

А. И. Пахирка

ПРИМЕНЕНИЕ МЕТОДА УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМ РАСПОЗНАВАНИЯ ЛИЦ

Представлен алгоритм обработки изображений лиц, включающий три этапа: нелинейное улучшение изображения (сжатие динамического диапазона), локализация лиц на основе цветовой сегментации кожи с последующим выделением антропометрических точек лица. Также рассмотрен процесс распознавания лиц на основе метода главных компонент.

Ключевые слова: сжатие динамического диапазона, локализация лиц, распознавание лиц.

Интерес к распознаванию лиц всегда был значительным, особенно в связи с возрастающими практическими потребностями, такими как биометрия, содержательный поиск изображений, компрессия видеоданных, организация видеоконференций, системы машинного зрения в робототехнике, интеллектуальные системы безопасности и контроля доступа.

Методы распознавания лиц могут быть разделены на две широкие категории: методы, основанные на извлечении особенностей изображения, и методы, основанные на представлении изображения лица. Первая группа методов использует свойства и геометрические отношения, такие как области, расстояния и углы между особыми точками изображения лица. Вторая группа методов рассматривает глобальные свойства изображения лица. Названные методы стараются представить данные изображения лица более эффективно, например, как набор главных векторов. Как правило, алгоритм распознавания лиц включает в себя три этапа: предварительную обработку изображения, локализацию лиц и непосредственное распознавание лиц. В данной работе представлен алгоритм, включающий нелинейное улучшение изображений (сжатие динамического диапазона), локализацию лиц на основе цветовой сегментации кожи, а также распознавание лиц на основе метода главных компонент [1].

На практике захватываемые цифровыми устройствами изображения зачастую отличаются от того, что видит наблюдатель. Это происходит потому, что устройство захвата получает физические значения световых данных, в то время как нервная система наблюдателя обрабатывает эти данные. Например, человек может четко видеть детали, как в глубоких тенях, так и в сильно освещенных областях, в то же время устройство захвата получит данную сцену со слишком темными тенями или засвеченными областями. Человек легко воспринимает сцены с широким диапазоном световых интенсивностей (HDR, High Dynamic Range – высокий динамический диапазон), при этом отношение между максимальной и минимальной яркостью превышает возможность устройства захвата или отображения.

Человек, наблюдая HDR-сцену, локально адаптирует каждую ее часть, благодаря чему может рассматривать детали в плохо освещенных областях так же хорошо, как и в ярко освещенных. Для цифровых устройств HDR-сцена требует сжатия, из-за чего захватываемое изображение зачастую теряет детали в плохо или ярко освещенных областях. Для устройств захвата это решается путем комбинирования изображений, снятых с разной экспозицией (тех-

нология HDR в фотографии), в результате которого получается единое изображение, содержащее все детали из всех исходных изображений, как в тенях, так и в освещенных областях. Однако остается проблема отображения данных изображений на цифровых устройствах, обладающих существенно меньшим диапазоном яркостей [2].

В данной работе для улучшения изображения применяется алгоритм Multi-Scale Retinex – MSR, имитирующий визуальную систему человека. MSR-алгоритм сжимает динамический диапазон изображения с сохранением (увеличением) локального контраста в плохо и ярко освещенных областях [3].

Классический многомерный MSR-алгоритм является взвешенной суммой одномерных SSR-алгоритмов (Single-Scale Retinex) для различных масштабов. Одномерная выходная функция i -го цветового канала $R_i(x, y, s)$ вычисляется следующим образом:

$$R_i(x, y, s) = \log\{I_i(x, y)\} - \log\{F(x, y, s) * I_i(x, y)\},$$

где $I_i(x, y)$ – входная функция i -го цветового канала по координатам x и y ; s – масштабный коэффициент; знак * обозначает свертку функций; $F(x, y, s)$ – гауссиан, определяемый как

$$F(x, y, s) = Ke^{-(x^2 + y^2)/s^2}.$$

При этом коэффициент K выбирается таким образом, чтобы выполнялось условие

$$\iint_{\Omega_{x,y}} F(x, y, \sigma) dx dy = 1,$$

где $\Omega_{x,y}$ – множество пикселей, принадлежащих всему изображению.

