

ОБРАБОТКА ИЗОБРАЖЕНИЙ, РАСПОЗНАВАНИЕ ОБРАЗОВ

СРАВНЕНИЕ АЛГОРИТМОВ ОПИСАНИЯ КОМПЛЕКСНОЗНАЧНОГО ПОЛЯ ГРАДИЕНТА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ЛИНЕЙНЫХ МЕТОДОВ СНИЖЕНИЯ РАЗМЕРНОСТИ

Е.А. Дмитриев¹, В.В. Мясников^{1,2}

¹ Самарский национальный исследовательский университет имени академика С.П. Королева,
443086, Россия, г. Самара, Московское шоссе, д. 34;

² ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН,
443001, Россия, г. Самара, ул. Молодогвардейская, д. 151

Аннотация

В работе представлен анализ различных подходов к построению описаний полей градиентов цифровых изображений. Анализируемые подходы основаны на известных методах снижения размерности данных, таких как методы главных и независимых компонент, метод дискриминантного анализа. Мы применяем указанные методы не к исходному изображению, представленному в виде двумерного поля яркости (полутоновому изображению), а к его вторичному представлению в виде двумерного поля градиента – комплекснозначному изображению. При этом рассматриваются случаи использования как поля градиента целиком, так и только его фазовой части. Дополнительно рассматриваются два независимых способа формирования окончательного описания искомого объекта: в виде коэффициентов разложения поля градиента по сформированному базису и с использованием оригинальной авторской конструкции модельно-ориентированных дескрипторов. Последние позволяют в два раза снизить число вещественных коэффициентов, используемых при описании искомого объекта. В качестве конкретной прикладной задачи, на которой проводятся исследования, выступает проблема распознавания лиц. Эффективность анализируемых подходов демонстрируется путём сравнения результатов их применения к изображениям из базы данных «Extended Yale Face Database B». Алгоритмом классификации выступает метод ближайшего соседа.

Ключевые слова: распознавание лиц, метод главных компонент, метод независимых компонент, линейный дискриминантный анализ, модельно-ориентированные дескрипторы, The Extended Yale Database B, описание изображения.

Цитирование: Дмитриев, Е.А. Сравнение алгоритмов описания комплекснозначного поля градиента цифровых изображений с использованием линейных методов снижения размерности / Е.А. Дмитриев, В.В. Мясников // Компьютерная оптика. – 2018. – Т. 42, № 5. – С. 822-828. – DOI: 10.18287/2412-6179-2018-42-5-822-828.

Введение

Описание изображений с использованием признаков – один из основных этапов в решении задач распознавания визуальных образов. Линейные методы снижения размерности описаний являются достаточно эффективными при реализации этого этапа. Основная их идея состоит в построении линейного пространства меньшей размерности (подпространства) и нахождении проекции исходных векторов-описаний (в случае изображений вектор формируется с использованием некоторой развёртки) на найденное подпространство. Ключевым моментом является построение «хорошего» с точки зрения критерия прикладной задачи подпространства. Существует большое число исследований, которые характеризуют эффективность методов линейного снижения размерности применительно к различным задачам как по отдельности [1, 2, 3], так и в результате взаимных сравнений [4, 5]. В некоторых случаях выводы работ оказываются противоположными [6, 7].

Несмотря на значительные достижения современных нейронных сетей [8], где описание изображения с помощью подпространства меньшей размерности

формируется неявно автоматически, мы показываем, что существует возможность значительного улучшения качества решения прикладной задачи (в нашей работе – распознавания изображений) и с применением классических линейных методов снижения размерности: метода главных компонент (*Principal Component Analysis* – PCA), метода независимых компонент (*Independent Component Analysis* – ICA), линейного дискриминантного анализа (*Linear Discriminant Analysis* – LDA), но используемых с некоторой дополнительной обработкой данных. В частности, в нашей работе мы использовали в качестве входных данных не сами (полутоновые) изображения, а поле градиента исходного изображения, которое может быть представлено как двумерное поле комплексных чисел. При этом рассматриваются случаи использования как поля градиента целиком, так и только его фазовой части. Изображения, содержащие только фазовую составляющую поля градиента, в существующей научной литературе называют *полем направлений* [9].

