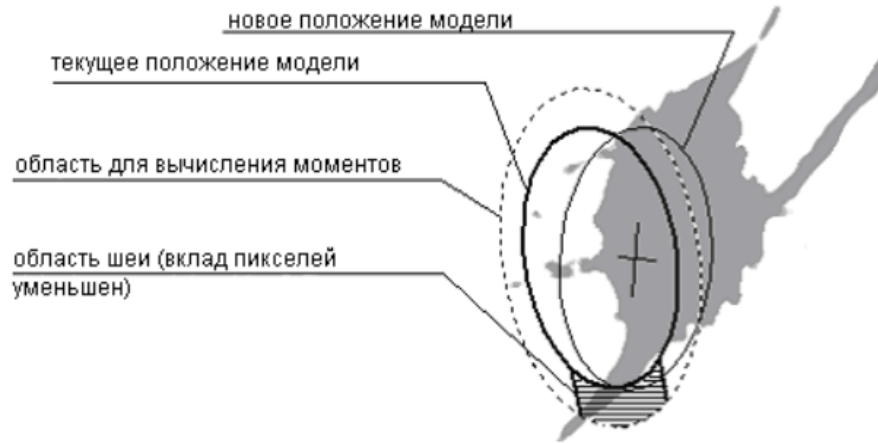


Рис. 1.8. Иллюстрация к методу статистических моментов для выделения области лица

Метод основан на вычислении статистических моментов множества пикселей цвета кожи, попадающих в текущую обрабатываемую область.

$$\begin{aligned}\mu_x &= \frac{\sum_{(x,y) \in S} y \cdot I(x, y)}{\sum_{(x,y) \in S} I(x, y)} \\ \mu_y &= \frac{\sum_{(x,y) \in S} x \cdot I(x, y)}{\sum_{(x,y) \in S} I(x, y)} \\ \mu_{ij} &= \frac{\sum_{(x,y) \in S} (x - \mu_x)^i \cdot (y - \mu_y)^j \cdot I(x, y)}{\sum_{(x,y) \in S} I(x, y)}\end{aligned}\quad (1.7)$$

В (2.8) S - обрабатываемая область изображения, $I(x, y)$ - вероятность принадлежности пикселя (x, y) к цвету кожи.

Инициализация модели производится либо положением лица на предыдущем кадре (в случае обработки видео), либо в центре масс связной компоненты пикселей кожи (как в случае с предыдущим методом). На каждом шаге деформации производится пересчет новых параметров эллипса. Пусть \vec{v}_k - вектор параметров эллипса на n -ом шаге (x, y, θ, a, b) . Здесь θ - угол наклона главной оси эллипса, a, b - длины осей эллипса. Полный цикл обнаружения области лица указанным методом описан в следующем

алгоритме:

0. Инициализация модели;

1. Количество итераций $k = 0$;

2. Пока изменение параметров модели $> \varepsilon$ и $k < N$:

3. Рассмотреть эллипс увеличенного размера $(x_k, y_k, \theta_k, \alpha \cdot a_k, \alpha \cdot b_k)$,
 $\alpha > 1$;

4. Подсчитать центр масс (μ_x, μ_y) и центральные моменты второго порядка $\mu_{20}, \mu_{02}, \mu_{11}$ множества пикселей, попавших в рассматриваемую область;

5. Подсчитать новые параметры эллипса по полученным статистикам
 $v_{k+1} = \left(\mu_x, \mu_y, \text{atan} \left(\frac{\mu_{11}}{\mu_{20} - \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \right), a_{k+1}, b_{k+1} \right)$;

6. $k = k + 1$, перейти на 2;

Новые длины осей эллипса a_{k+1}, b_{k+1} вычисляются по следующей формуле:

$$a_{k+1} = \frac{1}{1 + \beta} \cdot (a_k + d_k \cdot a_k \cdot \beta) \quad (1.8)$$

$$b_{k+1} = \frac{1}{1 + \beta} \cdot (b_k + d_k \cdot b_k \cdot \beta) \quad (1.9)$$

Где d_k - среднее расстояние Махаланобиса от центра эллипса лица до точек цвета кожи, находящихся внутри эллипса увеличенного размера, β - коэффициент, регулирующий скорость адаптации (по результатам экспериментов принятый равным 0.19).

Обычно, двух-трех итераций пересчета параметров эллипса достаточно для вычисления нового положения лица, даже при достаточно больших перемещениях между кадрами.

Сравнение методов с существующими

Были проведены испытания по сравнению ряда существующих в литературе методов с предлагаемыми. Для этого были реализованы следующие методы: метод связных компонент, метод интегральных проекций, метод радиального расширения, деформируемая модель из статьи /75/. Испытания показали, что предлагаемый метод обладает существенно более высокой устойчивостью к шумным исходным данным и ошибкам цветовой сегментации и быстрой сходимостью, добиваясь этого без увеличения вычислительной сложности алгоритма.

Простые схемы алгоритмов гарантируют возможность получения быстрой программной реализации. Несмотря на простоту, она обладает высокой устойчивостью к шумным и неточным исходным картам вероятности кожи. Предложенные методы также обладают важным свойством устойчивости к выбору начального приближения, что выгодно отличает их от многих методов распознавания, основанных на деформируемых моделях. Примеры работы приведены на рис. 1.9. Описанная схема обладает большой гибкостью и позволяет с легкостью встраивать дополнительные условия и ограничений на возможную форму выделяемых областей (пропорции, положение, ориентации эллипса).

Основная составляющая устойчивости обоих методов - жесткое ограничение на возможные формы модели (эллипс). Испытания показали, что метод эллиптической деформации больше подходит для применения при обнаружения лиц на статических изображениях среднего и высокого разрешения (выше точность, больше скорость на больших разрешениях), в то

время как метод статистических моментов больше подходит для отслеживания перемещения лица в видеопотоке (быстрее сходится).

Учитывая устойчивость алгоритмов к точности начального приближения на основе описанной схемы, строится следующий метод обнаружения лиц на изображениях:

1. Провести сегментацию кожи, создать карту вероятности кожи;
2. Выделить связанные компоненты пикселей с высокой кожной вероятностью или подсчитать интегральные проекции карты кожи;
3. Инициализировать модель в центрах масс связанных компонент или пиках интегральных проекций;
4. Произвести обнаружение эллиптических областей пикселей кожи;
5. Оценить их размеры, пропорции, форму и зафиксировать их как потенциальные лица, если проверка дает положительный результат;
6. Убрать обнаруженные области из карты вероятностей;
7. Повторять процедуру с шага 2, пока обнаруживаются новые области.

1.3. Заключение

Основными проблемами, с которыми сталкиваются методы, использующие связанные компоненты или интегральные проекции, являются: ошибочное разделение региона лица на несколько регионов и ошибочное объединение региона лица с каким-либо другим регионом. Все эти проблемы успешно решаются при применении описанных алгоритмов.