

РЕИДЕНТИФИКАЦИЯ ЛЮДЕЙ НА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

Н.А. Ломов, С.В. Сидякин

Введение

В настоящее время всё большее распространение получают системы видеонаблюдения. Вместе с ростом числа камер растёт и объём образующихся видеоархивов. Одной из основных задач, решаемых в видеонаблюдении, является поиск конкретного человека в видеоархиве. В рамках этой задачи по указанному на одном кадре человеку необходимо найти все его появления в видеоархиве, тем самым отследив его перемещения. Эта задача называется задачей реидентификации. Она является довольно сложной, поскольку человек может появляться на видео при разном освещении и ракурсе съёмки, в различных позах и на изменчивом фоне.

При решении задачи реидентификации, как правило, предполагается, что задача обнаружения и локализации людей на видео решена предварительно, поэтому работа ведётся не с самим видеоархивом, а с набором кадров, каждый из которых содержит одного человека. Большинство методов реидентификации можно разделить на две стадии: (1) извлечение признаков для объектов [1, 2, 3, 4] и (2) сопоставление объектов [5, 6, 7]. Наиболее распространённый подход, применяемый на первой стадии – вычисление для изображения различных цветовых и текстурных дескрипторов. Примерами таких признаков могут служить цветовые гистограммы [1, 2, 3], результаты обработки изображения фильтрами Шмид и Габора [4] или признаки, полученные при помощи алгоритма SIFT [8]. Вторая стадия обычно сводится к настройке по обучающей выборке метрики в пространстве изображений и её использованию в метрических методах классификации.

Рассмотренный в этой статье метод реидентификации основан на алгоритме, предложенном в работе [9]. Главной особенностью этого алгоритма является извлечение признаков, связанных с набором из небольшого числа именованных цветов. Такие признаки обладают большой информативностью [10] и демонстрируют хорошую устойчивость к изменениям условий съёмки [11].

1. Расчёт вероятностных распределений

Основная идея предлагаемого подхода к вычислению признаков заключается в том, что мы считаем, что каждый пиксель изображения имеет вероятностное распределение, которое зависит от цвета пикселя, на множестве из 16 имён цветов. Список основных цветов и их HTML-кодов приведён ниже.

black	#000000	gray	#808080
maroon	#800000	red	#FF0000
green	#008000	lime	#00FF00
olive	#808000	yellow	#FFFF00
navy	#000080	blue	#0000FF
purple	#800080	fuchsia	#FF00FF
teal	#008080	aqua	#00FFFF
silver	#C0C0C0	white	#FFFFFF

Табл. 1. Палитра основных цветов

Для большей устойчивости модели к изменению условий освещения мы будем учитывать не точное значение цвета пикселя, а его индекс. Во всех используемых в алгоритме цветовых моделях цвет задаётся тремя координатами (a, b, c) , принадлежащими отрезку $[0, 1]$; на кодирование каждой компоненты отводится 8 бит, следовательно, каждая компонента может принимать 256 различных значений. Каждый цветовой канал мы равномерно разобьём на 32 уровня, что обеспечит разбиение цветового куба на 32^3 ячеек. Значение индекса цвета определяется ячейкой, к которой цвет относится, и вычисляется по следующей формуле:

$$I(a, b, c) = \left\lfloor \frac{255}{8} a \right\rfloor \times 32^2 + \left\lfloor \frac{255}{8} b \right\rfloor \times 32 + \left\lfloor \frac{255}{8} c \right\rfloor + 1. \quad (1.1)$$

Таким образом, индекс цвета может принимать значения от 1 до 32^3 , и к каждому индексу относятся $\left(\frac{256}{32}\right)^3 = 512$ цветов.

Распределения цветов, относящихся к одному и тому же индексу, совпадают. Пусть $Z = [z_1, z_2, \dots, z_{16}]$ – множество имён основных цветов, d – значение индекса пикселя, а $W = \{w_1, w_2, \dots, w_{512}\}$ – соответствующее этому индексу множество цветов. Тогда вероятность отнесения пикселя с индексом d к имени z вычисляется по формулам:

$$p(z|d) = \sum_{n=1}^{512} p(z|w_n) p(w_n|d), \quad (1.2)$$

$$p(w_n|d) = \frac{\exp(-\alpha \|w_n - \mu\|^2)}{\sum_{l=1}^{512} \exp(-\alpha \|w_l - \mu\|^2)}, \quad (1.3)$$

$$p(z|w_n) = \begin{cases} \frac{\exp(-\|z - w_n\|^2 / \frac{1}{K-1} \sum_{z_l \neq z} \|z_l - w_n\|^2)}{\sum_{p=1}^K \exp(-\|z_p - w_n\|^2 / \frac{1}{K-1} \sum_{z_s \neq z_p} \|z_s - w_n\|^2)}, & \text{если } z \in KNN(w_n), \\ 0, & \text{иначе} \end{cases} \quad (1.4)$$

где K обозначает число ближайших соседей, μ – среднее значение w_n ($n = 1, \dots, 512$). В формуле (1.4) z_p, z_l и z_s относятся к K ближайшим соседям цвета w_n . Под расстояниями между цветами понимаются расстояния между их координатами в Евклидовом пространстве.

