

Автоматическое выделение на цветных цифровых изображениях лица человека и его характерных черт

С.Ю. Желтов, Ю.В. Визильтер, М.В. Ососков, С.Л.Каратеев, И.В. Бекетова

ФГУП Государственный научно-исследовательский институт авиационных систем

125319, Москва, ул. Викторенко, 7 , тел. 157-91-58,

e-mail: viz@gosniias.ru

1. Введение.

В последние годы в работах по машинной графике и машинному зрению значительное внимание уделяется проблемам синтеза искусственных персонажей, действующих в виртуальной среде под управлением своих реальных прототипов в реальном масштабе времени. При этом внешность, жесты и мимика виртуальных персонажей формируются путем анализа реального изображения человека, находящегося перед видеокамерой. Актуальный характер сравнительно новых исследований указанных технологий анимации продиктован быстро растущим рынком соответствующих web-приложений и игровых программ, и в меньшей степени, традиционными применениями, такими как дистанционное управление сложными или опасными процессами.

К сожалению, до настоящего времени выделение на изображении человеческого лица, распознавание жестыкуляции и мимики безусловно относится к трудным задачам компьютерного анализа изображений, хотя бы на том основании, что число степеней свободы только кисти руки измеряется десятками, а число точек, выделяемых на лице человека подходит к тремстам. Конкретные трудности, с которыми пришлось столкнуться при решении описываемой задачи, были связаны с наличием активной мимики лица, отображение которой, с одной стороны, и являлось целью проекта, но с другой стороны, делало яркостно-геометрические свойства основных черт лица гораздо более изменчивыми и трудно формализуемыми, чем это обычно рассматривается в литературе при описании подобных задач анализа лица.

В работе представлены результаты разработки алгоритмического обеспечения для автоматического обнаружения лица и его характерных крупных элементов (глаз, рта, бровей, носа), слежения за этими объектами и их последующего отображения в реальном масштабе времени на персональном компьютере типа IBM PC. В качестве устройства ввода используется цветная видеокамера низкого разрешения. Лицо пользователя расположено примерно анфас к камере и не закрыто никакими другими предметами. Допускаются небольшие повороты и наклоны лица. Условия освещенности обычные, офисные.

Разработанная процедура анализа изображения включает следующие основные блоки:

- выделение лица и слежение за ним;
- выделение бровей;
- выделение носа;
- выделение рта.

Рассмотрим подробнее каждый из этих алгоритмических блоков.

2. Алгоритмы сегментации лица.

2.1. Цветовые модели.

Назначение цветовой модели - дать средства описания цвета в пределах некоторого цветового охвата. Наиболее часто в компьютерном зрении используются модели RGB, CMY, YUV, YCbCr, HSV. RGB (Red, Green, Blue - красный, зеленый, синий) - аппаратно-ориентированная модель, используемая в дисплеях для аддитивного формирования оттенков самосветящихся объектов (пикселей экрана). Система координат RGB - куб с началом отсчета (0,0,0), соответствующим черному цвету (рис.1). Максимальное значение RGB - (1,1,1) соответствует белому цвету. В модели RGB не разделены яркостная и оттеночная компоненты

цвета, здесь легко указать яркости для одного из основных цветов, но затруднительно указать оттенок с требуемым цветовым тоном (например, телесным) и насыщенностью.

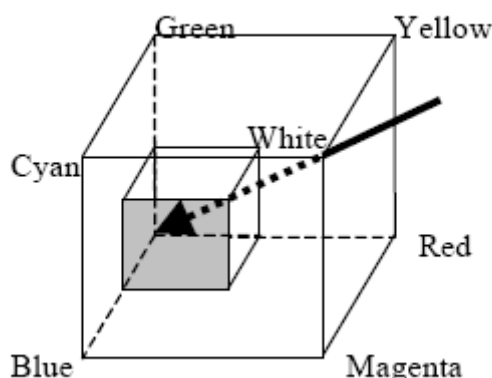


Рис.1. Цветовая модель RGB.

