

ОПИСАНИЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ МОДЕЛЬНО-ОРИЕНТИРОВАННЫХ ДЕСКРИПТОРОВ

В.В. Мясников^{1,2}

¹ Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия,

² Институт систем обработки изображений РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, Самара, Россия

Аннотация

В работе предлагается подход к построению описания изображений, основанный на их представлении с использованием набора модельно-ориентированных дескрипторов. Каждый дескриптор характеризует «схожесть» анализируемого изображения, представленного в виде комплекснозначного поля градиента, с некоторой предварительно выбранной моделью этого дескриптора. Для синтеза моделей дескрипторов в работе предлагается использовать метод главных компонент, применяемых для множества реализаций комплекснозначных градиентных полей. Предлагаемый подход позволяет получить описание комплекснозначного поля градиента анализируемого изображения в виде набора вещественных величин – признаков дескрипторов – из интервала $[0, 1]$, которые одновременно характеризуют и фазовую, и абсолютную составляющую градиента изображения. Эффективность предлагаемого подхода демонстрируется на примере решения задачи распознавания лиц путем сопоставления с решением-прототипом (методом «собственных лиц»), использующим непосредственно полутонные изображения. Сравнение выполняется с использованием классификатора по ближайшему соседу.

Ключевые слова: цифровые изображения, признаки, модельно-ориентированные дескрипторы, анализ, распознавание.

Цитирование: Мясников, В.В. Описание изображений с использованием модельно-ориентированных дескрипторов / В.В. Мясников // Компьютерная оптика. – 2017. – Т. 41, №6. – С. 888-896. – DOI: 10.18287/2412-6179-2017-41-6-888-896.

Введение

Важным этапом решения широкого класса задач обработки, анализа, понимания и поиска цифровых изображений является выбор *признаков*, часто именуемых в современной литературе *дескрипторами* [1]. Существуют различные классификации признаков, отражающие специфику подходов к их получению. Например, в монографиях [2, 3] признаки подразделяются на геометрические, топологические, вероятностные и спектральные. В работах [1, 4] локальные дескрипторы подразделяют на вероятностные, спектральные, спектрально-частотные и дифференциальные. Независимо от используемой классификации существуют признаки, которые могут сочетать в себе свойства, характерные для разных категорий. К таким дескрипторам относится *модельно-ориентированный дескриптор*, предложенный и развиваемый автором [4–6]. Его отличием от существующих является наличие априори заданного (или заранее определенного по обучающему множеству изображений) вероятностного распределения поля градиента, характеризующего *модель анализируемого изображения* и/или решаемую задачу. Признаком дескриптора является нормированная величина плотности вероятностей с аргументом в виде конкретного поля градиента. Она оказывается зависима как от реализации (конкретного изображения), так и от модели (распределения вероятностей), что позволило охарактери-

зовать этот дескриптор как *модельно-ориентированный* (МО).

Характеризуя предлагаемый подход, следует отметить его особенности и отличия от других известных решений. Градиент в задачах распознавания и поиска изображений используется достаточно широко. Наиболее известными дескрипторами, основанными на вычислении градиента являются HOG [7] и SIFT [8] дескрипторы. Однако указанные и большинство других известных дескрипторов и признаков [1] не используют градиент изображения как векторное поле (используемая ниже интерпретация – комплекснозначное изображение/сигнал) полноценно. А именно: фазовая или амплитудная составляющая градиента в известных решениях игнорируется полностью или частично. Наиболее яркими примерами такого подхода являются работы [9–12]. В частности, если в [9] фазовая составляющая градиента игнорируется полностью, то в работах [10–12], наоборот, игнорируется амплитуда. Работа [12] вообще полностью посвящена методам и алгоритмам обработки специального подкласса изображений, характеризующегося только направлением изменения функции яркости, – так называемого *поля направлений*.

Попытки полноценно использовать многозначные (в том числе комплекснозначные) изображения предпринимались исследователями в различных направлениях. Например, в работе [13] был предложен специализированный многокомпонентный анализ; существует целый цикл работ по использо-

ванию алгебры Клиффорда при работе с многокомпонентными изображениями (стр. 231–254 в [14]); статистическая обработка и анализ комплексных сигналов/изображений рассматривался в [15 – 18], а эссе по использованию гиперспектральных алгебр при работе с сигналами представлено в [19]. К сожалению, утверждать о возможности полноценной работы с многозначными (в данном конкретном случае комплекснозначными) изображениями в целях их анализа и распознавания невозможно. Данная работа представляет один из возможных подходов к использованию поля градиента для решения задач анализа и распознавания цифровых изображений.

Работа построена следующим образом. В первом параграфе даются основные определения и приводится краткое описание МО-дескриптора поля градиента, более подробное изложение можно найти в авторской работе [4], где также представлены несколько простейших параметрических моделей распределения вероятностей поля градиента и соответствующих признаков МО-дескриптора. Во втором параграфе представлена общая схема предлагаемого метода описания изображений с использованием МО-дескриптов. В третьем параграфе указанная общая схема детализируется: указываются алгоритмы предварительной обработки, вводится новая параметрическая модель МО-дескриптора и приводятся соответствующие ей признаки, предложен способ синтеза моделей МО-дескриптора. В заключительном параграфе представлены результаты исследования эффективности предлагаемого метода описания изображений на примере решения задачи распознавания лиц. Наконец, в заключение работы приведены выводы, благодарности и список использованной литературы.

