

УДК 004.932

## АЛГОРИТМЫ ВЗАИМНОЙ ТРАНСФОРМАЦИИ ИЗОБРАЖЕНИЙ ДЛЯ СИСТЕМ ОБРАБОТКИ И ПОИСКА ВИЗУАЛЬНОЙ ИНФОРМАЦИИ

Г.А. Кухарев<sup>a</sup>, Ю.Н. Матвеев<sup>b,c</sup>, А.Л. Олейник<sup>c</sup>

<sup>a</sup> Западноромский технологический университет в Щецине, Щецин, 70-310, Польша

<sup>b</sup> ООО «ЦРТ-инновации», Санкт-Петербург, 196084, Российская Федерация

<sup>c</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

Адрес для переписки: matveev@speechpro.com

### Информация о статье

Поступила в редакцию 29.11.16, принята к печати 25.12.16

doi: 10.17586/2226-1494-2017-17-1-62-74

Язык статьи – русский

**Ссылка для цитирования:** Кухарев Г.А., Матвеев Ю.Н., Олейник А.Л. Алгоритмы взаимной трансформации изображений для систем обработки и поиска визуальной информации // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 1. С. 62–74. doi: 10.17586/2226-1494-2017-17-1-62-74

### Аннотация

**Предмет исследования.** Исследованы методы и алгоритмы взаимной трансформации связанных пар изображений для расширения возможностей технологий кросс-модального мультимедийного поиска (CMMR). Подробно рассмотрена задача взаимного преобразования изображений лиц различной природы (например, фотографий и рисунков). Подобная задача широко представлена на практике. Исследования в данной области основаны на существующих базах данных. Предложенные в настоящей работе алгоритмы могут быть применены для произвольных пар связанных изображений за счет унифицированного математического описания. **Метод.** Предложены три алгоритма преобразования изображений. Первый алгоритм основан на анализе главных компонент и преобразовании Карунена–Лоэва (1DPCA/1DKLT). В отличие от известного аналога, предложенный алгоритм не использует обучающую выборку на этапе трансформации. Второй алгоритм предполагает генерацию популяции изображений. Третий алгоритм реализует трансформацию на основе двумерного анализа главных компонент и преобразования Карунена–Лоэва (2DPCA/2DKLT). **Основные результаты.** Проведенные эксперименты по трансформации изображений лиц и генерации их популяций позволили выявить особенности каждого из алгоритмов. Первый алгоритм позволяет строить точную и устойчивую модель перехода в рамках заданного набора изображений, второй алгоритм дополняет существующие базы новыми изображениями, а третий алгоритм выполняет трансформацию за пределами базы обучения. **Практическая значимость.** С учетом анализа особенностей каждого из предложенных алгоритмов даны рекомендации по их применению. Возможные сценарии включают построение моделей перехода для связанных пар изображений, их взаимную трансформацию в рамках базы и за ее пределами, а также генерацию популяций с целью повышения репрезентативности существующих наборов данных. Предложенные алгоритмы могут быть применены для повышения надежности распознавания лиц по изображениям различной физической природы, а также для решения широкого спектра других задач кросс-модального мультимедийного поиска.

### Ключевые слова

кросс-модальный мультимедийный поиск, метод главных компонент, изображения лиц, скетч, фоторобот

### Благодарности

Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

## MUTUAL IMAGE TRANSFORMATION ALGORITHMS FOR VISUAL INFORMATION PROCESSING AND RETRIEVAL

G.A. Kukharev<sup>a</sup>, Yu.N. Matveev<sup>b,c</sup>, A.L. Oleinik<sup>c</sup>

<sup>a</sup> West Pomeranian University of Technology Szczecin, Szczecin, 70-310, Poland

<sup>b</sup> “STC-Innovation”, Saint Petersburg, 196084, Russian Federation

<sup>c</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

Corresponding author: matveev@speechpro.com

### Article info

Received 29.11.16, accepted 25.12.16

doi: 10.17586/2226-1494-2017-17-1-62-74

Article in Russian

**For citation:** Kukharev G.A., Matveev Yu.N., Oleinik A.L. Mutual image transformation algorithms for visual information processing and retrieval. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2017, vol. 17, no. 1, pp. 62–74. doi: 10.17586/2226-1494-2017-17-1-62-74