HSV (Hue, Saturation, Value - цветовой тон, насыщенность, мера яркости) - модель, ориентированная на человека и обеспечивающая возможность явного задания требуемого оттенка цвета. Среди прочих используемых в настоящее время моделей, эта модель отражает физические свойства цвета и наиболее точно соответствует способу восприятия цвета человеческим глазом. Оттенок цвета (H) ассоциируется в человеческом сознании с обусловленностью окраски предмета определенным типом пигмента, краски, красителя. Составляющая H - это длина световой волны, отраженной или прошедшей через объект. Данная составляющая является объективной и отражает физическую сущность цвета. Насыщенность (S) характеризует степень, силу, уровень выражения оттенка цвета. Этот атрибут в человеческом сознании связан с количеством (концентрацией) пигмента, краски, красителя. Насыщенность позволяет оценить, насколько "глубоким" и "чистым" является данный цвет, то есть как он отличается от ахроматического. Она определяет соотношение серого цвета и выбранного оттенка и выражается в долях от 0 (серый) до 1 (полностью насыщенный). Мера яркости (V) характеризует относительную освещенность или затемненность цвета (интенсивность цвета), поэтому она измеряется в диапазоне от 0 (черный) до 1 (белый). При увеличении яркости цвет становится более светлым (белым). Составляющая V является нелинейной, что близко соответствует восприятию человеком светлых и темных цветов. Две последние составляющие являются субъективно определяемыми, так как отражают психологические особенности восприятия цвета.

Подпространство, определяемое данной моделью, соответствует проекции цветового пространства на главную диагональ RGB-куба в направлении от белого к черному и являет собой перевернутый шестигранный конус (рис.2). По вертикальной оси конуса задается V - мера яркости. Значению $V=0$ соответствует вершина конуса, значению $V=1$ - основание конуса; цвета при этом наиболее интенсивны. Цветовой тон H задается углом, отсчитываемым вокруг вертикальной оси. В частности, 0° - чистый красный цвет (Red), 60° - желтый (Yellow), 120° - зеленый (Green), 180° - голубой (Cyan), 240° - синий (Blue), 300° - пурпурный (Magenta), т.е. дополнительные цвета расположены друг против друга (отличаются на 180°). Насыщенность S определяет, насколько близок цвет к "чистому" пигменту и меняется от 0 на вертикальной оси V до 1 на боковых гранях шестигранного конуса. Точка $V=0$, в которой находится вершина конуса, соответствует черному цвету. Значение S при этом может быть любым в диапазоне 0-1. Точка с координатами $V=1, S=0$ - центр основания конуса соответствует белому цвету. Промежуточные значения координаты V при $S=0$, т.е. на оси конуса, соответствуют серым цветам. Если $S=0$, то значение оттенка H считается неопределенным.

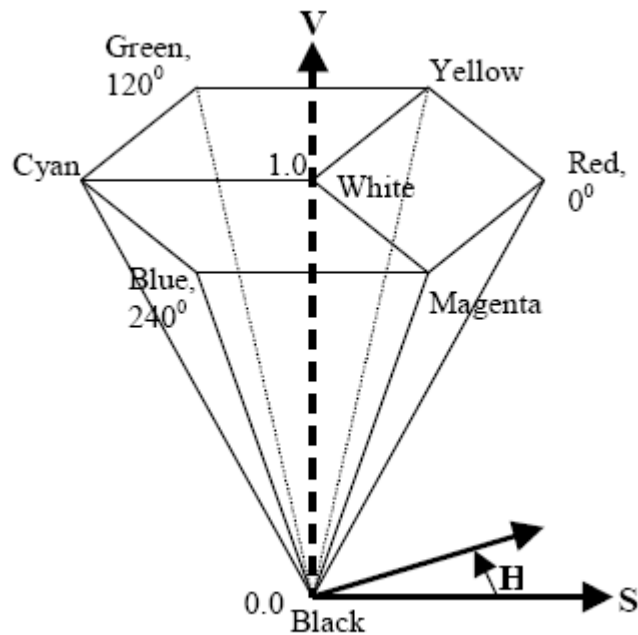


Рис.2. Цветовая модель HSV.

Цветовая модель HSV может быть получена из модели RGB следующим образом:

$$H_1 = \arccos \left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right)$$

$$H = \begin{cases} H_1 & , \text{ if } B \leq G \\ 360^\circ - H_1 & , \text{ if } B > G \end{cases}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

$$V = \frac{\max(R, G, B)}{255}$$

Данное преобразование является нелинейным, что затрудняет его практическое использование. Аппроксимация преобразования линейными зависимостями существенно упрощает алгоритм перевода из RGB в HSV и обеспечивает высокую вычислительную эффективность. Цветовая модель HSV является наиболее удобным представлением цветных изображений для их цветовой сегментации.