1. Модельно-ориентированный дескриптор поля градиента и его признаки

Пусть величина

$$g(n_1, n_2), \quad (n_1, n_2) \in D,$$

$$D \equiv \{(n_1, n_2) : n_1 = 0, N_1 - 1, n_2 = 0, N_2 - 1\}$$

обозначает цифровое поле градиента (двухкомпонентное изображение), в котором каждый отсчет содержит информацию о градиенте в следующем виде: величина $|g(n_1, n_2)|$ определяет модуль градиента в соответствующей позиции, а направление (ориентацию по отношению к осям координат) вектора градиента в этой позиции определяет величина $\text{arg}(g(n_1, n_2))$. Анализируемый фрагмент цифрового изображения целиком обозначается далее как $\bar{g} \equiv \{g(n_1, n_2)\}_{(n_1, n_2) \in D}$, а \bar{G} – случайное градиентное поле с плотностью вероятностей $p_{\bar{G}}(\cdot)$, характеризующей модель наблюдения и/или решаемую задачу. Модельно-ориентированный дескриптор (МО-дескриптор) поля градиента \bar{g} (дескриптор по отношению к модели \bar{G}) определяется в работе [4] как вектор из 11 компонентов вида

$$\begin{pmatrix} p_{\bar{G}}(\bar{g}) \\ \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \min_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}) \\ \max_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}) \\ \min_{\text{arg } \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{\text{arg } \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} \min_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}) \\ \min_{\text{arg } \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \min_{|\bar{g}|} \max_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} \max_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}) \end{pmatrix}, \tag{1}$$

каждая из которых отражает схожесть (в абсолютной величине) реализации градиентного поля с моделью по модулю, фазе и/или совместно. Причем для компонент этого вектора оказываются справедливыми ряд очевидных соотношений, например,

$$\min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \leq p_{\bar{G}}(\bar{g}) \leq \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}). \tag{2}$$

Учитывая соотношения (2), признаки МО-дескриптора были определены как относительные положения компонент дескриптора (например, $p_{\bar{G}}(\bar{g})$) в интервале от их минимального до максимального значений. Всего конструируются семь основных признаков, конкретные выражения для которых приведены в работах [4, 5]. Вид первых трех признаков:

$$\begin{aligned} \vartheta_0 &= \frac{\varphi(p_{\bar{G}}(\bar{g})) - \varphi(\min_{|\bar{g}|} p_{\bar{G}}(\bar{g}))}{\varphi(\max_{|\bar{g}|} p_{\bar{G}}(\bar{g})) - \varphi(\min_{|\bar{g}|} p_{\bar{G}}(\bar{g}))}, \\ \vartheta_1 &= \frac{\varphi(p_{\bar{G}}(\bar{g})) - \varphi(\min_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}))}{\varphi(\max_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g})) - \varphi(\min_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g}))}, \\ \vartheta_2 &= \frac{\varphi(p_{\bar{G}}(\bar{g})) - \varphi(\min_{\text{arg } \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}))}{\varphi(\max_{|\bar{g}|} \max_{\text{arg } \bar{g}} p_{\bar{G}}(\bar{g})) - \varphi(\min_{\text{arg } \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}))}, \end{aligned} \tag{3}$$

где $\varphi: \mathbf{R}_+ \rightarrow \mathbf{R}$ – произвольная монотонно возрастающая функция. Признаки ϑ_i характеризуют степень схожести анализируемого градиентного поля на потенциально возможные реализации случайного поля \bar{G} (на модель). Большие значения означают большую схожесть, а меньшие – меньшую. При этом если ϑ_0 , например, определяет степень схожести поля модуля градиента, то ϑ_1 – степень схожести фазового рисунка и т.п. На все множество этих признаков ниже будем ссылаться как на вектор $\bar{\vartheta} \equiv (\vartheta_0, \vartheta_1, \dots, \vartheta_6)^T$. В работе [4] доказывается следующее утверждение:

$$\forall i = \overline{0,6} \quad 0 \leq \vartheta_i \leq 1.$$

Имея ряд заданных случайных полей градиента \overline{G}_k ($k = \overline{0, K-1}$), для конкретного анализируемого изображения можно рассчитать ряд нормированных по диапазону признаков, например ϑ_k^0 ($k = \overline{0, K-1}$), которые в совокупности его характеризуют. На основании полученных значений может быть произведен анализ обрабатываемого изображения и его распознавание. Примеры для нескольких простейших моделей случайных полей градиента, удобных для распознавания «каркасных» объектов, можно найти в предшествующих работах [4, 5].

2. Общая схема предлагаемого метода описания изображений с использованием МО-дескрипторов

Общая схема описания изображения с целью их анализа и распознавания с использованием МО-дескрипторов представлена на рис. 1 (компоненты схемы, ограниченные пунктирной линией).

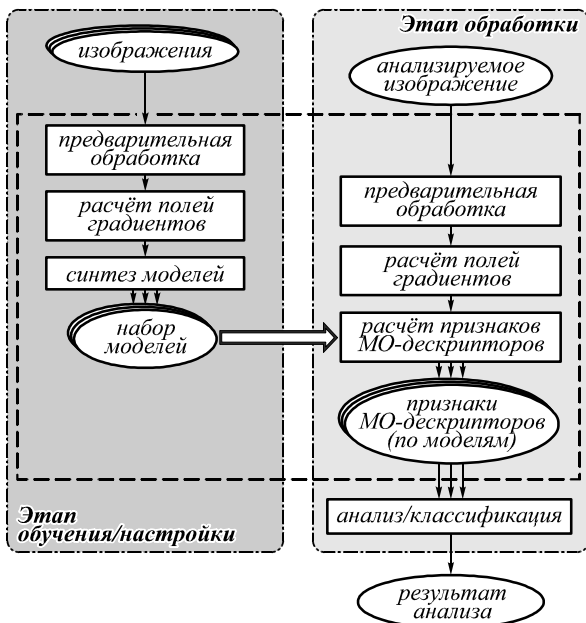


Рис. 1. Общая схема метода описания изображений с использованием МО-дескрипторов (пунктирная линия)