Дополнительно рассматриваются два способа получения окончательного представления при синтези-

рованном базисе подпространства, построенном с использованием обозначенных методов линейного снижения размерности. Первый предполагает традиционный подход, когда признаки рассчитываются как коэффициенты разложения поля градиента (или поля направлений) по сформированному базису. Второй способ использует оригинальную авторскую конструкцию, предложенную в работах [10, 11], и названную *модельно-ориентированным дескриптором*. В отличие от традиционного подхода, формирующего комплекснозначный вектор признаков определённой длины, модельно-ориентированный дескриптор позволяет получить вещественный вектор той же длины. При этом он также учитывает и фазовую, и амплитудную составляющие и базиса, и анализируемого поля градиента. Но число вещественных чисел, формирующих описание, оказывается в два раза меньшим. Дополнительным потенциальным преимуществом второго способа является возможность использования многих готовых методов классификации и распознавания, ориентированных на работу с вещественными числами.

Мотивацией написания работы послужил факт отсутствия исследований по применению линейных методов снижения размерности данных (PCA, ICA, LDA) для описания комплекснозначных полей градиентов. Для полутоновых изображений похожие исследования представлены в работе [4], где авторы провели их сравнительный анализ на примере решения задачи распознавания лиц базы данных «FERET». Результаты тех исследований показали, что ни один из указанных методов (или метрик, используемых при классификации по ближайшему соседу) не обеспечивает уверенного преимущества по сравнению с другими методами и метриками. Особый интерес для нас также представляет сравнение решений, которые используют в качестве первичного стандартное полутоновое представление поля яркости и нестандартное представление в виде комплекснозначного поля градиента.

Дальнейшая часть работы построена следующим образом. В первом параграфе приводятся основные используемые термины и понятия изображения, полей градиента и направлений. Во втором кратко представлены используемые методы линейного снижения размерности данных, анализируемые способы построения окончательных описаний, включая модельно-ориентированный дескриптор. В третьем параграфе представлены результаты исследований представленных выше подходов на примере решения задачи распознавания лиц с использованием базы данных «Extended Yale Face Database» [12, 13]. В заключение работы приведены основные выводы и список использованных источников.

1. Основные понятия

Пусть $I(n_1, n_2) \in \mathbf{R}$ – *цифровое изображение*, где $(n_1, n_2) \in D$, $D = \{(n_1, n_2) : n_1 = \overline{0, N_1 - 1}, n_2 = \overline{0, N_2 - 1}\}$, а $g(n_1, n_2) \in \mathbf{C}$ – *градиентное поле* этого изображения, которое можно получить на практике, например, с

помощью оператора Собеля [14]. Модуль $|g(n_1, n_2)|$ и аргумент $\arg(g(n_1, n_2))$ отсчетов поля градиента характеризуют соответственно амплитуду и фазу (направление) градиента в конкретной позиции. Вещественное изображение $\arg(g(n_1, n_2))$ называется *полем направлений*.

Для удобства дальнейшего изложения представим двумерные поля градиента и направлений в виде векторов соответственно \mathbf{v} и $\boldsymbol{\varphi}$ следующим образом:

$$\mathbf{v}(n_1 N_2 + n_2) \equiv g(n_1, n_2),$$

$$\boldsymbol{\varphi}(n_1 N_2 + n_2) \equiv \arg(g(n_1, n_2)).$$

Первым этапом рассматриваемых подходов к формированию описаний является построение базиса подпространства с помощью одного из методов линейного снижения размерности PCA, ICA и LDA. Второй этап заключается в формировании окончательного описания анализируемого изображения или поля градиента в виде вектора комплексных или вещественных коэффициентов (*вектора признаков*) с помощью построенного базиса. Эти алгоритмы представлены в следующем параграфе.