2.2. Цветовая сегментация.

Цветовая сегментация производится по предварительно построенной модели распределения цвета кожи в цветовом пространстве. Построение модели рассматривается как задача поиска группы кластеров, соответствующих оттенкам кожи. Структура и параметры модели формируются путем обработки изображений участков кожи, выделенных вручную. Распределение, отражающее соотношение параметров цвета H и S получено в результате обработки порядка 900 изображений фрагментов кожи размерами 3x3 пикселей (рис.3).

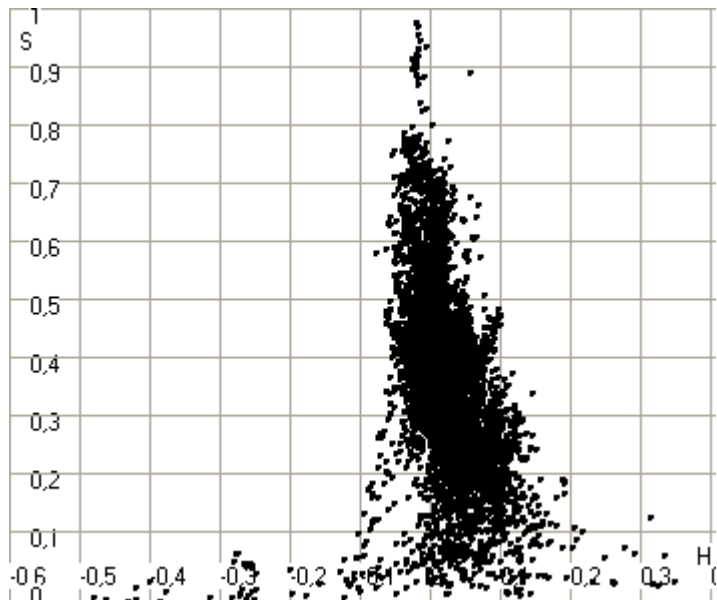


Рис.3. Взаимное соотношение компонент H и S на изображениях кожи.

Аналогичным образом получено распределение, отражающее соотношение параметров цвета H и V (рис.4).

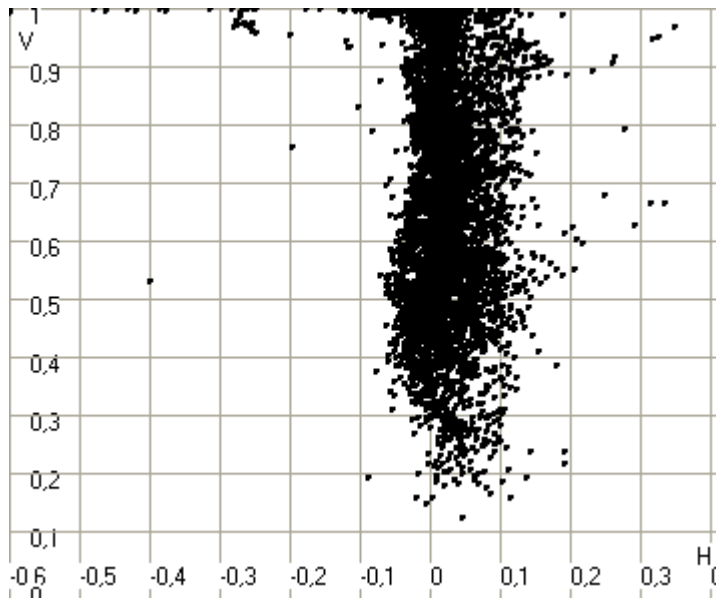


Рис.4. Взаимное соотношение компонент H и V на изображениях кожи.

«Трубки» распределений (H,S) и (H,V) , характеризующие оттенки кожи, являются достаточно компактными в цветовом пространстве и составляют группу первичных признаков на этапе построения модели. «Трубки» распределений (H,S) и (H,V) могут быть разделены на подобласти, образующие группы кластеров Σc_i . На этапе сегментации каждому пикселю изображения ставится в соответствие вектор параметров цвета (H,S,V) и проверяется принадлежность параметров цвета к одному из кластеров модели кожи Σc_i . Область изображения кожи формируется из пикселей, векторы параметров которых вошли в один из кластеров. На рис.5 – 1.2.7 приведены результаты работы алгоритма цветовой пиксельной сегментации лиц с различными оттенками кожи.