#### Abstract

**Subject of Research.** The paper deals with methods and algorithms for mutual transformation of related pairs of images in order to enhance the capabilities of cross-modal multimedia retrieval (CMMR) technologies. We have thoroughly studied the problem of mutual transformation of face images of various kinds (e.g. photos and drawn pictures). This problem is widely represented in practice. Research in this area is based on existing datasets. The algorithms we have proposed in this paper can be applied to arbitrary pairs of related images due to the unified mathematical specification. **Method.** We have presented three image transformation algorithms. The first one is based on principal component analysis and Karhunen-Loève transform (1DPCA/1DKLT). Unlike the existing solution, it does not use the training set during the transformation process. The second algorithm assumes generation of an image population. The third algorithm performs the transformation based on two-dimensional principal component analysis and Karhunen-Loève transform (2DPCA/2DKLT). **Main Results.** The experiments on image transformation and population generation have revealed the main features of each algorithm. The first algorithm allows construction of an accurate and stable model of transition between two given sets of images. The second algorithm can be used to add new images to existing bases and the third algorithm is capable of performing the transformation outside the training dataset. **Practical Relevance.** Taking into account the qualities of the proposed algorithms, we have provided recommendations concerning their application. Possible scenarios include construction of a transition model for related pairs of images, mutual transformation of the images inside and outside the dataset as well as population generation in order to increase representativeness of existing datasets. Thus, the proposed algorithms can be used to improve reliability of face recognition performed on images of various kinds. Moreover, these techniques can be applied to address a wide variety of other CMMR problems.

#### Keywords

cross-modal multimedia retrieval, principal component analysis, face images, sketch, facial composite

#### Acknowledgements

This work was partially financially supported by the Government of the Russian Federation, Grant 074-U01.

### Введение

Одна из первых статей, представляющих технологию поиска цифровых изображений по словесному описанию (Content-Based Image Retrieval, CBIR) [1], появилась около 25 лет назад. К настоящему времени технология CBIR достигла внушительных успехов, в том числе за счет использования многоуровневых поисковых систем и механизмов, включающих интерактивные стратегии поиска. Однако, несмотря на это, пользователь зачастую не в состоянии связать полученные результаты поиска и введенный запрос. Нередко возникает ситуация, когда пользователь не может понять, насколько хорошо результаты поиска соответствуют запросу, или получает нерелевантные, но подходящие по внешним признакам (например, по цветовой гамме) результаты.

Такие случаи, как правило, являются следствием двух основных причин. Первая из них – ориентация технологии CBIR на класс цифровых изображений, представляющих только типичную визуальную информацию – изображения, полученные только в видимом свете. Вторая причина – «семантический разрыв» между низкоуровневым обобщенным представлением изображения (цветовая гамма, текстура и форма объектов, входящих в изображение) и семантикой этого изображения, необходимой для его понимания, описания и дальнейшего его использования.

Неоднозначность запроса на поиск и его результата в совокупности с отмеченными выше причинами стимулировали появление новых тенденций в технологии поиска информации по содержанию.

Одна из них – расширение некоторого набора изображений дополнительными модальностями, например, звуком и текстом. Эта тенденция определилась как технология «Cross-Modal Multimedia Retrieval» (CMMR) [2]. Другая тенденция – составление отдельного класса из семантически связанных изображений, но представляющих различные способы их получения. Примерами здесь могут быть изображения лиц (ИЛ), используемые в биометрических системах поиска и представленные в видимом и инфракрасном («тепловом») спектрах, а также в форме рисованных ИЛ и композиционных фотороботов (которые далее обозначаются общим термином «скетч»). В англоязычной литературе такие изображения относятся к классу гетерогенных ИЛ, а соответствующая технология их поиска определилась как «Heterogeneous Face Recognition» (HFR) [3, 4].