На вход алгоритму будут подаваться изображения, использующие цветовую модель RGB. Однако признаки для изображения будут вычислены не только в этом пространстве, но и в трёх других: нормализованное rgb , $l_1l_2l_3$ и HSV. Поэтому вероятностные распределения необходимо рассчитать во всех четырёх цветовых пространствах, причём это можно сделать предварительно. При заданных координатах цвета в пространстве RGB можно однозначно определить координаты в трёх других пространствах. Преобразования производятся по формулам:

$$T_{rgb}(R, G, B) = \begin{cases} \left(\frac{R}{R+G+B}, \frac{G}{R+G+B}, \frac{B}{R+G+B} \right), & \text{если } R+G+B \neq 0 \\ (0, 0, 0), & \text{иначе} \end{cases} \quad (1.5)$$

Пусть $D = (R - G)^2 + (R - B)^2 + (G - B)^2$. Тогда

$$T_{l_1l_2l_3}(R, G, B) = \begin{cases} \left(\frac{(R - G)^2}{D}, \frac{(R - B)^2}{D}, \frac{(G - B)^2}{D} \right), & \text{если } D \neq 0 \\ \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right), & \text{иначе} \end{cases} \quad (1.6)$$

Пусть $MAX = \max(R, G, B)$, $MIN = \min(R, G, B)$. Тогда

$$H = \begin{cases} 0, & \text{если } MAX = MIN \\ \frac{1}{6} \times \frac{G - B}{MAX - MIN} + 0, & \text{если } MAX = R \text{ и } G \geq B \\ \frac{1}{6} \times \frac{G - B}{MAX - MIN} + 1, & \text{если } MAX = R \text{ и } G < B \\ \frac{1}{6} \times \frac{B - R}{MAX - MIN} + \frac{1}{3}, & \text{если } MAX = G \\ \frac{1}{6} \times \frac{R - G}{MAX - MIN} + \frac{2}{3}, & \text{если } MAX = B, \end{cases} \quad (1.7)$$

$$S = \begin{cases} 0, & \text{если } MAX = 0 \\ 1 - \frac{MIN}{MAX}, & \text{иначе} \end{cases} \quad (1.8)$$

$$V = MAX, \quad (1.9)$$

$$T_{HSV}(R, G, B) = (H, S, V). \quad (1.10)$$

При расчёте вероятностных распределений различие между цветовыми пространствами проявляется только в координатах основных цветов. Формула вычисления расстояния между цветами и координаты цветов, соответствующих определённому индексу, остаются прежними, несмотря на то что пространства имеют различную геометрию и в некоторых из них возможны далеко не все комбинации компонент. В пространствах rgb и $l_1l_2l_3$ координаты основных цветов могут совпадать. Предлагается не совершать по этому поводу никаких дополнительных действий, но стоит учесть, что в формуле (1.4) может возникнуть деление положительного числа на ноль. В этом случае будем условно считать, что:

$$\frac{a}{0} = \inf, \text{ если } a > 0, \quad (1.11)$$

$$\exp(-\inf) = 0. \quad (1.12)$$

В итоге для каждого пространства мы должны получить матрицу $P = (p_{ij})$ размера $32^3 \times 16$, где p_{ij} – вероятность отнесения цвета с индексом i к j -му имени в соответствующем пространстве.

2. Признаковое описание изображения

Пусть изображение имеет размеры $H \times W$. Разделим изображение на шесть вертикальных частей, присвоив каждому пикселю метку:

$$l_i = \left\lfloor 6 \times \frac{y_i + 1}{H} \right\rfloor + 1, \quad (2.1)$$

где y_i – y -координата i -го пикселя ($y_i \in \{1, \dots, H\}$). Вычислим распределения пикселей каждой части изображения по именам основных цветов. Будем считать, что пиксели неравноценны, и им присвоены веса ω_i . Сначала вычислим распределения пикселей каждой части изображения по индексам. Доля j -го индекса в части k равна:

$$q_{jk} = \frac{\sum_{\{i|l_i=k\}} [I_i = j] \omega_i}{\sum_{\{i|l_i=k\}} \omega_i}, \quad (2.2)$$

где I_i – индекс i -го пикселя.

При этом мы получим матрицу $Q = (q_{ij})$ размера $32^3 \times 6$. Тогда, вычислив матрицу

$$G = P^T Q, \quad (2.3)$$

мы найдём необходимые распределения. Матрица G имеет размер 16×6 , и g_{ik} задаёт долю i -го основного цвета в k -й части изображения.

Веса ω_i могут быть получены в результате сегментации изображения на объект (человека) и фон. Не используя сегментацию, можно присвоить пикселям следующие априорные веса:

$$\omega_i = \exp\left(-\frac{(y_i - \mu)^2}{2\sigma^2}\right), \quad (2.4)$$

$$\text{где } \mu = \frac{W}{2}, \text{ а } \sigma = \frac{W}{4}.$$

Как было отмечено выше, не во всех используемых цветовых пространствах цвета полностью заполняют куб. Это приводит к тому, что некоторые значения индексов оказываются невозможными, и их доли всегда будут равны нулю. Можно значительно сократить признаковое описание объекта, исключив из него заведомо нулевые признаки. Зададим характеристическую функцию индекса в следующем виде:

$$\chi(j) = \begin{cases} 1, & \text{если } \exists(R, G, B) : I(T(R, G, B)) = j \\ 0, & \text{иначе} \end{cases} \quad (2.5)$$

Значения характеристической функции для цветовых пространств можно рассчитать предварительно, перебрав все возможные комбинации компонент R , G и B . Зная значения характеристической функции, можно оставлять в матрицах P и Q из формулы (2.3) только те строки, характеристическая функция номеров которых равна 1, что сократит вычисления и сэкономит память. Число возможных индексов для каждого пространства приведено ниже.