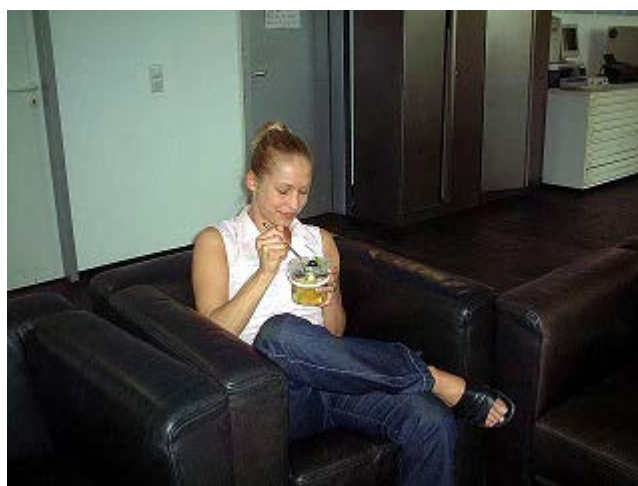


Рис.5. Пример цветовой пиксельной сегментации кожи.



Рис.6. Пример цветовой пиксельной сегментации кожи.



Рис.7. Пример цветовой пиксельной сегментации кожи.

В результате сегментации изображение имеет ряд недостатков в виде мелких областей, линий и отдельных пикселей, явно не относящихся к области кожи. Поэтому для удаления таких областей к изображению применяются такие операции математической морфологии, как дилатация и эрозия.

После того, как на изображении выделены пиксели, цвет которых соответствует оттенкам кожи, производится слияние выделенных пикселей в систему кластеров по мере близости их местоположения. Далее производится комплексная проверка кластеров для выявления кластеров, действительно являющихся лицами. Правило, которому должен отвечать фрагмент изображения, на котором присутствует лицо, достаточно очевидно: лицо обычно симметрично. На рис.8 – 1.2.11 показаны примеры работы алгоритма обнаружения лиц на цветных изображениях. Фрагменты изображения с обнаруженными лицами выделены рамкой и снабжены меткой.



Рис.8. Пример работы алгоритма обнаружения лиц

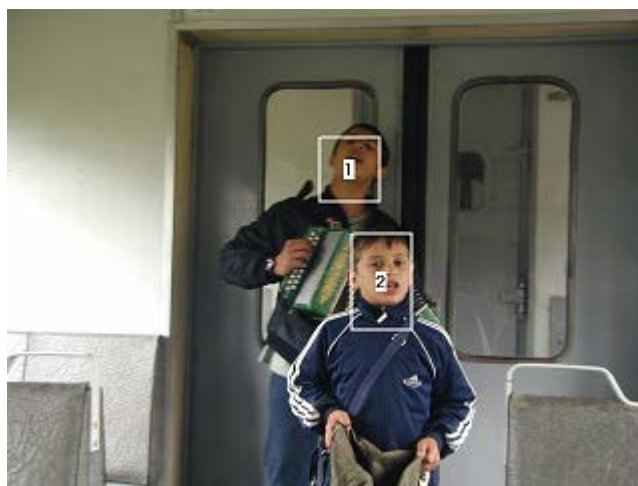


Рис.9. Пример работы алгоритма обнаружения лиц.



Рис.10. Пример работы алгоритма обнаружения лиц.



Рис.11. Пример работы алгоритма обнаружения лиц.

3. Алгоритмы выделения глаз и бровей.

Процедура выделения глаз и бровей состоит из следующих основных процедур:

- определение зоны глаз/бровей с помощью специального оператора интереса;
- корреляционная привязка точки переносицы;
- обнаружение бровей;
- обнаружение глаз.

3.1. Определение зоны глаз/бровей с помощью специального оператора интереса.