В исходном пространстве признаков в технологии CMMR используются наборы данных разной модальности (например, изображение и текст), а в технологии HFR используются наборы данных разной физической природы («мультисенсорной природы»), например, скетч и фотопортрет, фотография в видимом и в ближнем инфракрасном спектре. Чтобы реализовать поиск, оперировать подобными данными и интерпретировать их, необходимо связать их между собой (например, изображение и выбранную для него текстовую информацию или два изображения разной сенсорной природы). Далее, если предположить, что между данными из выбранных модальностей есть взаимное соответствие, то при запросе на поиск данных, входящих в один из наборов, можно легко найти соответствующий ему результат из другого набора. Необходимое взаимное соответствие (подобное взаимному индексированию) достигается на этапах обучения по выборкам многомодальных данных. Один из подходов такого обучения реализуется с помо-

стью методов проекции на собственные базисы – анализа главных компонент (Principal Component Analysis, PCA), проекций на латентные структуры (Projection Latent Structures, Partial Least Squares, PLS) и канонического корреляционного анализа (Canonical Correlation Analysis, CCA). Эти методы позволяют отобразить два набора исходных данных (часто не коррелированных и семантически не связанных между собой) в новое общее подпространство признаков, в котором их проекции сильно коррелируют. Последнее и позволяет связать выбранные пары данных с помощью общей модели, чтобы представлять и понимать одни наблюдения, действия (или явления) через другие, устраняя тем самым полностью или частично, упомянутый выше «семантический разрыв» [5].

С учетом изложенного в настоящей работе предлагаются новые алгоритмы взаимной трансформации одних изображений в другие в наборах пар изображений. Рассмотрены различные сценарии их применения. В первом из них (этот простой сценарий рассмотрен в [6]) предполагается наличие семантической и корреляционной связи между изображениями. Во втором присутствует семантическая связь между парами изображений без корреляции. В третьем сценарии изображения в соответствующих парах не только не коррелируют между собой, но и могут быть семантически не связаны между собой. Представленный в данной работе набор алгоритмов позволяет обрабатывать изображения в условиях любого из перечисленных сценариев, что существенно расширяет возможности использования предлагаемых решений в задачах и технологиях CMMR и HFR.

В основу предложенных в данной работе алгоритмов положена идея реконструкции скетча по исходному ИЛ [6], которая была расширена и обобщена в [7]. Этот подход включает проекцию на собственные базисы и опирается на предположение о структурном подобии ИЛ в виде фотографий и скетчей. Здесь и далее предполагается, что фотографии и скетчи представлены в виде цифровых изображений.

В настоящей работе рассмотрены различные варианты применения одномерного и двумерного метода главных компонент (1DPCA/1DKLT и 2DPCA/2DKLT) [8] к задаче взаимной трансформации ИЛ.

### Базовый алгоритм трансформации

Алгоритм [6] изначально был разработан для взаимной трансформации фотографий лиц и соответствующих им скетчей. Он заключается в вычислении собственного базиса для скетчей (фотографий), с помощью которого любой новый скетч (фотография) может быть представлен в виде линейной комбинации скетчей (фотографий) из обучающей выборки. С учетом вышеупомянутого предположения о структурном подобии скетчи (фотографии) в этой линейной комбинации заменяются на фотографии (скетчи). Суммирование взвешенных изображений обучающей выборки завершает процесс трансформации.

Применение PCA для анализа исходных данных позволяет выделить общие характеристики анализируемых изображений (в данном случае – фотографий или скетчей). Однако этот метод не предполагает построения какой-либо регрессионной модели. Это значительно упрощает процедуру анализа, однако накладывает дополнительные ограничения в виде предположения о структурном подобии фотографий и скетчей. Другим недостатком данного метода является явное использование всей обучающей выборки в процессе каждой трансформации скетча в фотографию (и наоборот), что отрицательным образом сказывается на вычислительной сложности. Более того, авторы алгоритма [6] не рассматривали возможность его применения в условиях, когда предположение о структурном подобии не выполняется.

В настоящей работе предложено несколько вариантов расширения функциональных возможностей и области применения базового алгоритма.