Пространство	Число индексов
RGB	32768
Rgb	1489
$l_1 l_2 l_3$	127
HSV	31249

Табл. 2. Число возможных индексов для цветовых пространств

В работе [9] предлагается также включать в признаковое описание изображения взвешенные цветовые гистограммы. В случае разбиения каждого канала на 32 уровня гистограммы совпадают со взвешенным распределением пикселей изображения по индексам:

$$h_j = \frac{\sum_i [I_i = j] \omega_i}{\sum_i \omega_i}. \quad (2.6)$$

Однако, даже если исключить заведомо нулевые ячейки гистограммы, описание каждого объекта будет содержать очень много признаков: $16 \times 6 \times 4 + 32768 + 1489 + 127 + 31249 = 66017$, а при наличии сегментации признаков будет вдвое больше. Забегая вперёд, отметим, что к матрице объектов–признаков будет применён метод главных компонент, и столь внушительный размер этой матрицы приведёт к ощутимым временным затратам. Поэтому мы используем другой подход: вычислим гистограммы с меньшим числом ячеек, но отдельно для каждой части изображения. В пространствах RGB и HSV каждый канал разобьём на 8 уровней, а в пространствах rgb и $l_1l_2l_3$ – на 16. Номер ячейки гистограммы вычисляется по формуле:

$$S(a, b, c) = \left\lfloor \frac{255q}{256} a \right\rfloor \times q^2 + \left\lfloor \frac{255q}{256} b \right\rfloor \times q + \left\lfloor \frac{255q}{256} c \right\rfloor + 1, \quad (2.7)$$

где q – число градаций цветового канала. Нам вновь потребуется понятие характеристической функции, на этот раз для ячейки гистограммы:

$$\xi(j) = \begin{cases} 1, & \text{если } \exists(R, G, B) : S(T(R, G, B)) = j \\ 0, & \text{иначе} \end{cases} \quad (2.8)$$

Пространство	Число индексов
RGB	512
Rgb	361
$l_1l_2l_3$	61
HSV	512

Табл. 3. Число возможных ячеек гистограммы для цветовых пространств

Тогда, отбросив в гистограммах заведомо нулевые ячейки, мы получим признаковое описание объекта, состоящее из $(16 \times 4 + 512 + 361 + 61 + 512) \times 6 = 9060$ признаков (табл. 3).

3. Метод KISSME

В задачах идентификации широко используются метрические методы, основанные на сравнении новых объектов с уже имеющимися эталонами.

Пусть признаковые описания x_i и x_j пары объектов представляют собой вектора-столбцы вещественных чисел одинаковой длины. В качестве меры расстояния между объектами мы будем использовать расстояние Махаланобиса [7]:

$$d_M^2(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j), \quad (3.1)$$

где M – положительно определённая матрица.

Нам нужно подобрать такую матрицу M , для которой расстояния между объектами из одного класса (изображениями одного и того же человека) получались бы как можно меньше, а расстояния между объектами из разных классов – как можно больше. Фактически, имея пару объектов, мы должны решить, судя по расстоянию между ними, относятся они к одному классу или нет:

$$\tau(x_i, x_j) = \begin{cases} 1, & \text{если } d_M^2(x_i, x_j) < a \text{ (делаем вывод, что класс один и тот же)} \\ 0, & \text{если } d_M^2(x_i, x_j) \geq a \text{ (делаем вывод, что классы разные).} \end{cases} \quad (3.2)$$

Задаче можно придать вероятностную интерпретацию [8]. Пусть у нас есть вероятностное пространство пар объектов, классы которых совпадают и пространство пар объектов, классы которых отличаются. Выдвинем две альтернативные гипотезы:

H_0 – объекты принадлежат к разным классам;

H_1 – объекты принадлежат к одному и тому же классу.

Тогда вероятность появления пары объектов из первого пространства можно обозначить за $p(x_i, x_j | H_0)$, пары объектов из второго – за $p(x_i, x_j | H_1)$. С точки зрения статистического вывода оптимальным будет решение, основанное на отношении этих вероятностей:

$$\delta(x_i, x_j) = \log \left(\frac{p(x_i, x_j | H_0)}{p(x_i, x_j | H_1)} \right). \quad (3.3)$$

При высоком значении $\delta(x_i, x_j)$ мы отдаём предпочтение гипотезе H_0 , при низком значении – гипотезе H_1 .