По аналогии со многими задачами распознавания и классификации изображений, процесс может быть поделен на два этапа: этап настройки/обучения и этап обработки (этапы на схеме ограничены штрихпунктирными линиями). На этапе *настройки* по обучающему множеству предварительно обработанных примеров-изображений рассчитываются соответствующие им поля градиентов, по которым синтезируется (формируется и настраивается) набор моделей дескрипторов – случайных градиентных полей \overline{G}_k ($k = \overline{0, K-1}$). В рамках данной работы для синтеза моделей предлагается использовать метод главных компонент (PCA – Principal Component Analysis) над набором комплекснозначных гради-

ентных полей [15–18, 20]. Из альтернативных способов, потенциально допускающих использование известных математических методов, можно указать также линейный дискриминантный анализ, метод независимых компонент, EM-алгоритм. Их сравнительный анализ выходит за рамки настоящей работы и может стать предметом отдельного исследования, как и разработка оригинального метода. Настраиваемые модели выступают на этапе обработки в качестве своеобразного «базиса» для анализируемых изображений. То есть позволяют описать анализируемое изображение в виде набора вещественных признаков ϑ_k^0 ($k = \overline{0, K-1}$) в диапазоне $[0, 1]$, характеризующих это изображение с точки зрения каждой из моделей. Эти признаки и используются в дальнейшем для анализа или классификации искомого изображения. А учитывая развитый математический аппарат обработки вещественных данных, анализ и/или классификация могут быть реализованы с использованием хорошо известных методов и алгоритмов: метода опорных векторов, деревьев решений и регрессии, нейронных сетей, ближайших соседей и др. (в данной работе при проведении экспериментальных исследований в качестве классификатора используется метод ближайшего соседа).

Следует отметить, что предлагаемый метод описания и анализа изображения использует идею *беспризнаковых методов распознавания* [21], поскольку анализируемый объект описывается через схожесть его с другими объектами, задаваемыми в виде случайных полей градиента \overline{G}_k ($k = \overline{0, K-1}$).

Также следует отметить связь предлагаемого метода описания с рядом известных решений, в частности с дескрипторами HOG (histogram of gradients) [7] и SIFT (Scale Invariant Feature Transform) [8]. Оба эти дескриптора используют вектор (гистограмму), каждый отсчет которого определяет частоту появления в анализируемом изображении (для локальной обработки – фрагменте изображения) градиента определенной ориентации/фазы. Такой дескриптор является достаточно удобным и мощным средством анализа, распознавания и интерпретации изображений: обнаружения людей, распознавания номерных знаков и др. Достаточно очевидно, что отсчеты такого дескриптора по отдельности могут быть получены как значение первой компоненты МО-дескриптора для достаточно простой (зависящей только от фазовой составляющей) модели градиентного поля, характеризующей пороговой величиной фазового рассогласования с предопределенной фазовой константой (константа определяет допустимые отклонения в направлении градиента). Очевидно, что такое решение могло бы быть улучшено, если модели МО-дескрипторов настраивать по обучающему множеству изображений. На взгляд автора, одним из возможных решений здесь может стать метод главных компонент, используемый ниже.

3. Формальное описание предлагаемого метода

В рамках настоящего параграфа представлено более детальное описание основных блоков предлагаемого метода построения описания изображений с использованием МО-дескрипторов, указанных на рис. 1 (компоненты схемы, ограниченные пунктирной линией).

3.1. Предварительная обработка

Данный шаг предлагаемого метода является опциональным и зависящим от задачи. В рамках экспериментов, описываемых ниже, использовалось два варианта предварительной обработки, описываемые известными поэлементными преобразованиями [2]:

- нормализация яркости и контраста,
- эквализация гистограммы.

В случае *нормализации яркости и контраста* функция яркости $f(n_1, n_2)$ анализируемого изображения преобразовывалась по линейному закону

$$f(n_1, n_2) \leftarrow af(n_1, n_2) + b$$

с такими параметрами a и b , чтобы среднее и выборочная дисперсия преобразованного изображения стали наперед заданными.

В случае *эквализации гистограммы* функция яркости $f(n_1, n_2)$ анализируемого изображения преобразовывались по закону, напрямую зависящему от исходного (эмпирического) распределения $P_f(f_0) \equiv P[f < f_0]$ функции яркости:

$$f(n_1, n_2) \leftarrow f_{\min} + (f_{\max} - f_{\min}) P_f[f(n_1, n_2)].$$

Здесь f_{\min}, f_{\max} – требуемые минимальное и максимальное значения функции яркости итогового изображения. На практике обычно полагают: $f_{\min} = 0, f_{\max} = 255$. Само распределение $P_f(f_0)$ функции яркости оценивается по входному изображению.

3.2. Расчет поля градиента

По цифровому изображению $f(n_1, n_2)$ градиентное поле $\bar{g} \equiv \{g(n_1, n_2)\}_{(n_1, n_2) \in D}$ может быть получено с использованием любого известного способа, например, с использованием *оператора Превитта* [2].

3.3. Модели градиентного поля, признаки МО-дескриптора

В качестве моделей будем использовать поля градиента с независимыми отсчетами [4]. Тогда плотность вероятностей конкретного градиентного поля \bar{G}_k определена в виде:

$$p_{\bar{G}_k}(\bar{g}) = \prod_{(n_1, n_2) \in D} p_{G_k(n_1, n_2)}(g(n_1, n_2)), \quad (4)$$

где $p_{G_k(n_1, n_2)}(\cdot)$ – плотность вероятностей градиента в конкретном отсчете (n_1, n_2) цифрового изображения. Дальнейшие рассуждения в этом пункте проведем для конкретного градиентного поля \bar{G}_k , опуская соответствующий индекс: \bar{G} .