Тогда многомерная выходная функция i -го цветового канала $R_{M_i}(x, y, w, s)$ определяется как

$$R_{M_i}(x, y, w, \sigma) = \sum_{n=1}^N w_n R_i(x, y, \sigma_n),$$

где $w = (w_1, w_2, \dots, w_m)$, $m = 1, 2, \dots, M$ – весовой вектор одномерных выходных функций i -го цветового канала $R_i(x, y, s)$; $s = (s_1, s_2, \dots, s_n)$, $n = 1, 2, \dots, N$ – вектор масштабов одномерных выходных функций. При этом $\sum_{n=1}^N w_n = 1$. Раз-

мерность вектора масштабов обычно выбирается не меньше 3. В различных источниках приводятся разные рекомендуемые значения масштабов, в наших экспериментах они составили 15, 90, 180. Весовой вектор w , как правило, имеет элементы с равными значениями.

Функциональная схема модуля улучшения изображения показана на рис. 1. Переход к цветовому простран-

ству $YCbCr$ обусловлен тем, что в нем яркостная компонента представлена независимо от других. Поэтому алгоритм применяется только к Y -компоненте, не затрагивая Cb и Cr , что, в свою очередь, увеличивает скорость работы алгоритма. Для Гауссова размытия применяется рекурсивный алгоритм фильтрации, который аппроксимирует гауссиан, с расчетом коэффициентов фильтра для желаемого значения сигма (s), такое представление фильтра работает значительно быстрее, чем стандартная фильтрация с использованием ядра свертки [4].

Пример работы MSR-алгоритма для изображения с низкой освещенностью показан на рис. 2.

После улучшения изображения следует этап локализации лиц на основе цветовой сегментации кожи. Процесс локализации лица на изображении можно разделить на две составляющие:

- выделение участков изображения, имеющих цвет, сходный с цветом кожи человека (цветовая сегментация);
- анализ выделенных после сегментации регионов.

Определение цвета кожи позволяет существенно образом сократить область поиска и является первым шагом во многих методах локализации лица. Человеческая кожа имеет характерный оттенок, позволяющий успешно сегментировать кожу на цветных изображениях. Независим

мость оттеночной компоненты цвета кожи от ориентации лица, а также его небольшая зависимость от яркости освещения, делают цвет устойчивым признаком кожи. Преимуществами метода цветовой сегментации кожи являются:

- малая вычислительная сложность;
- устойчивость к изменению масштаба и повороту лица;
- устойчивость к изменению освещения;
- устойчивость к изменению выражения лица и частичному перекрытию лица другим объектом сцены.

Цветовая сегментация кожи человека на изображении требует построение определяющих правил, которые будут разделять пиксели цвета кожи и пиксели, не относящиеся по цвету к коже. Для этого вводится метрика, позволяющая измерить расстояние между цветом пикселя и тоном кожи. Даная метрика является моделью распределения цвета кожи в выбранном цветовом пространстве.

Используем метрику для нормализованного RGB-цветового пространства, для которого распределение цвета кожи представлено на рис. 3:

$$\text{Skin}(r, g) = \begin{cases} 1 & \text{if } (g < g_u) \cdot (g > g_d) \cdot (W > 0,0004), \\ 0 & \text{otherwise,} \end{cases}$$