1. Представлена версия алгоритма, не использующая в явном виде обучающую выборку на этапе трансформации. Вместо этого в процессе анализа главных компонент один раз формируются только две матрицы взаимной трансформации изображений, которые далее и используются, что позволяет снизить вычислительные затраты на этапе преобразования скетча в фотографию (и наоборот).
2. На основе алгоритма с использованием матриц трансформации разработан алгоритм генерации популяций изображений.
3. Предложен метод построения матриц взаимной трансформации изображений в рамках двумерного анализа главных компонент и преобразования Карунена–Лоэва (2DPCA/2DKLT), который по сравнению с базовым алгоритмом обладает лучшей обобщающей способностью.

### Преобразования изображений в связанных парах на основе матриц взаимной трансформации

Чтобы отказаться от использования обучающей выборки на этапе трансформации, необходима некоторая модель перехода между изображениями различной природы. В данном разделе представлена модель на основе матриц перехода «скетч–фотография»  $T_{SP}$  и «фотография–скетч»  $T_{PS}$ . Ниже описан алгоритм, позволяющий их построить. В табл. 1 представлен список условных обозначений. Обучающая выборка включает два набора изображений: набор фотографий и набор скетчей. Они преобразуются в векторы посредством конкатенации их столбцов. Полученные векторы объединяются в матрицы данных  $S$  и  $P$ .

**Замечание.** Здесь и далее предполагается, что исходные выборки фотографий и скетчей  $S$  и  $P$  центрированы относительно средних изображений.

Обозначение	Описание
$K$	количество пар «фотография–скетч» в базе
$M$	количество строк в каждом из изображений
$N$	количество столбцов в каждом из изображений
$\mathbf{S}, \mathbf{P}$	матрицы скетчей и фотографий (имеют размер $MN \times K$ )
$\mathbf{U}_S, \mathbf{U}_P$	матрицы собственных векторов для ковариационных матриц $\mathbf{S}\mathbf{S}^T$ и $\mathbf{P}\mathbf{P}^T$ (имеют размер $MN \times K$ )
$\Lambda_S, \Lambda_P$	диагональные матрицы собственных чисел для матриц Грама $\mathbf{S}^T\mathbf{S}$ и $\mathbf{P}^T\mathbf{P}$ (имеют порядок $K$ )
$\mathbf{V}_S, \mathbf{V}_P$	матрицы собственных векторов для матриц Грама $\mathbf{S}^T\mathbf{S}$ и $\mathbf{P}^T\mathbf{P}$ (имеют порядок $K$ )
$\mathbf{S}_{new}, \mathbf{P}_{new}$	вектор-столбец, содержащий $MN$ элементов, полученный конкатенацией столбцов нового скетча (фотографии), подлежащего обработке
$\mathbf{J}_S, \mathbf{J}_P$	векторы-столбцы, содержащие по $K$ элементов, представляющие собой проекции $\mathbf{S}_{new}$ и $\mathbf{P}_{new}$ на собственные пространства
$\mathbf{T}_{SP}, \mathbf{T}_{PS}$	матрицы перехода «скетч–фотография» и «фотография–скетч» (имеют размер $MN \times K$ )

Таблица 1. Список основных условных обозначений

Матрицы данных  $\mathbf{S}$  и  $\mathbf{P}$  могут быть представлены в виде произведений трех матриц посредством сингулярного разложения:

$$\mathbf{S} = \mathbf{U}_S \Lambda_S^{\frac{1}{2}} \mathbf{V}_S^T, \mathbf{P} = \mathbf{U}_P \Lambda_P^{\frac{1}{2}} \mathbf{V}_P^T.$$

Так как в большинстве случаев  $K < MN$ , ковариационные матрицы  $\mathbf{S}\mathbf{S}^T$  и  $\mathbf{P}\mathbf{P}^T$  являются вырожденными. В связи с этим сначала следует вычислять матрицы  $\mathbf{V}_S, \Lambda_S$  и  $\mathbf{V}_P, \Lambda_P$  как решения задач на собственные значения для матриц Грама  $\mathbf{S}^T\mathbf{S}$  и  $\mathbf{P}^T\mathbf{P}$ . Затем могут быть вычислены матрицы  $\mathbf{U}_S$  и  $\mathbf{U}_P$ :

$$\mathbf{U}_S = \mathbf{S} \left( \mathbf{V}_S \Lambda_S^{-\frac{1}{2}} \right), \mathbf{U}_P = \mathbf{P} \left( \mathbf{V}_P \Lambda_P^{-\frac{1}{2}} \right).$$

По сути, данная процедура представляет собой анализ главных компонент, выполняемый отдельно на обучающей выборке фотографий и скетчей.