Первичный детектор зоны глаз/бровей основан на самых общих предположениях о чертах лица. Предполагается что на лице имеются «горизонтальные черты», что выражается в том что производная яркости по оси Y больше производной яркости по оси X. Предполагается также, что яркость горизонтальных черт лица ниже, чем яркость пикселей лица. Исходя из этих соображений, разработанный оператор интереса в каждом пикселе вычисляет следующий отклик по окрестности:

$$\Sigma (255 - \min\{I[x,y]\}) * I_y[x,y] / I_x[x,y]$$

При этом положение глобального максимума горизонтальной проекции соответствующего синтетического изображения, собранной в области верхней половины лица, соответствует бровям.

3.2. Корреляционная привязка точки переносицы

Как было установлено в ходе исследований, переносица является наиболее стабильной точкой лица в условиях сильного изменения формы изображения лица при активной мимике, что позволяет успешно проследивать ее положение с помощью корреляционного метода. Процедура корреляционного слежения осуществляется стандартным образом. Для достижения большего быстродействия и устойчивости, корреляционное слежение реализовано с использованием пирамиды изображений. При этом на N-м уровне пирамиды используется прореживание области поиска и корреляционного окна в 2^{N+1} раз (по площади). Для N=3 это дает уменьшение времени слежения примерно в 100 раз.

На каждом кадре также осуществляется дополнительная коррекция положения эталона на основе рассмотренного выше оператора выделения горизонтальных черт.

3.3. Выделение бровей.

Блок выделения бровей включает следующие этапы:

- формирование зоны поиска бровей
- сегментация зоны поиска оператором Марра

- маскирование с помощью "теневого" и "цветовой" масок
- собственно выделение бровей

Формирование зоны поиска бровей осуществляется, исходя из ранее сформированной прямоугольной области интереса, содержащей лицо и прямоугольной области корреляционного слежения за переносицей. Далее используется стандартная процедура сегментации изображения оператором Марра ([2]). После этого вычисляется теневая маска, формируемая как совокупность пикселей, яркость которых ниже среднего значения, по вертикальному окну на всю высоту области лица. Эта «теневая», а также ранее сформированная «цветовая» маска применяются к изображению, полученному оператором Марра.



Рис.12. Результат применения теневого и цветовой масок к препарату Марра в области глаз/бровей. После этого осуществляется собственно процедура выделения бровей. Она состоит из следующих этапов:

- выделение вертикальных сегментов препарата Марра
- прометка крайних сегментов снизу-вверх (от глаз) и сверху-вниз (от лба)
- горизонтальное связывание помеченных вертикальных сегментов
- уточнение положения центра переносицы
- выделение нижней и верхней границ бровей
- аппроксимация границ бровей сплайнами по трем точкам

При уточнении положения центра переносицы рассматриваются вертикальная и горизонтальная проекции «теневого» маски в области поиска бровей. При аппроксимации границ бровей сплайнами, в качестве крайних опорных точек сплайна выбираются начальная и конечная точки найденных провязок максимальной длины, а в качестве внутренней точки принимается точка провязки, находящаяся на максимальном расстоянии от прямой, соединяющей крайние точки.

3.4. Обнаружение открытых глаз

Зона поиска глаз формируется на основании информации о трех точках границы брови и оценки расстояния между глазами. Далее осуществляется свертка изображения с маской, соответствующей лапласиану одномерной гауссианы, с последующей бинаризацией полученного изображения по знаку.

Для обнаружения глаз используется строчный оператор, состоящий из трех площадок. Ширина центральной площадки, соответствующей ширине зрачка глаза, определяется из бинарного препарата Марра. Ширина крайних площадок оператора фиксирована. В области центральной и крайних площадок определяется среднее значение цвета расстояния между цветами. Кроме этого для каждой из площадок вычисляется среднее значение яркости и определяется отношение сигнал/шум. Из всех кандидатов выбирается точка с максимальным отношением расстояния между цветами к среднему значению яркости в центральной площадке.

4. Алгоритм выделения носа.

Блок выделения носа состоит из двух основных процедур:

- первичное детектирование и слежение за положением нижней границы носа;
- прослеживание контуров носа.

Первичный детектор носа осуществляет установку корреляционного окна в начальный момент или после срыва слежения. Первичный детектор также позволяет подправить положение корреляционного окна при его смещении относительно объекта слежения. Первичный детектор носа, как и первичный детектор переносицы (см. выше), основан на операторе горизонтальных черт лица. В качестве носа (ноздрей) принимается первый существенный максимум аккумулятора горизонтальных черт, если двигаться по оси Y' локальной системы координат, связанной с центрами глаз, вниз от переносицы.