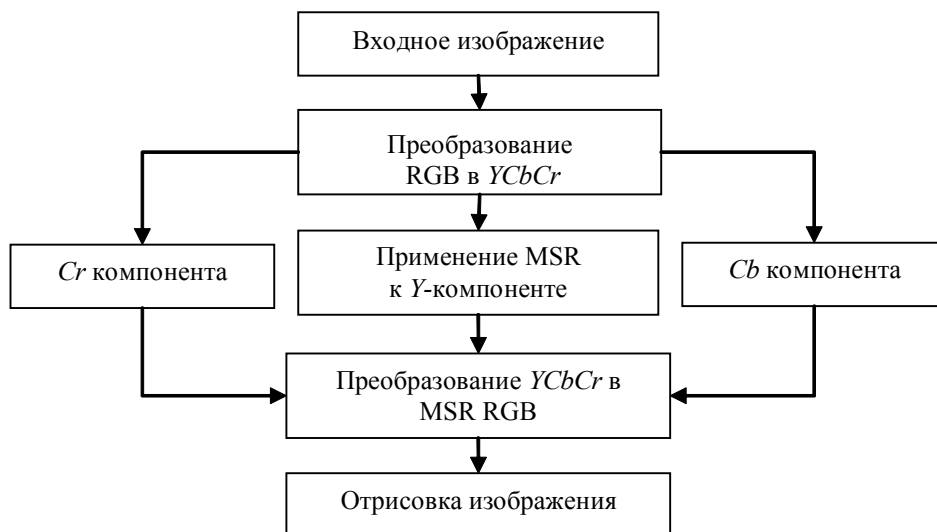


Рис 1. Функциональная схема применения MSR-алгоритма к изображению

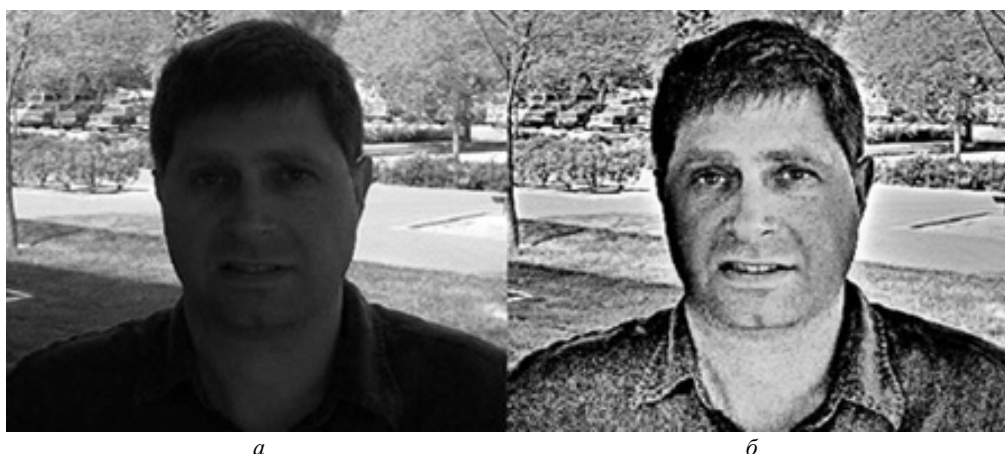


Рис 2. Пример работы MSR-алгоритма:

a – входное изображение; b – выходное изображение после обработки MSR-алгоритмом

где u – верхняя граница; d – нижняя граница. Значения g_u , g_d , W определяются как

$$g_u = J_u r^2 + K_u r + L_u, \quad g_d = J_d r^2 + K_d r + L_d,$$

$$W = (r - 0,33)^2 + (g - 0,33)^2,$$

а коэффициенты принимают следующие значения:

$$J_u = -1,377, \quad K_u = 1,074, \quad L_u = 0,145,$$

$$J_d = -0,776, \quad K_d = 0,560, \quad L_d = 0,177.$$

Пример сегментации кожи представлен на рис. 4. Сегментированное изображение подвергается морфологической обработке (сжатие с последующим расширени-

ем), которая позволяет разъединить плохо связанные регионы и удалить регионы малого размера (шум). Далее производится маркировка связанных областей (рис. 5), после чего в каждой области ищутся антропометрические точки (глаза, нос, губы).

Распознавание лиц осуществляется с помощью метода главных компонент (Principal Components Analysis, PCA), который позволяет уменьшить размерность данных за счет минимизации потерь информации. Метод главных компонент (МГК) один из наиболее часто используемых методов для снижения размерности данных в системах распознавания и сжатия. Данный метод проеци-