2. Алгоритмы линейного снижения размерности и окончательного описания

Метод главных компонент (PCA)

Метод PCA позволяет получить базис подпространства, оптимальный по нескольким критериям [15]. Один из критериев оптимальности – максимальный разброс проекций векторов на найденное подпространство. Для того чтобы получить базис подпространства по методу PCA по набору централизованных векторов (векторов-столбцов \mathbf{x} , в качестве которых далее могут выступать вектора \mathbf{v} или $\boldsymbol{\varphi}$), формирующих матрицу данных \mathbf{X} , необходимо для выборочной ковариационной (для комплексных векторов – эрмитовой) матрицы данных:

$$\mathbf{C} = \mathbf{X}\mathbf{X}^H,$$

найти базис собственных векторов \mathbf{w} , соответствующих наибольшим собственным числам λ :

$$\mathbf{C}\mathbf{w} = \lambda\mathbf{w}.$$

Чтобы применить алгоритм PCA к набору векторов $\{\boldsymbol{\varphi}_j\}_{j \in J}$, характеризующих ориентации градиента, необходимо дополнительно ввести меру, которая позволяет оценить расстояние между такими векторами. В работе [16] предлагается использовать следующую величину, характеризующую разброс комплекснозначных векторов на N -мерной сфере единичного радиуса ($N = N_1 N_2$):

$$d^2(\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_j) \equiv \frac{1}{2} \left\| e^{i\boldsymbol{\varphi}_k} - e^{i\boldsymbol{\varphi}_j} \right\|_2^2 = \sum_{n=1}^{N-1} \left(1 - \cos(\boldsymbol{\varphi}_k(n), \boldsymbol{\varphi}_j(n)) \right).$$

В результате, перед применением алгоритма PCA построения базиса для вектора ориентаций градиента,

полученного из двумерного поля направлений, необходимо спроецировать исходный вектор на N -мерную сферу единичного радиуса.

Линейный дискриминантный анализ (LDA)

В отличие от метода PCA, в LDA производится попытка максимизировать разброс проекций векторов различных классов, минимизируя одновременно разброс проекций внутри классов. Данные условия записываются в виде критерия Фишера [15]:

$$J(\mathbf{w}) = \frac{\mathbf{w}^H \mathbf{S}_b \mathbf{w}}{\mathbf{w}^H \mathbf{S}_w \mathbf{w}} \rightarrow \max_{\mathbf{w}}$$

где \mathbf{S}_b и \mathbf{S}_w – соответственно межклассовая и внутриклассовая ковариационные матрицы [15]; \mathbf{w} – базисный вектор подпространства.

Задача стоит в максимизации критерия Фишера [15]. Решением является собственное разложение матрицы $\mathbf{S}_w^{-1} \mathbf{S}_b$. Учитывая специфику определения матрицы \mathbf{S}_b , максимальный размер системы базисных векторов, получаемой с помощью метода LDA для данных большой размерностью, не превосходит числа классов.

Метод независимых компонент (ICA)

Суть метода ICA заключается в нахождении такого подпространства, где компоненты проекций векторов оказываются статистически независимы [17]. Один из способов нахождения базиса по ICA – максимизация критерия негауссовости. Популярной и эффективной реализацией данного подхода является алгоритм FastICA, который описан в работе [18].

Основные этапы алгоритма FastICA включают в себя [18]: приведение ковариационной матрицы исходных данных к единичной матрице, начальную инициализацию компонент базисных векторов, оптимизации критерия негауссовости с помощью метода градиентного спуска, использование алгоритма ортогонализации Грамма–Шмидта для получения ортонормированной системы базисных векторов. В качестве критерия негауссовости часто используется ко-

$$y_j = \frac{\sum_{(n_1, n_2)} |g(n_1, n_2)| B_j(n_1, n_2)^2 \left| \cos(\arg(g(n_1, n_2)) - \phi_j(n_1, n_2)) \right|}{\sum_{(n_1, n_2)} |g(n_1, n_2)| B_j(n_1, n_2)^2},$$

здесь $B_j(n_1, n_2)$ и $\phi_j(n_1, n_2)$ – модуль и аргумент отсчёта (n_1, n_2) модели дескриптора, задаваемой конкретным базисным вектором.

3. Результаты экспериментов

Исследование предложенных методов описания изображений проводилось на примере решения задачи распознавания лиц. В качестве тестовых данных использовались изображения из базы данных «Extended Yale Face Database B» [12, 13]. Данная база содержит 38 персон при 64 различных вариантах освещения. Поскольку часть файлов повреждена, доступными для экспериментов являются 2420 изображений. Размер каждого изображения составляет 192×168 .