На этапе прослеживания принята следующая яркостно-геометрическая модель носа: яркий контур посередине и два темных контура слева и справа (рис.13).

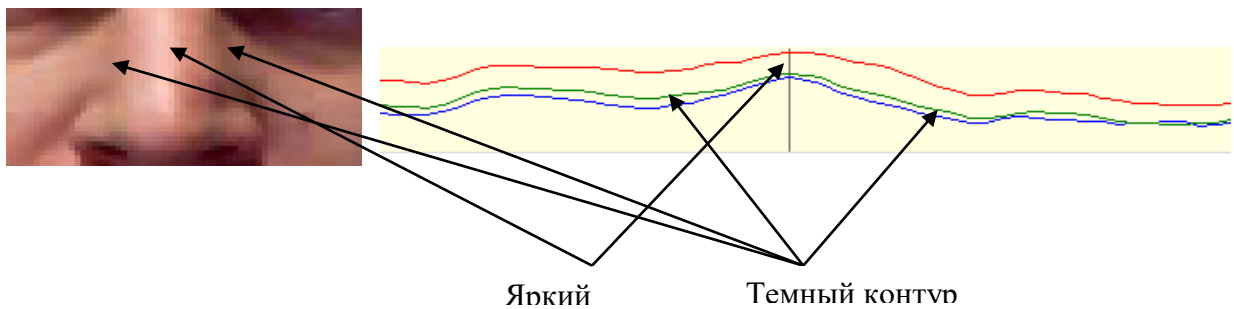


Рис.13 Яркостно-геометрическая модель носа

Для отображения носа средствами машинной графики необходимо в области носа выделять краевые линии и центральную линию. Для этого во всех горизонтальных профилях области интереса производится поиск глобального максимума и двух существенных боковых минимумов яркости. После этого используется прослеживание методом динамического программирования. Результатом прослеживания линий носа являются: конечные точки линии S , соединяемые прямой линией; конечные точки линий L,R , соединяемые сплайн-кривой.

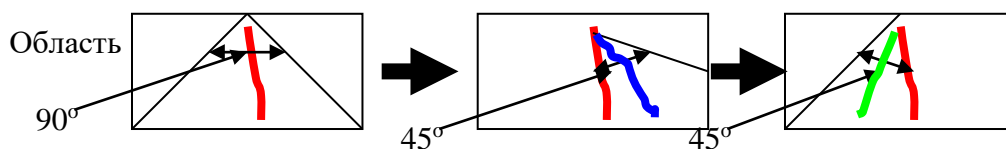


Рис.14. Последовательность получения линий носа

5. Алгоритмы выделения рта.

С учетом различий в мимике, процедура выделения рта распадается на две основные алгоритмические ветви:

- выделение линии губ в случае закрытого рта
- выделение внутренней области рта в случае открытого рта.

В качестве яркостно-геометрической модели закрытого рта используется следующая модель: имеется достаточно острый глобальный минимум на всех вертикальных профилях в области рта (рис.15). Для подчеркивания данного признака используется операция морфологической нормализации фона на основе оператора открытия ([3]). Далее также производится прослеживание линии рта методом динамического программирования.



Рис.15. Яркостно-геометрическая модель закрытого рта. Прослеживание линии из середины области с помощью двух оптимальных путей

Для ограничения линии рта справа и слева используются линии, образованные краями носа (рис.16).

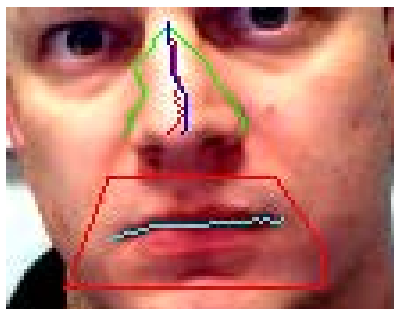


Рис.16. Прослеживание линии закрытого рта.