Плотность вероятностей отсчета конкретного градиентного поля зададим в параметрическом виде

$$p_{G(n_1, n_2)}(g(n_1, n_2)) = \begin{cases} a_{(n_1, n_2)} \cdot \exp \left[\left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right| \right], & 0 \leq |g(n_1, n_2)| \leq g_{\max}; \\ 0, & |g(n_1, n_2)| > g_{\max}. \end{cases} \quad (5)$$

Здесь, по аналогии с работой [4]: $a_{(n_1, n_2)}$ – нормирующий множитель; $\beta_{(n_1, n_2)}$ – величина, характеризующая степень «выраженности» градиента в соответствующем отсчете изображения, λ – показатель степени (в работе использовались значения «1» и «2»); $\phi_{(n_1, n_2)}$ – величина, определяющая наиболее вероятное значение фазы градиента в соответствующем отсчете изображения; g_{\max} – наибольшее значение для модуля градиента на изображении.

Тогда значение плотности (4) имеет вид:

$$p_{\bar{G}}(\bar{g}) = A \exp \left(\sum_{(n_1, n_2)} \left(\left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right| \right) \right),$$

$$\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A.$$

Здесь

$$A = \prod_{(n_1, n_2)} a_{(n_1, n_2)}.$$

Полагая функцию $\varphi: \mathbf{R}_+ \rightarrow \mathbf{R}$ в виде

$$\varphi(x) = \ln \left(\frac{x}{A} \right), \quad (6)$$

имеем признаки (3) в следующем виде:

$$\vartheta_0 = \frac{\sum_{(n_1, n_2)} \left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right|}{g_{\max} \sum_{(n_1, n_2)} \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right|},$$

$$\vartheta_1 = \frac{\sum_{(n_1, n_2)} \left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right|}{\sum_{(n_1, n_2)} \left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda}, \quad (7)$$

$$\vartheta_2 = \frac{\sum_{(n_1, n_2)} \left| g(n_1, n_2) \right| \cdot \beta_{(n_1, n_2)}^\lambda \cdot \left| \cos(\arg g(n_1, n_2) - \phi_{(n_1, n_2)}) \right|}{g_{\max} \sum_{(n_1, n_2)} \beta_{(n_1, n_2)}^\lambda}.$$

Любую из этих величин или все вместе можно использовать в качестве признака изображения. Величина ϑ_1 , используемая в работе далее, принимает тем большее значение, чем ближе фазовые составляющие модели и анализируемого поля градиента в отсчетах с высокой степенью «выраженности».

3.4. *Настройка моделей: метод на основе PCA*

Для проведения расчета признака ϑ_1^k (7) по каждой из моделей \bar{G}_k ($k = \overline{0, K-1}$) необходимо в процессе обучения (настройки) системы определить набор числовых параметров $\beta_{(n_1, n_2)}^k \phi_{(n_1, n_2)}^k$. Для этого предлагается использовать метод главных компонент (PCA – Principal Component Analysis). Специфика его использования в данной работе заключается в том, что данные представляют собой набор комплекснозначных градиентных полей [15].

Для удобства изложения представим двумерное комплекснозначное поле $g(n_1, n_2)$ в виде одномерного комплексного вектора \bar{v} с компонентами $v(n)$, положив, например:

$$v(n_1 N_2 + n_2) = g(n_1, n_2). \tag{8}$$

Суть метода главных компонент, как известно [15–18, 20], состоит в построении оптимального в смысле ряда критериев (в частности, по критерию среднеквадратической ошибки представления) разложения исходного набора из M векторов $\{\bar{v}_m\}_{m=0}^{M-1}$ по новому базису – векторам $\{\bar{e}_k\}_{k=0}^{K-1}$ ($K \leq M$) (т.н. преобразование Кархунена–Лоэва):

$$\bar{v}_m \cong \sum_{k=0}^{K-1} \gamma_k^m \bar{e}_k + \bar{v}_E, \quad m = \overline{0, M-1}. \tag{9}$$

Здесь вектор \bar{v}_E – вектор средних для комплексных векторов $\{\bar{v}_m\}_{m=0}^{M-1}$. Заметим, что, поскольку исходные вектора комплексные, получаемый базис также оказывается комплексным, так же как и коэффициенты разложения γ_k^m .

Базис $\{\bar{e}_k\}_{k=0}^{K-1}$ – суть *собственные векторы эрмитовой (выборочной) матрицы ковариации С*

$$C \equiv \sum_{m=0}^{M-1} (\bar{v}_m - \bar{v}_E) (\bar{v}_m^* - \bar{v}_E^*)^T$$

исходного множества центрированных векторов, соответствующие K наибольшим собственным числам (у эрмитовой матрицы все собственные числа λ_k вещественны):

$$C \bar{e}_k = \lambda_k \bar{e}_k, \quad k = \overline{0, K-1}.$$

В выражении для матрицы C символ T – знак транспонирования соответствующего вектора.

Получив базис $\{\bar{e}_k\}_{k=0}^{K-1}$ представления комплексных векторов $\{\bar{v}_m\}_{m=0}^{M-1}$ и соответствующих полей градиента и учитывая связь компонент векторов и отсчетов двумерных полей в виде (8), параметры моделей (5) предлагается задать следующим образом:

$$\beta_{(n_1, n_2)}^k = |e_k(n_1 N_2 + n_2)|,$$

$$\phi_{(n_1, n_2)}^k = \arg(e_k(n_1 N_2 + n_2)), \quad k = \overline{0, K-1}. \tag{10}$$

Число K , определяющее количество используемых моделей дескрипторов, в конкретной задаче может выступать как независимый или зависимый (например, от точности представления (9)) параметр.