Далее вычисляются матрицы трансформации  $\mathbf{T}_{SP}$  и  $\mathbf{T}_{PS}$ . Определение этих матриц учитывает предположение о структурном подобии:

$$\mathbf{T}_{SP} \stackrel{\text{def}}{=} \mathbf{P} \left( \mathbf{V}_S \Lambda_S^{-\frac{1}{2}} \right), \mathbf{T}_{PS} \stackrel{\text{def}}{=} \mathbf{S} \left( \mathbf{V}_P \Lambda_P^{-\frac{1}{2}} \right).$$

Столбцы этих матриц можно рассматривать в качестве «смешанных» базисов для систем «скетчи–фотографии». Они могут быть использованы для аппроксимации фотографии (скетча) по соответствующему вектору из собственного пространства скетчей (фотографий). Вычисление матриц  $\mathbf{T}_{SP}, \mathbf{T}_{PS}$  завершает процесс построения модели взаимного перехода.

**Замечание.** В представленном методе обучающая выборка не используется в явном виде на этапе генерации. Вместо нее используются матрицы  $\mathbf{T}_{SP}$  и  $\mathbf{T}_{PS}$ . Если  $K$  велико, имеет смысл отбросить столбцы матриц  $\mathbf{U}_S, \mathbf{T}_{SP}$  и  $\mathbf{U}_P, \mathbf{T}_{PS}$ , соответствующие малым собственным значениям. Это поможет значительно снизить вычислительные затраты без существенных потерь качества генерации.

Взаимная трансформация изображений осуществляется следующим образом. Новый скетч (или фотография) преобразуется в вектор  $\mathbf{S}_{new}$  и (или  $\mathbf{P}_{new}$ ) посредством конкатенации столбцов. После этого посредством KLT вычисляется проекция на собственное пространство:

$$\mathbf{J}_S = \mathbf{U}_S^T \mathbf{S}_{new}, \mathbf{J}_P = \mathbf{U}_P^T \mathbf{P}_{new}. \tag{1}$$

Далее выполняется умножение на матрицу трансформации:

$$\mathbf{P}_{recon} = \mathbf{T}_{SP} \mathbf{J}_S, \mathbf{S}_{recon} = \mathbf{T}_{PS} \mathbf{J}_P.$$

Процесс трансформации завершается преобразованием  $\mathbf{P}_{recon}$  (или  $\mathbf{S}_{recon}$ ) в изображение размера  $M \times N$  (операция, обратная конкатенации столбцов).

В табл. 3 приведены псевдокоды для описанных процедур построения модели и трансформации.

В силу субъективности семантического восприятия изображений лиц человеком и отсутствия однозначной связи между фотографиями и скетчами в парах обучающего набора предложить метрику качества генерируемых изображений и меру подобия между ними практически невозможно. Исходя из этого, выводы о качестве работы алгоритма сделаны на основе визуальной оценки полученных изображений.

Тестирование выполнено на обучающем наборе связанных пар изображений, примеры которых показаны в первых двух строках (рядах) на рис. 1. При этом первая строка определяет набор данных  $\mathbf{P}$ , а вторая строка – набор данных  $\mathbf{S}$ . Набор  $\mathbf{P}$  составлен из изображений лиц, представленных в видимом свете. Набор  $\mathbf{S}$  составлен из изображений разных категорий: нарисованные художником скетчи (столбцы 1 и 2) [9, 10], изображения в ближнем инфракрасном спектре (столбцы 3 и 4), контурные скетчи (столбцы 5 и 6) [11] и изображения морд собак (столбцы 7 и 8) [12]. Заметим, что в обучающем наборе столбцы 1, 2, 3 и 4 включают семантически связанные между собой изображения в парах, столбцы 5 и 6 включают изо-

бражения со слабой корреляционной связью, а столбцы 7 и 8 – изображения с высокой корреляцией яркостных гистограмм и геометрических форм, но совершенно не связанные семантически. Таким образом, предположение о структурном подобии может выполняться не для всех изображений использованного обучающего набора.