Предположим, что пространства в гипотезах H_0 и H_1 задаются распределениями из одного и того же параметрического семейства. Удобнее рассматривать не пару признаковых описаний объектов x_i и x_j а разность между ними x_{ij} :

$$\delta(x_i, x_j) = \log \left(\frac{p(x_{ij}|H_0)}{p(x_{ij}|H_1)} \right) = \log \left(\frac{f(x_{ij}|\theta_0)}{f(x_{ij}|\theta_1)} \right), \quad (3.4)$$

где $f(x|\theta)$ – плотность вероятности случайной величины x , взятой из распределения с параметрами θ .

Мы будем считать, что случайная величина x_{ij} имеет многомерное нормальное распределение, матрица ковариации которого зависит от того, совпадают ли классы. В формуле для плотности вероятности у этого распределения под экспонентой стоит как раз расстояние Махаланобиса до матожидания случайной величины.

Оценка сходства пары объектов имеет вид:

$$\delta(x_i, x_j) = \log \left(\frac{\frac{1}{\sqrt{2\pi|\Sigma_{y_{ij}=0}|}} \exp(-\frac{1}{2} x_{ij}^T \Sigma_{y_{ij}=0}^{-1} x_{ij})}{\frac{1}{\sqrt{2\pi|\Sigma_{y_{ij}=1}|}} \exp(-\frac{1}{2} x_{ij}^T \Sigma_{y_{ij}=1}^{-1} x_{ij})} \right), \quad (3.5)$$

где $y_{ij}=1$, если объекты принадлежат к одному и тому же классу, и $y_{ij}=0$, если объекты принадлежат к разным классам.

Матрицы ковариации $\Sigma_{y_{ij}=0}$ и $\Sigma_{y_{ij}=1}$ оценим по методу максимального правдоподобия:

$$\Sigma_{y_{ij}=1} = \frac{1}{[y_{ij}=1]} \sum_{y_{ij}=1} (x_i - x_j)(x_i - x_j)^T, \quad (3.6)$$

$$\Sigma_{y_{ij}=0} = \frac{1}{[y_{ij}=0]} \sum_{y_{ij}=0} (x_i - x_j)(x_i - x_j)^T. \quad (3.7)$$

Отметим, что (x_i, x_j) и (x_j, x_i) при $i \neq j$ – это разные пары объектов, поэтому матожидание x_{ij} равно нулю.

Логарифмируя выражение в скобках из формулы (4.5), получаем:

$$\delta(x_{ij}) = \frac{1}{2} x_{ij}^T \Sigma_{y_{ij}=1}^{-1} x_{ij} + \log \left(|\Sigma_{y_{ij}=1}| \right) - \frac{1}{2} x_{ij}^T \Sigma_{y_{ij}=0}^{-1} x_{ij} - \log \left(|\Sigma_{y_{ij}=0}| \right). \quad (3.8)$$

Сдвигая эту величину на константу, имеем:

$$\delta(x_{ij}) = x_{ij}^T (\Sigma_{y_{ij}=1}^{-1} - \Sigma_{y_{ij}=0}^{-1}) x_{ij}. \quad (3.9)$$

Мы получили нечто очень похожее на формулу для расстояния Махаланобиса, которая, напомним, имеет вид:

$$d_M^2(x_i, x_j) = (x_i - x_j)^T M(x_i - x_j).$$

Однако матрица

$$\hat{M} = \Sigma_{y_{ij}=1}^{-1} - \Sigma_{y_{ij}=0}^{-1}. \quad (3.10)$$

может и не быть положительно определённой, то есть иметь отрицательные или нулевые собственные значения. В этом случае нужно найти проекцию \hat{M} матрицы \hat{M} на конус положительно определённых матриц. Эта процедура заключается в замене неположительных собственных значений на небольшое положительное число с сохранением всех собственных векторов.

Таким образом, оптимизация метрики по методу KISSME сводится к вычислению матриц ковариации по формулам (3.6) и (3.7) и преобразованию матрицы \hat{M} к положительно определённой, если это необходимо. В имеющейся реализации метода для обучения используются все возможные «схожие» пары изображений и такое же число «несхожих» пар, которые выбираются случайным образом.

4. СМС-кривая

В методах, решающих задачу идентификации, могут использоваться самые разные способы получения признакового описания объекта и задания меры близости этих объектов. Однако, как правило, каждый класс в таких задачах характеризуется небольшим числом эталонов (зачастую единственным эталоном), поэтому итоговое решение чаще всего принимается на основании метода ближайших соседей. В случае большого числа классов стандартные критерии качества, например, доля правильных ответов алгоритма, могут оказаться нерепрезентативными, поскольку, во-первых, результат, скорее всего, окажется невысоким, а во-вторых, во многих прикладных задачах обращают внимание не на единственного ближайшего соседа, а на k ближайших соседей.

Поэтому в таких задачах считается приемлемой ситуация, когда «парный» объект не является ближайшим соседом рассматриваемого, но попадает в число k его ближайших соседей. СМС (Cumulative Match Characteristic)-кривая отражает зависимость доли ситуаций, когда нужный объект попадает в область рассмотрения, от числа ближайших соседей.