4. *Экспериментальное исследование предлагаемого метода описания*

Исследование эффективности предложенного описания изображений с использованием МО-дескрипторов проводилось на примере решения задачи распознавания лиц [22]. В качестве тестовых данных использовались изображения лиц базы данных «Extended Yale Face Database В» [23, 24]. По информации разработчика базы данных, указанной на сервере «<http://vision.ucsd.edu/datasetsAll>», база данных содержит в целом 16128 изображений 28 человек, снятых с 9 позиций при 64 вариантах освещения, также для каждого человека присутствует один дополнительный снимок (65-й). По факту, доступный на официальном сайте архив размером 57123 Кб содержит данные при 64 вариантах освещения для 38 персон. Поскольку часть файлов повреждена, доступными для экспериментов являются 2420 изображений 38 персон.

Следует также отметить, что, несмотря на значительное число работ по решению задачи распознавания лиц с использованием различных методов классификации и подходов, использовать результаты какой-либо из них для выполнения *корректного сравнения* (то есть сравнения выбранных представлений и/или методов классификации в одних и тех же условиях) оказывается довольно затруднительно: авторы используют различные методы предобработки, различные фрагменты исходных изображений, различные подмножества тестовых баз данных, различные методы обучения (при обучении может быть один объект в классе, несколько в классе, может использоваться или не использоваться метод кросс-валидации, подвыборки могут быть детерминированные или случайные и др.) и оценки показателей эффективности, различное подмножество признаков описания и т.п. В результате показатели качества распознавания в работах варьируются даже для одной тестовой базы от, например, 0,54 (стр. 32 в монографии [22]) до практически идеального распознавания.

Учитывая все вышесказанное, для оценки эффективности предлагаемого подхода описания изображения представляется целесообразным проведение идентичных экспериментов по распознаванию и сравнению получаемых результатов с существующим методом-прототипом. Поскольку для синтеза модели в предлагаемом подходе применялся метод PCA, сравнение проводилось с «классическим» методом «собственных лиц» [25]. То есть прототипом описания являлись коэффициенты разложения полутоновых изображений по базису, получаемого методом главных компонент [20, 22, 25].

Для определенности, эксперименты проводятся при следующих условиях, позволяющих проверить получаемые в итоге результаты:

- тестовая база данных содержит 2420 изображений для 38 персон (см. описание исходных данных выше);
- изображения использовались с пониженным разрешением, то есть вместо исходных изображений размера 192×168 использовались картинки 96×84, полученные заменой каждого фрагмента 2×2 на единственный отсчет с усредненным значением яркости;
- из 64 изображений для каждого человека первые 32 изображения используются для обучения (настройки), а последующие 27 – для оценки показателей распознавания;
- в сопоставляемых подходах используется один и тот же метод предварительной обработки (варианты: нет предобработки, нормализация яркости и контраста, эквализация гистограммы), который применялся до снижения разрешения изображений;
- в качестве классификатора использовался метод ближайшего соседа с евклидовой метрикой;
- качественные показатели сопоставлялись при одинаковом количестве признаков-описателей (вещественных);
- в качестве показателя качества выступало относительное число верно классифицированных изображений (1026 изображений по 38 классам).

Примеры комплекснозначных «собственных лиц», получаемых методом PCA и используемых для синтеза моделей МО-дескрипторов, для различных вариантов предварительной обработки представлены на рис. 2. Цвет характеризует локальную ориентацию поля градиента, яркость – его амплитуду. При этом красный и синий цвета соответствуют противоположным (вертикальным) направлениям градиента. Фиолетовый, таким образом, соответствует горизонтальному направлению.



Рис. 2. Примеры комплекснозначных «собственных лиц». Варианты предварительной обработки: отсутствует (а); эквализация гистограммы (б)

Собственно показатели качества решения искомой задачи приведены в табл. 1–3. В последнем столбце («выигрыш») указана величина (в процентах) относительного роста качественного показателя, вычисляемая как отношение величины приращения показателя качества к его значению для метода-прототипа.

Табл. 1. Относительное число верно классифицированных объектов: нет предварительной обработки

Число признаков	Показатель качества		
	Прототип	Признаки МО-дескрипторов	Выигрыш (%)
2	0,0419	0,0564	34,61
4	0,0848	0,1901	124,17
8	0,2098	0,3333	58,87
15	0,3168	0,4561	43,97
50	0,4347	0,582	33,89
100	0,4737	0,6199	30,86

Табл. 2. Относительное число верно классифицированных объектов: нормализация яркости/контраста

Число признаков	Показатель качества		
	Прототип	Признаки МО-дескрипторов	Выигрыш (%)
2	0,0395	0,0741	87,59
4	0,0593	0,2160	264,25
8	0,1741	0,4136	137,56
15	0,2901	0,5099	75,77
50	0,3802	0,5988	57,50
100	0,4173	0,6716	60,94

Табл. 3. Относительное число верно классифицированных объектов: эквализация гистограммы

Число признаков	Показатель качества		
	Прототип	Признаки МО-дескрипторов	Выигрыш (%)
2	0,0395	0,1321	234,43
4	0,0728	0,3642	400,27
8	0,2198	0,4728	115,10
15	0,3753	0,7148	90,46
50	0,5074	0,9049	78,34
100	0,5457	0,9247	69,45

По представленным результатам исследования можно сделать следующие выводы:

- предлагаемый метод описания изображения с использованием МО-дескрипторов стабильно дает более высокие показатели качества при любом количестве признаков и любом способе предварительной обработки (выигрыш всегда положительный),
- наилучшие качественные показатели предлагаемый метод (так же, как и метод-прототип) демонстрирует для случая предварительной обработки путем эквализации гистограммы (см. табл. 3);
- большой (относительный) выигрыш достигается при меньшем числе признаков;
- выигрыш предлагаемого метода по сравнению с методом-прототипом всегда существенен. Минимальный выигрыш для всех экспериментов оказывается 30 % (100 признаков и отсутствие предварительной обработки – табл. 1), максимальный – 400 %, то есть качество решения задачи возрастает в 4 раза! Для наилучшего результативного варианта предварительной обработки (эквализация, табл. 4) рост показателя качества всегда оказывается не менее 70 %!

Обобщая представленные выше промежуточные выводы, отражающие детальный анализ результатов, можно говорить о значительном приросте качественных показателей при использовании предло-

женного метода описания изображений с использованием признаков МО-дескрипторов по сравнению с методом-прототипом.

Выводы и рекомендации

В работе предложен и исследован метод построения описания изображения с использованием модельно-ориентированных дескрипторов. Метод предполагает построение описателей анализируемого изображения на основе сопоставления фактического поля градиента этого изображения с моделью МО-дескриптора. На примерах решения практической задачи распознавания изображений лиц продемонстрировано *значительное и стабильное преимущество предлагаемого метода* описания по сравнению с методом-прототипом, использующим тот же математический принцип (метод главных компонент) при обработке полутоновых изображений.

Разработка оригинального алгоритма синтеза моделей МО-дескриптора может являться дальнейшим направлением работ в этой области.

Благодарности

Исследование выполнено при финансовой поддержке:

- грантов РФФИ в рамках научных проектов № 15-07-01164-а, № 17-29-03190-офи-м;

- программы фундаментальных исследований Президиума РАН «Фундаментальные проблемы информатики и информационных технологий».

В качестве тестовых данных использовались изображения лиц базы данных «Extended Yale Face Database B» [23, 24].

Литература

1. **Mikolajczyk, K.** A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27, Issue 10. – P. 1615-1630. – DOI: 10.1109/TPAMI.2005.188.
2. Методы компьютерной обработки изображений / М.В. Гашников, Н.И. Глузов, Н.Ю. Ильясова, В.В. Мясников, С.Б. Попов, В.В. Сергеев, В.А. Соيفер, А.Г. Храмов, А.В. Чернов, В.М. Чернов, М.А. Чичева, В.А. Фурсов; под ред. В.А. Соифера. – Изд. 2-е, испр. – М.: Физматлит, 2003. – 784 с. – ISBN: 5-9221-0270-2.
3. **Duda, R.O.** Pattern classification and scene analysis / R.O. Duda, P.E. Hart. – New York: Wiley, 1973. – 512 p. – ISBN: 978-0471223610.
4. **Мясников, В.В.** Модельно-ориентированный дескриптор поля градиента как удобный аппарат распознавания и анализа цифровых изображений / В.В. Мясников // Компьютерная оптика. – 2012. – Т. 36, № 4. – С. 596-604.
5. **Мясников, В.В.** Метод обнаружения транспортных средств на цифровых аэрофото- и космических изображениях дистанционного зондирования земли / В.В. Мясников // Компьютерная оптика. – 2012. – Т. 36, № 3. – С. 429-438.
6. **Kuznetsov, A.V.** New algorithms for verifying the consistency between satellite images and survey conditions / A.V. Kuznetsov, V.V. Myasnikov // Pattern Recognition and Image Analysis. – 2016. – Vol. 26(3). – P. 593-596. – DOI: 10.1134/S1054661816030135.
7. **Dalal, N.** Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA. – 2005. – P. 886-893. – DOI: 10.1109/CVPR.2005.177.
8. **Lowe, D.G.** Distinctive image features from scale-invariant keypoints / D.G. Lowe // International Journal of Computer Vision. – 2004. – Vol. 60, Issue 2. – P. 91-110. – DOI: 10.1023/B:VISI.0000029664.99615.94.
9. **Gonzalez-Reyna, S.E.** Eigen-gradients for traffic sign recognition / S.E. Gonzalez-Reyna, J.G. Avina-Cervantes, S.E. Ledesma-Orozco, I. Cruz-Aceves // Mathematical Problems in Engineering. – 2013. – Vol. 2013. – 364305 (6 p.). – DOI: 10.1155/2013/364305.
10. **Hu, R.** A performance evaluation of gradient field HOG descriptor for sketch based image retrieval / Rui Hu, John Colloso // Computer Vision and Image Understanding. – 2013. – Vol. 117, Issue 7. – P. 790-806. – DOI: 10.1016/j.cviu.2013.02.005.
11. **Tzimiropoulos, G.** Principal component analysis of image gradient orientations for face recognition / G. Tzimiropoulos, S. Zafeiriou, M. Pantic // 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops (FG 2011). – 2011. – DOI: 10.1109/FG.2011.5771457.
12. **Храмов, А.Г.** Метод поля направлений в анализе и интерпретации диагностических изображений : дис. ... доктора техн. наук : 05.13.17 / Храмов Александр Григорьевич. – Самара, 2006. – 230 с.
13. **Harshman, R.A.** PARAFAC: Parallel factor analysis / R.A. Harshman, M.E. Lundy // Computational Statistics and Data Analysis. – 1994. – Vol. 18, Issue 1. – P. 39-72. – DOI: 10.1016/0167-9473(94)90132-5.
14. Geometric computing with Clifford algebras: Theoretical foundations and applications in computer vision and robotics / ed. by G. Sommer. – Berlin, Heidelberg: Springer Verlag, 2000. – 529 p. – ISBN: 978-3-540-41198-7.
15. **Horel, J.D.** Complex principal component analysis: Theory and examples / J.D. Horel // Journal of Climate and Applied Meteorology. – 1984. – Vol. 23. – P. 1660-1673. – DOI: 10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2.
16. **Schreier, P.J.** Statistical signal processing of complex-valued data: The theory of improper and noncircular signals / P.J. Schreier, L.L. Scharf. – Cambridge: Cambridge University Press, 2010. – 330 p. – ISBN: 978-0-521-89772-3.
17. **Jolliffe, I.T.** Principal component analysis / I.T. Jolliffe. – 2nd ed. – New York, Berlin, Heidelberg: Springer-Verlag, 2002. – ISBN: 0-387-95442-2.
18. **Бриллинджер, Д.П.** Временные ряды. Обработка данных и теория / Д.П. Бриллинджер. – пер. с англ. – М.: Мир, 1980. – 536 с.
19. **Alfsmann, D.** Hypercomplex algebras in digital signal processing: Benefits and drawbacks / D. Alfsmann, H.G. Gockler, S.J. Sangwine, T.A. Ell // 15th European Signal Processing Conference (EUSIPCO 2007). – 2007. – P. 1322-1326.
20. **Delac, K.** Independent comparative study of PCA, ICA, and LDA on the FERET data set / K. Delac, M. Grgic, S. Grgic // International Journal of Imaging Systems and Technology. – 2005. – Vol. 15, Issue 5. – P. 252-260. – DOI: 10.1002/ima.20059.
21. **Duin, R.P.W.** Featureless pattern classification / R.P.W. Duin, D. de Ridder, D.M.J. Tax // Kybernetika. – 1998. – Vol. 34(4). – P. 399-404.