Отметим, что на данных обучающей выборки реконструкция в результате трансформации изображений из набора  $S$  в изображения набора  $P$  в соответствующих парах выполняется безошибочно (см. строку 3). Очевидно, что такие результаты будут получены независимо от характеристик исходных пар изображений, поскольку эта реконструкция основана на использовании набора собственных изображений, точно соответствующего исходным данным, и не выходит за его рамки. С другой стороны, даже при сильном понижении разрешения новых данных и добавлении шума результаты трансформации можно считать удовлетворительными (см. строки 4 и 5). Исключение составляют только результаты с контурными скетчами (столбцы 5, 6 и строки 4, 5). Это связано с тем, что «собственные контурные скетчи» были поглощены шумом и потеряли полезную информацию, поскольку вся информация была сосредоточена в малом наборе точек, отличающихся от фона. Таким образом, предложенный алгоритм может быть использован для кодирования изображений с целью сокрытия их содержания, а также для поиска изображений в условиях низкокачественной поисковой информации.



Рис. 1. Обучающая выборка и результаты трансформации. В рядах 1 и 2 представлены обучающие наборы данных. Ряд 3 – результат реконструкции по изображениям ряда 2. Ряд 4 – изображения ряда 2 в низком разрешении с добавлением шума. Ряд 5 – результат трансформации изображений ряда 4

На рис. 2 представлены результаты трансформации скетчей из тестовой выборки. При этом в столбце 2 представлен скетч, сгенерированный из фотографии по способу [13]. Алгоритм позволяет обрабатывать такие скетчи без дополнительной настройки, несмотря на то, что их не было в обучающей выборке.

В целом результаты на тестовой выборке заметно хуже, чем на обучающей. При этом качество трансформации тем ниже, чем менее коррелированы фотографии и скетчи. Можно заметить, что результат трансформации зависит от того, были ли в обучающей выборке подобные изображения. Кроме того, алгоритм чувствителен к изменениям сдвига и масштаба.

Для получения удовлетворительных результатов за пределами обучающей выборки следует учитывать предположение о структурном подобии, использовать объемную и репрезентативную базу обучения и нормировать исходное изображение по положению и масштабу.



Устойчивая работа алгоритма в пределах заданного набора произвольных изображений позволяет строить надежную модель перехода для этого набора. Ниже показано, как такая модель может быть использована для генерации популяций.