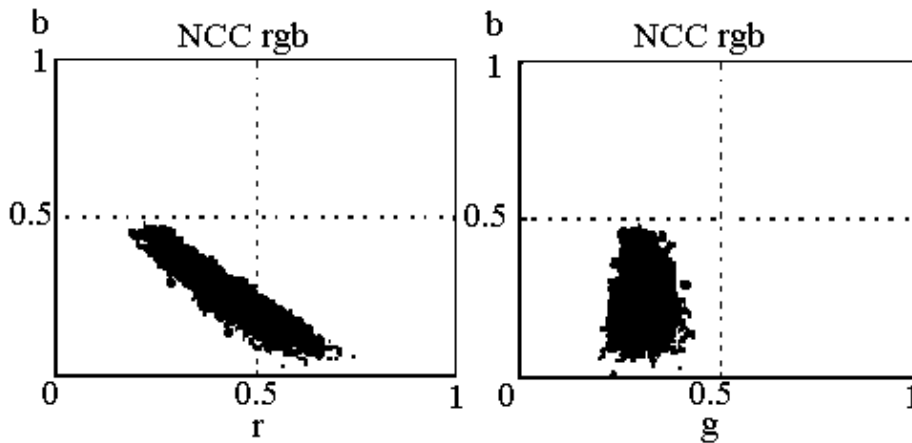


Рис 3. Распределение цвета кожи для нормализованного RGB-пространства



Рис. 4. Пример сегментации кожи: a – входное изображение; b – выходное изображение после применения метрики

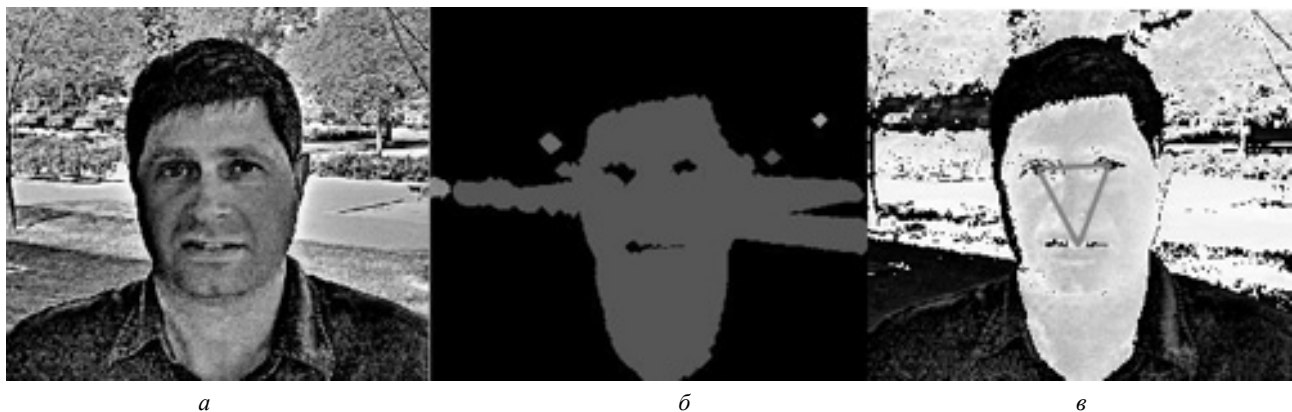


Рис. 5. Пример локализации лиц: a – входное изображение; b – морфологическая обработка с маркировкой связанных областей; c – выделение антропометрических точек лица

рует пространство изображения в пространство признаков меньшей размерности. Главная идея метода главных компонент состоит в представлении изображений лиц людей в виде набора главных компонент изображений, называемых «собственные лица» (Eigenfaces). Вычисление главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы, которая рассчитывается из изображения [5].

Любое изображение может рассматриваться как вектор из пикселей, каждое значение которого представлено значением в полутоновой градации. Например, изображение 8×8 пикселей может рассматриваться как вектор длиной в 64 пикселя. Такое векторное представление описывает входное пространство изображения. Для представления и распознавания лиц используем подпространство, созданное собственными векторами ковариационной матрицы исследуемых изображений. Собственные векторы, соответствующие ненулевым собственным значениям ковариационной матрицы, формируют ортогональный базис, который отображает изображения в N -мерное пространство [5]. Каждое изображение сохраняется в векторе размера N :

$$\mathbf{X} = [x_1^i \dots x_N^i]^T, \quad (1)$$

где x^i – эталонные изображения; \mathbf{X} – матрица эталонных изображений. Изображения центрируются вычитанием из каждого вектора усредненного изображения:

$$\bar{x}^i = x^i - m, \quad \text{где } m = \frac{1}{P} \sum_{i=1}^P x^i. \quad (2)$$