Пусть в нашей задаче имеется n пар объектов (x_{i1}, x_{i2}) , $1 \leq i \leq n$, причём объекты из одной пары принадлежат к одному классу, а объекты из разных пар – к разным. Для произвольного объекта, идущего в паре вторым, отсортируем объекты $x_{11}, x_{21}, \dots, x_{n1}$:

$$\rho(x_{j2}, x_j^{(1)}) \leq \rho(x_{j2}, x_j^{(2)}) \leq \dots \leq \rho(x_{j2}, x_j^{(n)}).$$

Номер объекта x_{i1} в подобной последовательности, построенной для x_{i2} , обозначим за $r(x_{i1})$. Тогда СМС ранга k вычисляется по формуле

$$CMC(k) = \frac{\sum_{i=1}^n [r(x_{i1}) \leq k]}{n}. \quad (4.1)$$

Для оценки качества работы алгоритма строится график зависимости СМС от k , который называется СМС-кривой. Особую популярность эта характеристика оценки качества получила в задачах биометрии.

5. Эксперименты

Для оценки качества реализованного метода были использованы два набора данных: VIPeR¹ и PRID 450S², изображения в которых получены при различном освещении и различных точках съёмки. Набор данных VIPeR содержит 632 пары изображений, PRID 450S – 450 пар. Значение α из формулы (1.3) было установлено равным 1. В каждом эксперименте данные были преобразованы методом главных компонент, для вычисления расстояний между объектами использовались первые 70 из них. Метод KISSME сравнивался со следующими способами задания метрики:

- 1) Mahalanobis Distance Learning (MAHAL)
- 2) Identity (IDENT) – расстояние Махаланобиса с единичной матрицей
- 3) Information Theoretic Metric Learning (ITML)
- 4) Large Margin Nearest Neighbor (LMNN)

Подробное описание этих метрик можно найти в работе [2]. На одной половине доступных пар объектов метрический метод обучался, по другой – оценивал качество. Каждый эксперимент повторялся 11 раз, на графиках приведены медианные характеристики.



Рис.1. Примеры сегментации с помощью метода Stel Component Analysis

Для сегментации изображений был использован метод, предложенный в работе [12] с параметрами: число структурных элементов – 2, число компонент – 3, число итераций алгоритма – 50. Хорошая работа этого метода сегментации возможна при выполнении предположений о различимости пикселей разных сегментов как с точки зрения их пространственного расположения, так и с точки зрения их цвета. Если первое предположение в нашем случае справедливо, то второе из-за разнообразного фона сложно считать выполненным, поэтому результаты работы метода сегментации не всегда оказываются удовлетворительными (рис. 1).

Для набора данных VIPeR было получено четыре различных признаковых описания объектов:

- (a) Без учёта стандартных цветовых гистограмм и без учёта сегментации (384 признака)
- (b) С учётом гистограмм и без учёта сегментации (9060 признаков)
- (c) С учётом сегментации без гистограмм (768 признаков)

¹ Набор данных находится в открытом доступе по адресу <http://vision.soe.ucsc.edu/node/178>

² Набор данных находится в открытом доступе по адресу <http://lrs.icg.tugraz.at/download.php>

(d) С учётом и сегментации, и гистограмм (18120 признаков)

Результаты экспериментов (рис. 2) показывают, что если число ближайших соседей не превосходит 50, то метод KISSME всегда оказывается предпочтительнее аналогов независимо от способа формирования признакового описания. Добавление гистограммы повышает качество (рис. 3), особенно для большого числа ближайших соседей, однако не самая удачная сегментация может привести к снижению результатов.

Набор данных PRID 450S содержит готовые маски изображений, причём двух видов: полученные вручную и в результате работы алгоритма сегментации. Поэтому для этого набора данных было вычислено шесть признаковых описаний объектов:

- (a) Без учёта гистограмм и без учёта сегментации (384 признака)
- (b) С учётом гистограмм и без учёта сегментации (9060 признаков)
- (c) С учётом ручной сегментации без гистограмм (768 признаков)
- (d) С учётом ручной сегментации и гистограмм (18120 признаков)
- (e) С учётом автоматической сегментации без гистограмм (768 признаков)
- (f) С учётом автоматической сегментации и гистограмм (18120 признаков)

Результаты экспериментов (рис. 4) показывают, что метод KISSME снова является лучшим выбором при числе ближайших соседей не более 50. Если сегментация изображений близка к идеальной, её использование может значительно повысить качество сопоставления (рис. 5). Варианты, использующие гистограммы, снова показывают более высокие результаты.

Отметим также, что качество работы варианта без гистограммы и без сегментации выше на наборе данных VIPeR, хотя этот набор данных считается более сложным. Можно сделать вывод, что разнообразный фон (как в VIPeR) является для метода меньшей проблемой, нежели изменение освещения (как в PRID 450S).