эффицент эксцесса. Для случайной комплексной величины Y с нулевым средним коэффициент эксцесса определяется следующим образом [18]:

$$K = E\{|Y|^4\} - \left(E\{|Y|^2\}\right)^2 - |E\{Y^2\}|^2.$$

Формирование вектора описаний:

стандартный подход

Стандартным подходом к построению окончательного представления вектора признаков y анализируемого объекта с использованием построенного базиса $\{\mathbf{w}_j\}_{j \in J}$ является выбор в качестве компонент y_j вектора коэффициентов разложения анализируемого вектора/изображения по базису:

$$y_j \equiv \mathbf{x}^T \mathbf{w}_j = \sum_{n=0}^{N-1} x(n) w_j^*(n).$$

В конкретных ситуациях в качестве компонент вектора \mathbf{x} выступают либо отсчёты поля яркости, либо отсчёты поля направления, либо отсчёты поля градиента. Заметим, что для комплексного базиса коэффициенты также оказываются комплексными, то есть требуют в два раза больше N вещественных чисел. Сократить такое описание можно, используя модельно-ориентированные дескрипторы.

Формирование вектора описаний:

модельно-ориентированные дескрипторы

Модельно-ориентированные дескрипторы предложены одним из авторов настоящей работы в публикациях [10, 11]. Базисные вектора выступают в качестве моделей этих дескрипторов, а признаками для описания исходных изображений являются вещественные коэффициенты на отрезке $[0, 1]$. Они определяют «похожесть» анализируемого вектора или изображения на модель дескриптора. В случае двумерного поля градиента признак для конкретной j -й модели определяется по следующей формуле [10]:

Для определённости эксперименты проводились при следующих условиях, позволяющих проверить полученные результаты:

- тестовая база данных содержала 2420 изображений для 38 человек;
- размер изображений уменьшался в два раза, то есть использовались изображения размером 96×84 ;
- половина изображения для каждого человека использовалась для оценки подпространства, а последующие – для получения оценки вероятности верного распознавания (показатель качества);
- использовался один и тот же метод предварительной обработки для всех изображений – эквализация гистограммы, который применялся до

снижения размерности и показал наилучшие результаты в других исследованиях [10];

- в качестве классификатора выступал метод ближайшего соседа с метрикой L2;

- качественные показатели сопоставлялись при одинаковом количестве признаков-описателей, в таблицах ниже указывались данные и по длине формируемого вектора признаков, и по числу его вещественных компонент;

- показателем качества выступало относительное число верно классифицированных изображений.

Примеры базисных векторов, получаемых с помощью рассматриваемых линейных методов сниже-

ния размерности, представлены на рис. 1, 2. Цвет характеризует аргумент компоненты базисного вектора, а яркость – модуль компоненты. Красный и синий цвета соответствуют различным направлениям значимых аргументов.

Результаты экспериментов приведены в табл. 1, 2. Все предложенные методы описаний изображений сравнивались с «классическим» подходом с использованием «собственных лиц» [19]. Когда рассматриваемые алгоритмы снижения размерности используются для полутоновых изображений, а в качестве признаков выступают вещественные коэффициенты разложения по базисным векторам.

Табл. 1. Относительное число верно классифицированных объектов: проекции векторов

Длина вектора признаков	Показатель качества / Число вещественных коэффициентов			
	Яркостные изображения	Поле направлений	Поля градиентов	Используемый метод снижения размерности
2	0,0395 / 2	0,0850 / 4	0,1229 / 4	PCA
	0,0446 / 2	0,2855 / 4	0,1493 / 4	LDA
	0,0388 / 2	0,1237 / 4	0,1361 / 4	ICA
4	0,0728 / 4	0,3927 / 8	0,3985 / 8	PCA
	0,0990 / 4	0,6312 / 8	0,4348 / 8	LDA
	0,0833 / 4	0,4232 / 8	0,3845 / 8	ICA
8	0,2198 / 8	0,6510 / 16	0,6271 / 16	PCA
	0,1683 / 8	0,8036 / 16	0,8267 / 16	LDA
	0,3465 / 8	0,7821 / 16	0,5817 / 16	ICA
15	0,3754 / 15	0,8218 / 30	0,7351 / 30	PCA
	0,2360 / 15	0,9249 / 30	0,8936 / 30	LDA
	0,5116 / 15	0,8952 / 30	0,7046 / 30	ICA
38	0,4546 / 38	0,9472 / 76	0,7599 / 30	PCA
	0,3185 / 38	0,9835 / 76	0,9158 / 76	LDA
	0,5726 / 38	0,9769 / 76	0,7384 / 76	ICA
50	0,5074 / 50	0,9637 / 100	0,7417 / 100	PCA
	–	–	–	LDA
	0,6972 / 50	0,9876 / 100	0,6914 / 100	ICA
100	0,5457 / 100	0,9785 / 200	0,6988 / 200	PCA
	–	–	–	LDA
	0,6559 / 100	0,9975 / 200	0,6155 / 200	ICA