Эти векторы объединяются, образуя матрицу данных $N \times P$ (где P – количество изображений; \bar{x}^i – центрированное изображение) следующим образом:

$$\bar{\mathbf{X}} = [\bar{x}^1 \quad \bar{x}^2 \quad \dots \quad \bar{x}^P]. \quad (3)$$

Матрица данных $\bar{\mathbf{X}}$ умножается на транспонированную матрицу данных для расчета ковариационной матрицы

$$\mathbf{\Omega} = \bar{\mathbf{X}} \bar{\mathbf{X}}^T. \quad (4)$$

Эта ковариационная матрица $\mathbf{\Omega}$ имеет до P собственных векторов, связанных с ненулевыми собственными значениями, при этом $P < N$. Собственные векторы сортируются от большего значения к меньшему значению в соответствии с их собственными значениями. Собственный вектор с наибольшим собственным значением представляет самую большую дисперсию в изображениях.

Распознавание изображений с использованием проецирования на собственное пространство осуществляется в три этапа:

1. Создается собственное пространство из эталонных изображений (этап обучения).
2. Эталонные изображения проецируются в собственное пространство (этап обучения).
3. Спроецированное входное изображение сравнивается с проецированным тестовым изображением (этап распознавания).

Рассмотрим первый этап – создание собственного пространства, состоящего из следующих шагов:

– центрирование данных: каждое изображение центрируется вычитанием усредненного изображения из каж-

дого эталонного изображения. Усредненное изображение – это вектор-столбец, в который входят средние значения пикселей из всех пикселей эталонных изображений (выражение (2));

– создание матрицы данных: как только входные изображения центрированы, они комбинируются в матрицу данных $N \times P$ (выражение (3));

– создание ковариационной матрицы: матрица данных умножается на ее транспонированное представление (выражение (4));

– вычисление собственных векторов и собственных значений: собственные векторы и их собственные значения вычисляются из ковариационной матрицы

$$\mathbf{\Omega} \mathbf{V} = \mathbf{\Lambda} \mathbf{V},$$

где \mathbf{V} – набор собственных векторов связанных с собственными значениями $\mathbf{\Lambda}$;

– упорядочивание собственных векторов: упорядочиваются собственные векторы $v_i \in \mathbf{V}$ в соответствии с их собственными значениями $\lambda_i \in \mathbf{\Lambda}$ от большего значения к меньшему значению. Сохраняются собственные векторы с ненулевыми собственными значениями. Эта матрица собственных векторов является собственным пространством \mathbf{V} , где каждый столбец – собственный вектор:

$$\mathbf{V} = [v_1 \quad v_2 \quad \dots \quad v_p].$$

На втором этапе происходит проецирование эталонных изображений в собственное пространство. Каждое центрированное входное изображение \bar{x}^i проецируется в собственное пространство:

$$\tilde{x}^i = \mathbf{V}^T \bar{x}^i.$$

В рамках третьего этапа происходит распознавание входного изображения. Каждое входное изображение центрируется вычитанием усредненного изображения и затем проецируется в собственное пространство \mathbf{V} :

$$\bar{y}^i = y^i - m, \quad \text{где } m = \frac{1}{P} \sum_{i=1}^P x^i \quad \text{и} \quad \tilde{y}^i = \mathbf{V}^T \bar{y}^i.$$

Спроецированное входное изображение сравнивается со всеми спроецированными эталонными изображениями. Изображения могут сравниваться с использованием любой из простых метрик, например евклидовой.

В настоящий момент проводится разработка системы нелинейного улучшения изображений в разных цветовых пространствах, планируется использовать алгоритм Multi-Scale Retinex с восстановлением цветов для захвата и обработки видеоизображений, имеющих большой диапазон значений яркостей. Разрабатывается система захвата изображений лиц из видеоизображения, с последующей обработкой и приведением изображений к некоторому «усредненному» виду, снижением влияния освещения, корректировкой положения лица, выбором из видеоданных относительно лучшего изображения лица.