Алгоритм выделения открытого рта, как темной замкнутой области, запускается после алгоритма обнаружения закрытого рта. Если вертикальные размеры обнаруженной области меньше порога, то считается, что рот закрыт, и в этом случае результатом является сплайн-кривая, соединяющая несколько точек найденной области R . Иначе, в качестве результата выдается замкнутая линия, соединяющая несколько характерных точек обнаруженной области. Обнаружение темной области, основано на пороговой сегментации области рта. Порог вычисляется адаптивно в области интереса путем оптимального разделения бимодального распределения яркости по критерию Оцу. В качестве результата обнаружения выбирается связная область, которая пересекается с областью интереса и имеет наибольшую площадь (рис.17).

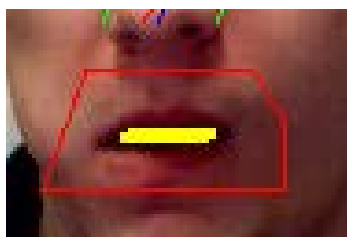


Рис.17. Обнаружение открытого рта как темной замкнутой области.

Таким образом, на каждом кадре оказываются выделенными основные черты лица.

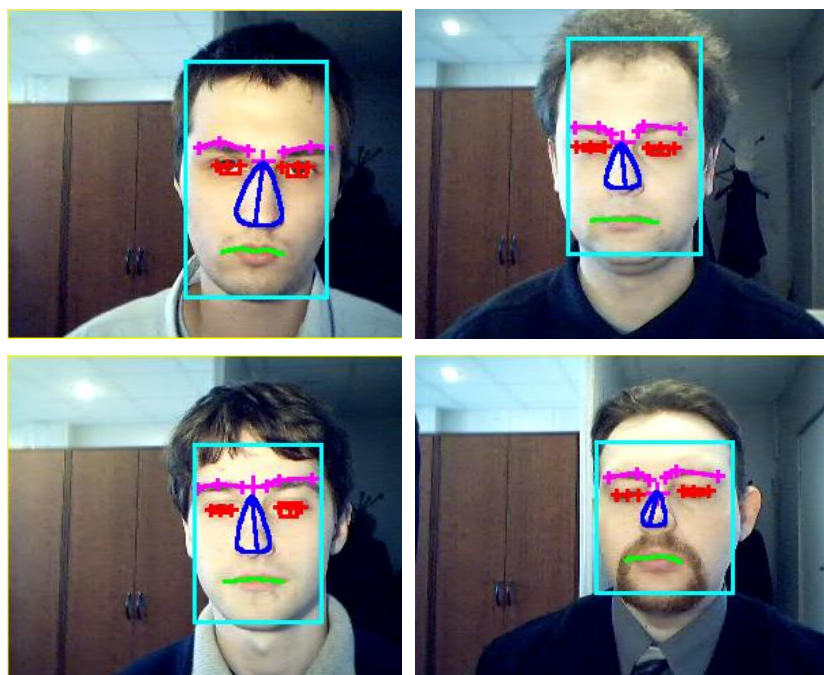


Рис.18. Примеры выделения на различных изображениях лица и его основных черт.

6. Заключение.

В работе представлены результаты разработки алгоритмического и программного обеспечения для автоматического обнаружения лица и его характерных крупных элементов (глаз, рта, бровей, носа), слежения за этими объектами и их последующего отображения в реальном масштабе времени на персональном компьютере типа IBM PC. В качестве устройства ввода используется цветная видео камера. На этапе выделения лица используется специальная процедура сегментации в цветовом пространстве HSV. На этапе выделения глаз и бровей используются специальные операторы интереса, а также оператор Марра. На этапе выделения носа и рта используются различные морфологические операторы. Результаты выделения границ бровей, носа и рта аппроксимируются сплайнами по нескольким характерным точкам. Вся выделенная таким образом информация представляется в виде векторных графических элементов для передачи по сетям или дальнейшей визуализации средствами машинной графики в реальном времени.

Разработанная система тестировалась на значительном объеме реальных видео последовательностей, снятых с помощью видеокамеры. При этом продемонстрировано приемлемое качество распознавания элементов лица.

Литература.

1. Дуда Р., Харт П. Распознавание образов и анализ сцен: Пер. с англ. – М.: Мир, 1976, с.511.
2. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов, М. Радио и связь, 1987.
3. J. Serra. Image Analysis and Mathematical Morphology. Academic Press. 1982.