Табл. 2. Относительное число верно классифицированных объектов: модельно-ориентированные дескрипторы

Длина вектора признаков	Показатель качества / Число вещественных коэффициентов			
	Яркостные изображения	Поле направлений	Поля градиентов	Используемый метод снижения размерности
2	0,0395 / 2	0,0495 / 2	0,1023 / 2	PCA
	0,0446 / 2	0,1007 / 2	0,0800 / 2	LDA
	0,0388 / 2	0,0281 / 2	0,0809 / 2	ICA
4	0,0728 / 4	0,3111 / 4	0,3696 / 4	PCA
	0,0990 / 4	0,2838 / 4	0,1650 / 4	LDA
	0,0833 / 4	0,1287 / 4	0,1922 / 4	ICA
8	0,2198 / 8	0,6163 / 8	0,6609 / 8	PCA
	0,1683 / 8	0,7797 / 8	0,3284 / 8	LDA
	0,3465 / 8	0,6023 / 8	0,4546 / 8	ICA
15	0,3754 / 15	0,7772 / 15	0,8300 / 15	PCA
	0,2360 / 15	0,8556 / 15	0,5116 / 15	LDA
	0,5116 / 15	0,7285 / 15	0,6254 / 15	ICA
38	0,4546 / 38	0,8672 / 38	0,8688 / 38	PCA
	0,3185 / 38	0,9637 / 38	0,7913 / 38	LDA
	0,5726 / 38	0,9769 / 38	0,7384 / 38	ICA
50	0,5074 / 50	0,9373 / 50	0,9125 / 50	PCA
	–	–	–	LDA
	0,6972 / 50	0,9504 / 50	0,7913 / 50	ICA
100	0,5457 / 100	0,9587 / 100	0,9208 / 100	PCA
	–	–	–	LDA
	0,6559 / 100	0,9950 / 100	0,9381 / 100	ICA

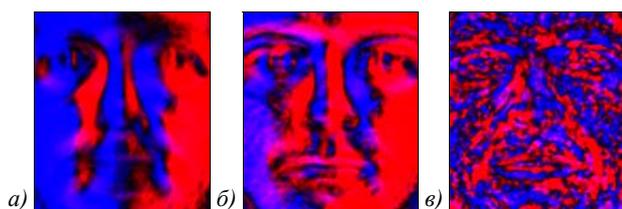


Рис. 1. Примеры базисных векторов с использованием ориентаций градиентов. Используемые методы снижения размерности: метод главных компонент (а); линейный дискриминантный анализ (б); метод независимых компонент (в)

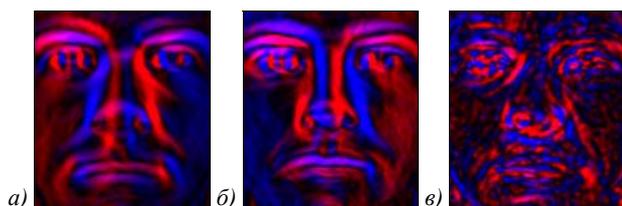


Рис. 2. Примеры базисных векторов с использованием градиентов. Используемые методы снижения размерности: метод главных компонент (а); линейный дискриминантный анализ (б); метод независимых компонент (в)

По представленным результатам можно сделать следующие выводы:

- предлагаемые подходы к описанию полей градиентов/направлений изображений лиц во всех случаях дают более высокое качество классификации по сравнению с «классическим» подходом, использующим поле яркости;
- использование поля направлений в большинстве случаев приводит к более высоким показателям качества по сравнению с полем градиента;
- модельно-ориентированные дескрипторы позволяют снизить число вещественных коэффициентов вектора признаков (при анализе комплекснозначных полей). Это может приводить как к некоторому снижению качественных показателей, так и к их значительному росту (см. вариант для 100 признаков у поля градиентов);
- выигрыш/проигрыш отдельных методов снижения размерности более значим в случае малых размерностей вектора признаков. Для больших размерностей (~100) разница между показателями качества оказывается менее значима (максимальное среднее квадратическое отклонение оценки вероятности составило для $K=1210$ экспериментов $0,5K^{-1/2}=0,0144$), что подтверждает выводы работы [4]: ни один из рассмотренных методов снижения размерности не обеспечивает уверенного преимущества по сравнению с другими.

Заключение

В работе исследованы различные подходы к построению описаний цифровых изображений с использованием линейных методов снижения размерности и дополнительных способов пред- и постобработки. На примере практической задачи распознавания лиц продемонстрировано, что наилучшие качественные пока-

затели достигаются при использовании фазовой составляющей поля градиента – поля направлений вместо исходного полутонового изображения. Показано, что выигрыш/проигрыш отдельных методов линейного снижения размерности более значим в случае малых размерностей получаемого вектора признаков. Также экспериментально показана эффективность использования модельно-ориентированных дескрипторов вместо традиционно применяемых коэффициентов разложения – качество решения задачи при этом возрастает, а число признаков описания снижается.

Дальнейшие направления исследований включают сравнение различных методов и подходов распознавания формируемых описаний с использованием различных тестовых баз данных (benchmark databases типа LFW, FERET и др.) и классификаторов в условиях ограниченных ресурсов.

Благодарности

Исследование выполнено при финансовой поддержке грантов РФФИ в рамках научных проектов № 18-01-00748-а, № 17-29-03190-офи-м в части «Алгоритмы линейного снижения размерности и окончательного описания», Министерства науки и высшего образования РФ в рамках выполнения работ по Государственному заданию ФНИЦ «Кристаллография и фотоника» РАН (Соглашение № 007-ГЗ/Ч3363/26) в части «Результаты экспериментов».

В качестве тестовых данных использовались изображения лиц базы данных «Extended Yale Face Database B» [12, 13].

Литература

1. **Vasan, K.K.** Dimensionality reduction using principal component analysis for network intrusion detection / K.K. Vasan, B. Surendiran // Perspectives in Science. – 2016. – Vol. 8. – P. 510-512. – DOI: 10.1016/j.pisc.2016.05.010.
2. **Wang, J.** Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis / J. Wang, C.-I. Chang // IEEE Transactions on Geoscience and Remote Sensing. – 2006. – Vol. 44, Issue 6. – P. 1586-1600. – DOI: 10.1109/TGRS.2005.863297.
3. **Phinyomark, A.** Application of linear discriminant analysis in dimensionality reduction for hand motion classification / A. Phinyomark, H. Hu, P. Phukpattaranont, C. Limsakul // Measurement Science Review. – 2012. – Vol. 12, No 3. – P. 82-89. – DOI: 10.2478/v10048-012-0015-8.
4. **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.
5. **Subasi, A.** EEG signal classification using PCA, ICA, LDA and support vector machines / A. Subasi, M.I. Gursoy // Expert Systems with Applications. – 2010. – Vol. 37, Issue 12. – P. 8659-8666. – DOI: 10.1016/j.eswa.2010.06.065.
6. **Hese, S.K.** Performance of PCA and LDA for face recognition / S.K. Hese, M.R. Banwaskar // International Journal of Advanced Research in Electronics and Communication Engineering. – 2013. – Vol. 2, Issue 2. – P. 149-154.
7. **Fernandes, S.** Performance analysis of PCA-based and LDA-based algorithms for face recognition / S. Fernandes, J. Bala // International Journal of Signal Processing