Таким образом, в статье предлагается усовершенствованный подход к распознаванию лиц по изображению, использующий алгоритм нелинейного улучшения изображения, который позволяет скомпенсировать тени и блики. Также проведенный анализ цветовых пространств позволяет повысить качество распознавания сегментов кожи и антропометрических точек лица.

Библиографические ссылки

1. Jain K., Flynn P., Ross A. Handbook of Biometrics. Springer, 2008.
2. Meylan L., Susstrunk S. Bio-inspired color image enhancement // SPIE Electronic Imaging. San Jose. 2004. P. 46–56.
3. Tao L., Asari K. V. Nonlinear enhancement of color images // SPIE Journal of Electronic Imaging. 2005. Vol. 14. P. 1.1–1.14.
4. Young T., Van Vliet L. J. Recursive Implementation of the gaussian filter // Signal Processing 44. Elsevier, 1995.
5. Yambor W. Analysis of PCA-based and Fisher discriminant-based image recognition algorithms : Technical Report CS-00-103, 2000.

A. I. Pakhirka

IMAGE ENHANCEMENT FOR FACE RECOGNITION SYSTEM

Three steps face recognition algorithm is proposed. We used the method of image enhancement based on high dynamic range compression, face detection algorithm based on skin color information, face recognition process based on principal components analysis method is considered as well.

Keywords: high dynamic range, face detection, face recognition.

© Пахирка А. И., 2010

УДК 681.332.53/519.676

Е. И. Алгазин

ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ ИНВАРИАНТНОЙ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ ПРИ НЕКОГЕРЕНТНОМ ПРИЕМЕ И НЕТОЧНОМ ОПРЕДЕЛЕНИИ ПОРОГОВ

Рассматривается инвариантная система обработки информации, основанная на линейном детекторе при неточном определении порогов. Проводится количественная оценка помехоустойчивости такой системы с последующим сравнением ее с помехоустойчивостью обычной бинарной системы при некогерентном приеме.

Ключевые слова: инвариантная система, помехоустойчивость.

Основным требованием к системам обработки информации является ее безыскаженная передача по каналам связи с переменными параметрами.

Существуют методы, которые сводятся к использованию АРУ, разнесенного приема, адаптивных методов с обучающим сигналом, систем с обратной связью.

Эти методы имеют как положительные, так и отрицательные стороны. Одним из отрицательных моментов указанных выше методов является трудность реализации алгоритмов передачи сигналов с многоуровневой амплитудной модуляцией.

В предлагаемой работе синтезирован алгоритм передачи многоуровневых амплитудно-модулированных сигналов по каналам с переменными параметрами и произведена количественная оценка помехоустойчивости при некогерентном приеме.

Постановка задачи. Имеется канал связи, ограниченный частотами f_n и f_b . Состояние канала связи определяется интервалом стационарности, внутри которого действие мультипликативной помехи описывается постоянством коэффициента передачи $k(t)$ на определенной частоте.

Алгоритм приема определяется несущей частотой, задаваемой как средняя частота канала, амплитуда которой промодулирована прямоугольными импульсами.

Требуется определить технические характеристики инвариантной системы передачи при неточном определении порогов.

Решение поставленной задачи. Каждый передаваемый блок будет содержать информационную часть и последовательность обучающих сигналов $S_{об}$.

На приемной стороне обучающие сигналы усредняются и используются для демодуляции информационной части блока.

При этом из-за изменения параметров канала связи информационные и обучающие сигналы зашумлены аддитивной помехой.

Для уменьшения влияния аддитивных шумов канала связи используется операция усреднения обучающих сигналов в каждом блоке [1].

Проведем анализ помехоустойчивости инвариантной системы (рис. 1), где использованы два канала обработки.

В первом канале, состоящем из синхронного детектора (СД) и первого решающего устройства (РУ1), производится оценка коэффициента передачи канала и дисперсии нормального шума, которые в дальнейшем используются для расчета порогов при демодуляции инвариантов.