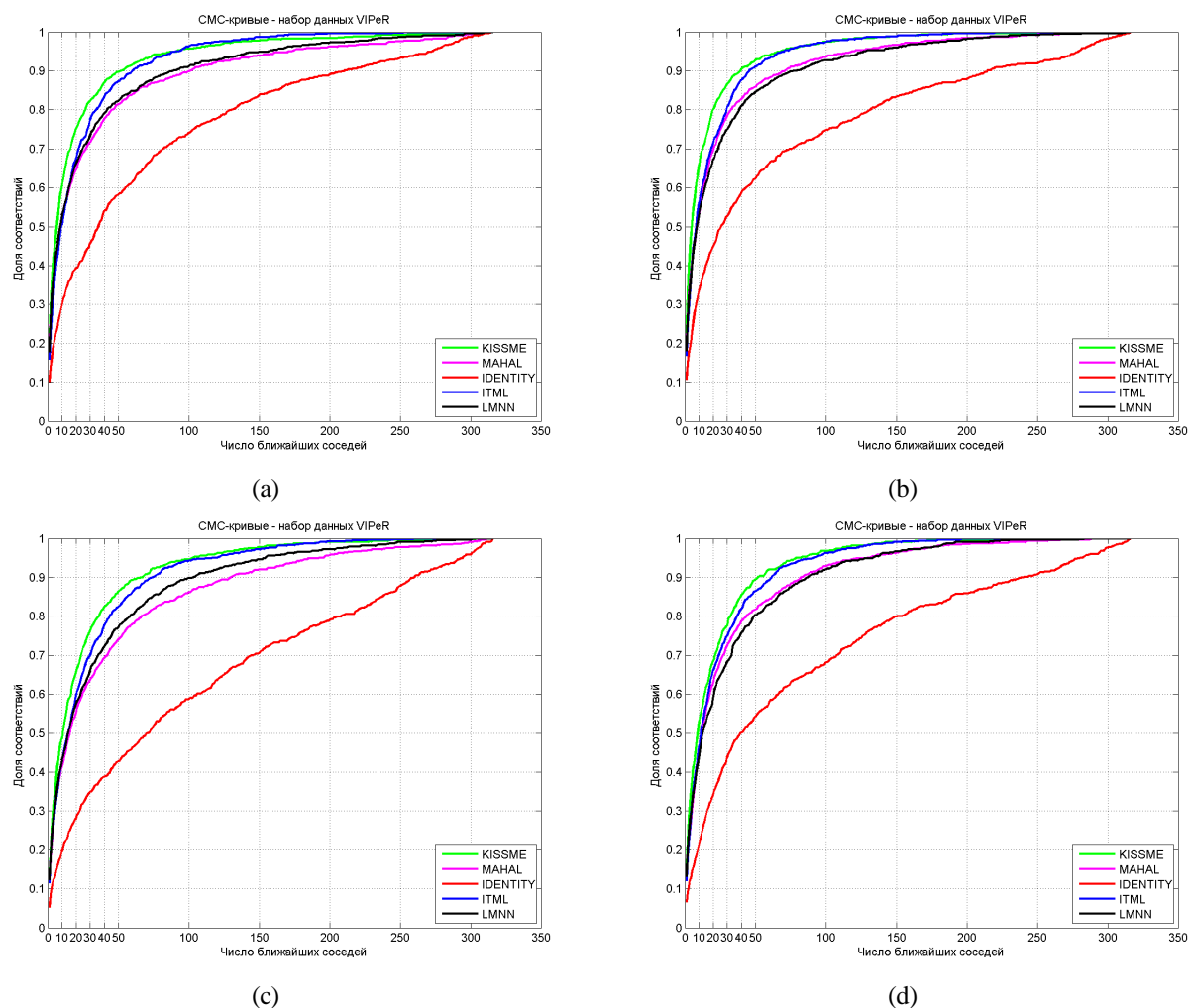


Рис. 2. Качество работы методов на наборе данных VIPeR: (a) без учёта цветowych гистограмм и без учёта сегментации (b) с учётом только гистограмм (c) с учётом только сегментации (d) с учётом и сегментации, и гистограмм

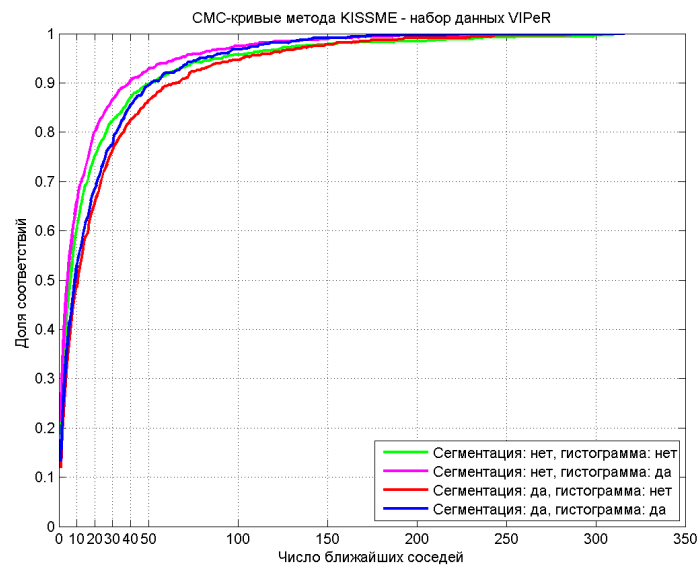
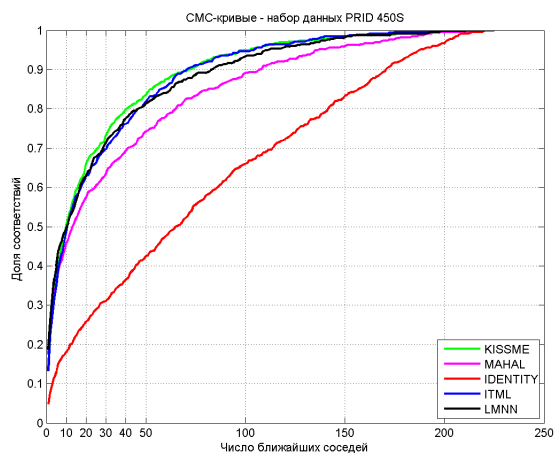
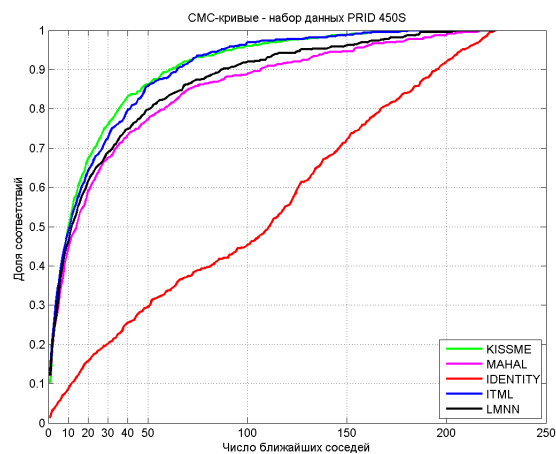