- Systems. – 2013. – Vol. 1, Issue 1. – P. 1-6. – DOI: 10.12720/ijsp.1.1.1-6.
8. **Schroff, F.** Facenet: A unified embedding for face recognition and clustering / F. Schroff, D. Kalenichenko, J. Philbin // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 815-823. – DOI: 10.1109/CVPR.2015.7298682.
 9. Методы компьютерной обработки изображений / М.В. Гашников, Н.И. Глумов, 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. Чичёва, В.А. Фурсов; под ред. В.А. Соифера. – Изд. 2-е, испр. – М.: Физматлит, 2003. – 784 с. – ISBN: 5-9221-0270-2.
 10. **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.
 11. **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.
 12. **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.
 13. **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.
 14. **Kanopoulos, N.** Design of an image edge detection filter using the sobel operator / N. Kanopoulos // IEEE Journal of Solid State Circuits. – 1988. – Vol. 23, Issue 2. – P. 358-367. – DOI: 10.1109/4.996.
 15. **Fukunaga K.** Introduction to statistical pattern recognition / K. Fukunaga. – San Diego: Academic Press, 1990. – 592 p. – ISBN: 0-12-269851-7.
 16. **Tzimiropoulos, G.** Principal component analysis of image gradient orientations for face recognition / G. Tzimiropoulos, S. Zafeiriou, M. Pantic // Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition and Workshops. – 2011. – P. 553-558. – DOI: 10.1109/FG.2011.5771457.
 17. **Hyvärinen, A.** Independent component analysis / A. Hyvärinen, J. Karhunen, E. Oja. – New York: Wiley-Interscience, 2001. – 504 p. – ISBN: 978-0-471-40540-5.
 18. **Ruan, Z.** On extending the complex FastICA algorithms to noisy data / Z. Ruan, L. Li, G. Qian // Neural Networks. – 2014. – Vol. 60. – P. 194-202. – DOI: 10.1016/j.neunet.2014.08.013.
 19. **Turk, M.** Face recognition using eigenfaces / M. Turk, A. Pentland // Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 1991. – P. 586-591. – DOI: 10.1109/CVPR.1991.139758.

Сведения об авторах

Дмитриев Егор Андреевич, 1994 года рождения, в 2018 году окончил Самарский национальный исследовательский университет по специальности «Информационная безопасность автоматизированных систем». Область научных интересов: компьютерное зрение, распознавание образов, защита информации, геоинформатика. E-mail: DmitrievEgor94@yandex.ru.

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 более 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: vmyas@geosamara.ru.

ГРНТИ: 28.21.15, 28.21.19.

Поступила в редакцию 21 июня 2018 г. Окончательный вариант – 31 июля 2018 г.

COMPARATIVE STUDY OF DESCRIPTION ALGORITHMS FOR COMPLEX-VALUED GRADIENT FIELDS OF DIGITAL IMAGES USING LINEAR DIMENSIONALITY REDUCTION METHODS

E.A. Dmitriev¹, V.V. Myasnikov^{1,2}

¹ Samara National Research University, 34, Moskovskoye shosse, Samara, 443086, Samara, Russia ;

² IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS, Molodogvardeyskaya 151, 443001, Samara, Russia

Abstract

The paper presents an analysis of various approaches to constructing descriptions for the gradient fields of digital images. The analyzed approaches are based on the well-known methods for data dimensionality reduction, such as Principal (PCA) and Independent (ICA) Component Analysis, Linear Discriminant Analysis (LDA). We apply these methods not to the original image, represented as a two-dimensional field of brightness (a halftone image), but to its secondary representation in the form of a two-dimensional gradient field, that is, a complex-valued image. In this case, approaches based on using both the entire gradient field and only its phase component are considered. In addition, two independent ways of forming the final description of the original object are considered: using expansion coefficients of the gradient field in a derived basis and using an original authors' design that is called

model-oriented descriptors. With the latter, the number of real coefficients used in the description of the original object can be halved. The studies are conducted via solving a face recognition problem. The effectiveness of the analyzed methods is demonstrated by applying them to images from Extended Yale Face Database B. The comparison is made using a nearest neighbor's classifier.

Keywords: face recognition, PCA, ICA, LDA, model-oriented descriptors, The Extended Yale Database B, image description.

Citation: Dmitriev EA, Myasnikov VV. Comparative study of description algorithms for complex-valued gradient fields of digital images using linear dimensionality reduction methods. *Computer Optics* 2018; 42(5): 822-828. – DOI: 10.18287/2412-6179-2018-42-5-822-828.