22. Advances in face image analysis: Techniques and technologies / ed. by Y.-J. Zhang. – Hershey, PA: IGI Global, 2011. – 350 p. – ISBN: 978-1-61520-991-0.
23. **Georghiades, A.S.** From few to many: Illumination cone models for face recognition under variable lighting and pose / A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – Vol. 23, Issue 6. – P. 643-660. – DOI: 10.1109/34.927464.
24. **Lee, K.C.** Acquiring linear subspaces for face recognition under variable lighting / K.C. Lee, J. Ho, D. Kriegman // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27, Issue 5. – P. 684-698. – DOI: 10.1109/TPAMI.2005.92.
25. **Turk, M.** Eigenfaces for recognition / M. Turk, A. Pentland // Journal of Cognitive Neuroscience. – 1991. – Vol. 3(1). – P. 71-86. – DOI: 10.1162/jocn.1991.3.1.71.

Сведения об авторе

Мясников Владислав Валерьевич, 1971 года рождения. В 1994 году окончил Самарский государственный аэрокосмический университет (СГАУ). В 1995 году поступил в аспирантуру СГАУ, в 1998 году защитил диссертацию на степень кандидата технических наук, а в 2008 – диссертацию на степень доктора физико-математических наук. В настоящее время работает профессором кафедры геоинформатики и информационной безопасности Самарского национального исследовательского университета имени академика С.П. Королева и одновременно ведущим научным сотрудником в Институте систем обработки изображений РАН – филиале ФНИЦ «Кристаллография и фотоника» РАН. Круг научных интересов включает компьютерное зрение, распознавание образов и искусственный интеллект, искусственные нейронные сети, цифровую обработку сигналов и изображений, геоинформатику. Имеет около 200 публикаций, в том числе более 100 статей и две монографии (в соавторстве). Член Российской ассоциации распознавания образов и анализа изображений. Страница в интернете: <http://www.ssau.ru/staff/62061001-Myasnikov-Vladislav-Valerevich>. E-mail: vmias@geosamara.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 30 августа 2017 г. Окончательный вариант – 19 октября 2017 г.

DESCRIPTION OF IMAGES USING MODEL-ORIENTED DESCRIPTORS

V.V. Myasnikov^{1,2}

¹Samara National Research University, Samara, Russia,

²Image Processing Systems Institute of RAS – Branch of the FSRC “Crystallography and Photonics” RAS, Samara, Russia

Abstract

The paper proposes an approach to constructing an image description using a set of model-oriented descriptors. Each descriptor characterizes the "similarity" of the analyzed image, represented as a complex-valued gradient field, to a pre-selected model of this descriptor. It is proposed that descriptor models should be synthesized using a method of principal components, or discriminant analysis, which has been applied to a diversity of complex-valued gradient field realizations. As a result, the proposed approach enables the complex-valued field of the gradient of the analyzed image to be described as a set of real quantities from the interval [0,1], capable of simultaneously characterizing the phase and magnitude of the image gradient. The effectiveness of the proposed approach is illustrated via solving a face recognition problem and comparing the result with prototype solutions (based on the principal component method and discriminant analysis), which directly utilize halftone images. The comparison is made using a nearest neighbor's classifier.

Keywords: digital images, descriptors, features, analysis, recognition, image retrieval.

Citation: Myasnikov VV. Description of images using model-oriented descriptors. Computer Optics 2017; 41(6): C. 888-896. DOI: 10.18287/2412-6179-2017-41-6-888-896.

Acknowledgements: The reported study was partially funded by RFBR according to the research project № 172901001 and by the Program of Fundamental Research of the Presidium of the Russian Academy of Sciences "Fundamental Problems of Informatics and Information Technologies". As the test data, images of the individuals of the database "Extended Yale Face Database B" were used [23, 24].