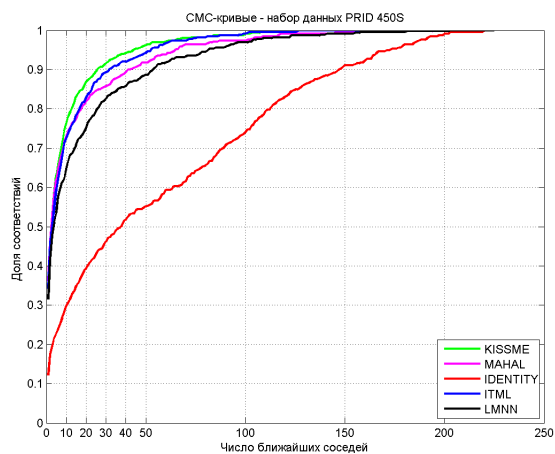
Рис. 3. Сводный график результатов метода KISSME на наборе данных VIPeR



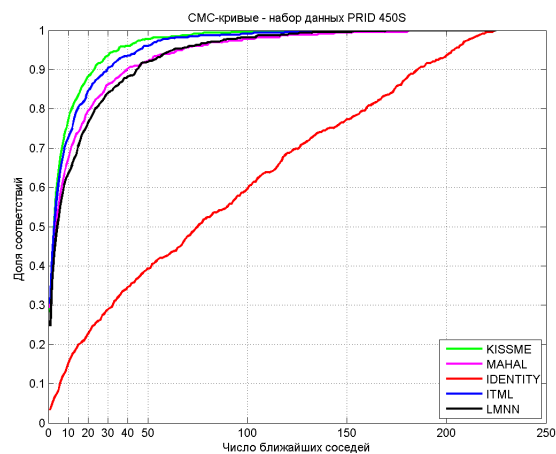
(a)



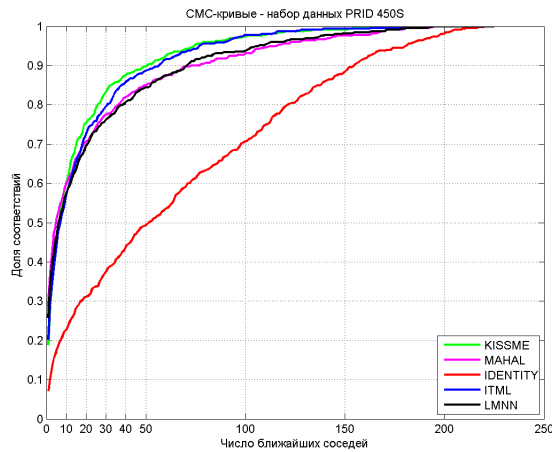
(e)



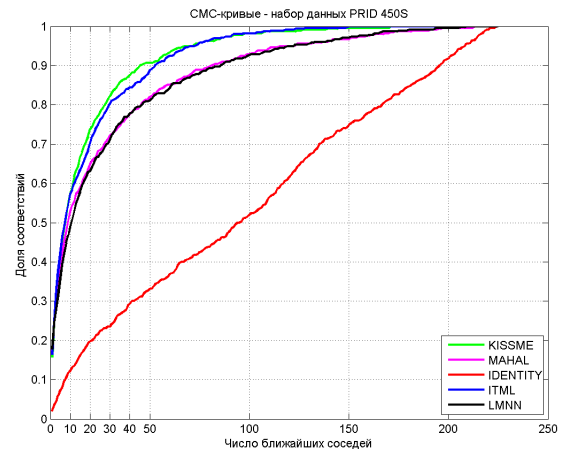
(c)



(b)



(d)



(f)

Рис. 4. Качество работы методов на наборе данных PRID 450S: (a) без учёта гистограмм и без учёта сегментации (b) с учётом гистограмм и без учёта сегментации (c) с учётом ручной сегментации без гистограмм (d) с учётом ручной сегментации и гистограмм (e) с учётом автоматической сегментации без гистограмм (f) с учётом автоматической сегментации и гистограмм