Acknowledgements: The work was funded by RFBR according to the research projects No. 18-01-00748, 17-29-03190 (“Algorithms of linear dimension reduction and final description”) and by the RF Ministry of Science and Higher Education within the State assignment to the FSRC “Crystallography and Photonics” RAS under contract No. 007-GZ/Ch3363/26 (“Experimental results”).

References

- [1] Vasan KK, Surendiran B. Dimensionality reduction using principal component analysis for network intrusion detection. *Perspectives in Science* 2016; 8: 510-512. DOI: 10.1016/j.pisc.2016.05.010.
- [2] Wang J, Chang C-I. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Trans Geosci Remote Sens* 2006; 44(6): 1586-1600. DOI: 10.1109/TGRS.2005.863297.
- [3] Phinyomark A, Hu H, Phukpattaranont P, Limsakul C. Application of linear discriminant analysis in dimensionality reduction for hand motion classification. *Meas Sci Rev* 2012; 12(3): 82-89. DOI: 10.2478/v10048-012-0015-8.
- [4] 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.
- [5] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst Appl* 2010; 37(12): 8659-8666. DOI: 10.1016/j.eswa.2010.06.065.
- [6] Hese SK, Banwaskar MR. Performance of PCA and LDA for Face Recognition. *International Journal of Advanced Research in Electronics and Communication Engineering* 2013; 2(2): 149-154.
- [7] Fernandes S, Bala J. Performance analysis of PCA-based and LDA-based algorithms for face recognition. *International Journal of Signal Processing Systems* 2013; 1(1): 1-6. DOI: 10.12720/ijsp.1.1.1-6.
- [8] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering. *Proc CVPR* 2015: 815-823. DOI: 10.1109/CVPR.2015.7298682.
- [9] Soifer VA, ed. *Computer image processing, Part II: Methods and algorithms*. Saarbrücken: VDM Verlag; 2009. ISBN: 978-3639175455.
- [10] Myasnikov VV. Description of images using model-oriented descriptors // *Computer Optics* 2017; 41(6): 888-896. DOI: 10.18287/2412-6179-2017-41-6-888-896.
- [11] 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.
- [12] Georghades AS, Belhumeur PN, Kriegman DJ. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans Pattern Anal Mach Intelligence* 2001; 23(6): 643-660. DOI: 10.1109/34.927464.
- [13] Lee KC, Ho J, Kriegman D. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans Pattern Anal Mach Intelligence* 2005; 27(5): 684-698. DOI: 10.1109/TPAMI.2005.92.
- [14] Kanopoulos N. Design of an image edge detection filter using the sobel operator // *IEEE Journal of Solid State Circuits* 1988; 23(2): 358-367. DOI: 10.1109/4.996.
- [15] Fukunaga K. *Introduction to statistical pattern recognition*. San Diego: Academic Press; 1990. ISBN: 0-12-269851-7.
- [16] Tzimiropoulos G, Zafeiriou S, Pantic M. Principal component analysis of image gradient orientations for face recognition. *Face and Gesture* 2011: 553-558. DOI: 10.1109/FG.2011.5771457.
- [17] Hyvärinen A, Karhunen J, Oja E. *Independent component analysis*. New York: Wiley-Interscience; 2001. ISBN: 978-0-471-40540-5.
- [18] Ruan Z, Li L, Qian G. On extending the complex FastICA algorithms to noisy data. *Neural Networks* 2014; 60: 194-202. DOI: 10.1016/j.neunet.2014.08.013.
- [19] Turk M, Pentland A. Face recognition using eigenfaces. *Proc IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1991: 586-591. DOI: 10.1109/CVPR.1991.139758.

Author's information

Egor Andreevich Dmitriev (b. 1994) graduated from Samara National Research University in 2018, majoring in Information Security of Automated Systems. His research interests are currently focused on computer vision, pattern recognition, information security, GIS. E-mail: DmitrievEgor94@yandex.ru.

Vladislav Valerievich Myasnikov (1971 b.), graduated (1994) from the S.P. Korolyov Samara State Aerospace University (SSAU). He received his PhD in Technical Sciences (1998) 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). He is a 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 June 21, 2018. The final version – July 31, 2018.