References

- [1] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence 2005; 27(10): 1615-1630. DOI: 10.1109/TPAMI.2005.188.
- [2] Soyfer VA, ed. Computer image processing methods [In Russian]. Moscow: "Fizmatlit" Publisher; 2003. ISBN: 5-9221-0270-2.
- [3] Duda RO, Hart PE. Pattern classification and scene analysis. New York: Wiley; 1973. 512. ISBN: 978-0471223610.
- [4] Myasnikov VV. Model-based gradient field descriptor as a convenient tool for image recognition and analysis [In Russian]. Computer Optics 2012; 36(4): 596-604.
- [5] Myasnikov VV. Method for detection of vehicles in digital aerial and space remote sensed images [In Russian]. Computer Optics 2012; 36(3): 429-438.

- [6] Kuznetsov AV, Myasnikov VV. New algorithms for verifying the consistency between satellite images and survey conditions. *Pattern Recognition and Image Analysis* 2016; 26(3): 593-596. DOI: 10.1134/S1054661816030135.
- [7] Dalal N, Triggs B. Histograms of oriented gradients for human detection. *Proc CVPR 2005*: 886-893. DOI: 10.1109/CVPR.2005.177.
- [8] Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comp Vision* 2004; 60(2): 91-110. DOI: 10.1023/B:VISI.0000029664.99615.94.
- [9] Gonzalez-Reyna SE, Avina-Cervantes JG, Ledesma-Orozco SE, Cruz-Aceves I. Eigen-gradients for traffic sign recognition. *Mathematical Problems in Engineering* 2013; 2013: 364305. DOI:10.1155/2013/364305.
- [10] Hu R, Collomosse J. A performance evaluation of gradient field HOG descriptor for sketch based image retrieval. *Computer Vision and Image Understanding* 2013; 117(7): 790-806. DOI: 10.1016/j.cviu.2013.02.005.
- [11] Tzimiropoulos G, Zafeiriou S, Pantic M. Principal component analysis of image gradient orientations for face recognition. *IEEE Int Conf Automatic Face and Gesture Recognition (FG 2011)* 2011. DOI: 10.1109/FG.2011.5771457.
- [12] Khramov AG. Direction field method and its application for medicine images analysis and interpretation [In Russian]. The thesis for the Doctoral degree in Technical Sciences. Samara; 2006.
- [13] Harshman RA, Lundy ME. PARAFAC: Parallel factor analysis. *Computational Statistics and Data Analysis* 1994; 18(1): 39-72. DOI: 10.1016/0167-9473(94)90132-5.
- [14] Sommer G, ed. *Geometric computing with Clifford algebras: Theoretical foundations and applications in computer vision and robotics*. Berlin, Heidelberg: Springer Verlag; 2000. ISBN: 978-3-540-41198-7.
- [15] Horel JD. Complex principal component analysis: Theory and examples. *Journal of Climate and Applied Meteorology* 1984; 23: 1660-1673. DOI: 10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2.
- [16] Schreier PJ, Scharf LL. *Statistical signal processing of complex-valued data: The theory of improper and noncircular signals*. Cambridge: Cambridge University Press; 2010. ISBN: 978-0-521-89772-3.
- [17] Jolliffe IT. *Principal component analysis*. 2nd ed. New York, Berlin, Heidelberg: Springer-Verlag; 2002. ISBN: 0-387-95442-2.
- [18] Brillinger DR. *Time series: Data analysis and theory*. Holden-Day, 1981. 540.
- [19] Alfsmann D, Gockler HG, Sangwine SJ, Ell TA. Hypercomplex algebras in digital signal processing: benefits and drawbacks. *EUSIPCO 2007*: 1322-1326.
- [20] Delac K, Grgic M, Grgic S. Independent comparative study of PCA, ICA, and LDA on the FERET data set. *Int J Imaging Syst Technol* 2005; 15(5): 252-260. DOI: 10.1002/ima.20059.
- [21] Duin RPW, de Ridder D, Tax DMJ. Featureless pattern classification. *Kybernetika* 1998; 34(4): 399-404.
- [22] *Advances in face image analysis: Techniques and technologies*. Ed by Zhang YJ. IGI Global, USA 2011. 350.
- [23] Georgiades AS, Belhumeur PN, Kriegman DJ. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2001; 23(6): 643-660. DOI: 10.1109/34.927464.
- [24] Lee KC, Ho J, Kriegman D. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2005; 27(5): 684-698. DOI: 10.1109/TPAMI.2005.92.
- [25] Turk M, Pentland A. Eigenfaces for recognition. *J Cogn Neurosci* 1991; 3(1): 71-86. DOI: 10.1162/jocn.1991.3.1.71.

Author's information

Vladislav Valerievich Myasnikov (1971 b.), graduated (1994) from the S.P. Korolyov Samara State Aerospace University (SSAU). He received his PhD in Technical Sciences (2002) and DrSc degree in Physics & Maths (2008). Currently he works as a professor at the Geoinformatics and Information Security department in Samara National Research University and, at the same time, as a leading researcher at the Institute of Image Processing Systems of the Russian Academy of Sciences, a branch of the Russian Academy of Sciences 'Crystallography and Photonics' RAS. The range of scientific interests includes computer vision, pattern recognition and artificial intelligence, artificial neural networks, digital processing of signals and images, and geoinformatics. He has about 200 publications, including more than 100 articles and two monographs (in co-authorship). Member of the Russian Association for Pattern Recognition and Image Analysis. <http://www.ssau.ru/staff/62061001-Myasnikov-Vladislav-Valerevich> . E-mail: vmyas@geosamara.ru .

Received August 30, 2017. The final version – October 19, 2017.