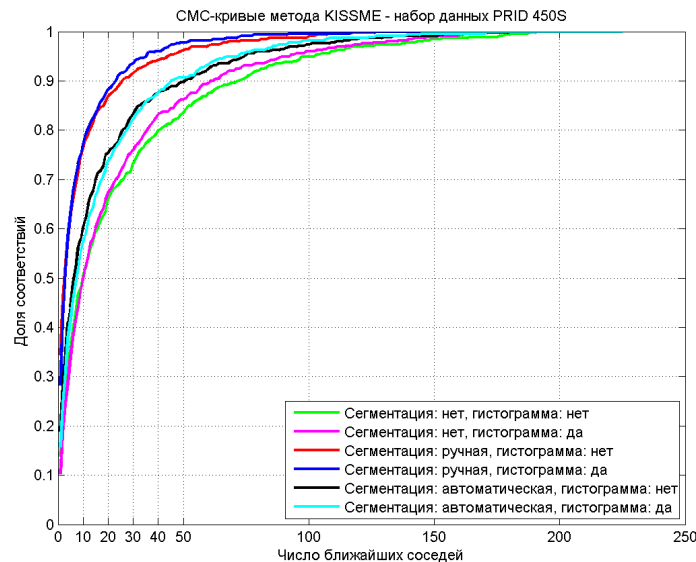


Рис. 5. Сводный график результатов метода KISSME на наборе данных PRID 450S.

Заключение

В работе рассматривается метод формирования признакового описания изображения, устойчивый к изменению условий освещения и ракурса съёмки. Описан способ расчёта признакового описания, позволяющий произвести большую часть вычислений предварительно и избежать вычисления неиспользуемых величин. Дескриптор был использован в задаче реидентификации людей на видеопоследовательностях. Рассмотрена модификация исходного алгоритма, позволяющая сократить объём вычислений без заметного снижения качества распознавания. Проведены эксперименты, использующие различные подмножества признаков и различные способы задания метрики. Высокое качество работы метода подтверждает результаты, приведённые в работе [9].

Метод реидентификации можно использовать для сегментированных изображений, поэтому один из путей его развития – поиск наиболее подходящего алгоритма сегментации. Кроме того, взгляд на задачу идентификации как на задачу классификации с двумя классами в вероятностной постановке позволяет применить к признаковому описанию распространённые техники машинного обучения и прикладной статистики.

Предлагаемый дескриптор может использоваться во многих задачах анализа и распознавания изображений.

Литература

1. **Farenzena, M.** Person reidentification by symmetry-driven accumulation of local features / L. Bazzani, A. Perina, V. Murino, M. Cristani // IEEE Conference on Computer Vision and Pattern Recognition, 2010. – P. 2360-2367.
2. **Gray, D.** Viewpoint invariant pedestrian recognition with an ensemble of localized features / H. Tao // Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302, Springer, Heidelberg, 2008. – P. 262-275
3. **Ma, B.** Local descriptors encoded by fisher vectors for person re-identification / Y. Su, F. Jurie // Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012 Ws/Demos, Part I. LNCS, vol. 7583, Springer, Heidelberg, 2012. – P. 413-422
4. **Liu, C.** Person re-identification: What features are important? / S. Gong, C.C. Loy, X. Lin // Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012 Ws/Demos, Part I. LNCS, vol. 7583, Springer, Heidelberg, 2012. – P. 391-401
5. **Hirzer, M.** Person re-identification by efficient impostor based metric learning / P.M. Roth, H. Bischof // Advanced Video and Signal-Based Surveillance (AVSS), 2012. – P. 203–208
6. **Köstinger, M.** Large Scale Metric Learning from Equivalence Constraints / M. Hirzer, P. Wohlhart, P. M. Roth, H. Bischof // IEEE Conference on Computer Vision and Pattern Recognition, 2012. – P. 2288-2295.
7. **Roth, P.M.** Mahalanobis Distance Learning for Person Re-Identification / M. Hirzer, M. Köstinger, C. Beleznai, H. Bischof // Advances in Computer Vision and Pattern Recognition, 2014. – P. 247-267.
8. **D. G. Lowe.** Object recognition from local scale-invariant features // International Conference on Computer Vision, 1999. – P. 1150-1157.
9. **Yang, Y.** Salient Color Names for Person Re-identification / J. Yang, J. Yan, S. Liao, D. Yi, S.Z. Li // European Conference on Computer Vision (ECCV), 2014. – P. 536-551.
10. **van de Weijer, J.** Learning color names for real-world applications / C. Schmid, J. Verbeek, D. Larlus // IEEE Trans. on Image Processing. 18(7), 2009. – P. 1512-1523.
11. **van de Weijer, J.** Applying color names to image description / C. Schmid // IEEE International Conference on Image Processing (ICIP), 2007. – P. 493-496.
12. **Jojic, N.** Stel component analysis: Modeling spatial correlations in image class structure / A. Perina, N. Jojic, M. Cristani, V. Murino // IEEE Conference on Computer Vision and Pattern Recognition, 2009. – P. 2044-2051.