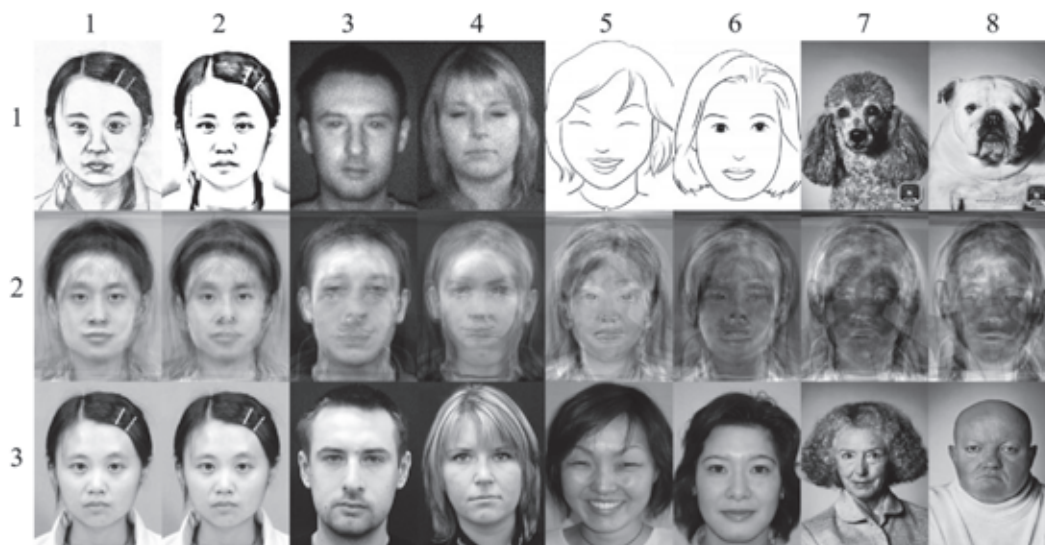


Рис. 2. Трансформация тестовых скетчей. Ряд 1 – исходные скетчи, ряд 2 – результат трансформации скетчей из ряда 1 в фотографию, ряд 3 – оригинальные фотографии

### Генерация популяций изображений

При обработке связанных пар изображений часто возникает проблема неоднозначности соответствия между ними. Например, скетчи, нарисованные разными художниками и (или) на основе показаний различных свидетелей, будут существенно отличаться друг от друга и от оригинальной фотографии, что видно из рис. 3. При этом в совокупности такой расширенный состав скетчей (фотороботов) улучшает решение задачи поиска или распознавания фотографии по скетчу.

С другой стороны, на практике зачастую приходится иметь дело с единственным свидетелем и, следовательно, с единственным скетчем, что приводит к малой репрезентативности исходных данных. Для ее повышения может быть использован представленный ниже подход. Он основан на генерации популяций скетчей и фотографий в рамках метода взаимной трансформации изображений.

Фотография	Скетчи				
	Художник-1	Художник-2	Художник-3	Художник-4	Художник-5

Рис. 3. Скетчи одного и того же человека, выполненные различными художниками [14]

Этот подход предполагает трансформацию скетча в популяцию фотоизображений (что необходимо в задачах распознавания фотографий по скетчам) и наоборот (что необходимо при создании расширенных баз фотографий и скетчей) [15, 16].

В табл. 2 приведены дополнительные обозначения, используемые в этом разделе.

Обозначение	Описание
$L$	заданное количество изображений в генерируемой популяции
$\xi_l, \eta_l$	случайные векторы-столбцы, содержащие $K$ элементов
$\tilde{H}_S, \tilde{H}_P$	матрицы, столбцы которых представляют собой модифицированные версии $J_S$ и $J_P$ (имеют размер $K \times L$ )
$P_{pop}, S_{pop}$	матрицы популяций фотографий и скетчей (имеют размер $MN \times L$ )

Таблица 2. Дополнительные обозначения, используемые для описания алгоритма генерации популяций

Для генерации популяции изображений используется та же модель, что и для их трансформации. При этом процедура трансформации несколько модифицируется. В проекции  $J_S$  и  $J_P$  (1) вводятся случайные изменения, в результате чего формируются матрицы модифицированных проекций:

$$\tilde{H}_S = [J_S + \xi_1, \dots, J_S + \xi_l, \dots, J_S + \xi_L],$$

$$\tilde{H}_P = [J_P + \eta_1, \dots, J_P + \eta_l, \dots, J_P + \eta_L].$$

Далее выполняется их трансформация:

$$P_{pop} = T_{SP}\tilde{H}_S, S_{pop} = T_{PS}\tilde{H}_P.$$

Процесс завершает преобразование каждого из столбцов матриц  $P_{pop}$  и  $S_{pop}$  в изображение размером  $M \times N$  (операция, обратная конкатенации столбцов).

На рис. 4 представлена схема алгоритма генерации популяций фотографий и скетчей. В табл. 3 приведены алгоритмы трансформации скетчей в фотографии и их популяции. Для краткости и ясности изложения алгоритмы обратного преобразования не приведены: они полностью аналогичны приведенным в табл. 3. Заметим, что алгоритм трансформации изображений является частным случаем алгоритма генерации популяций (при  $L = 1, \xi_1 \equiv \eta_1 \equiv 0$ ).

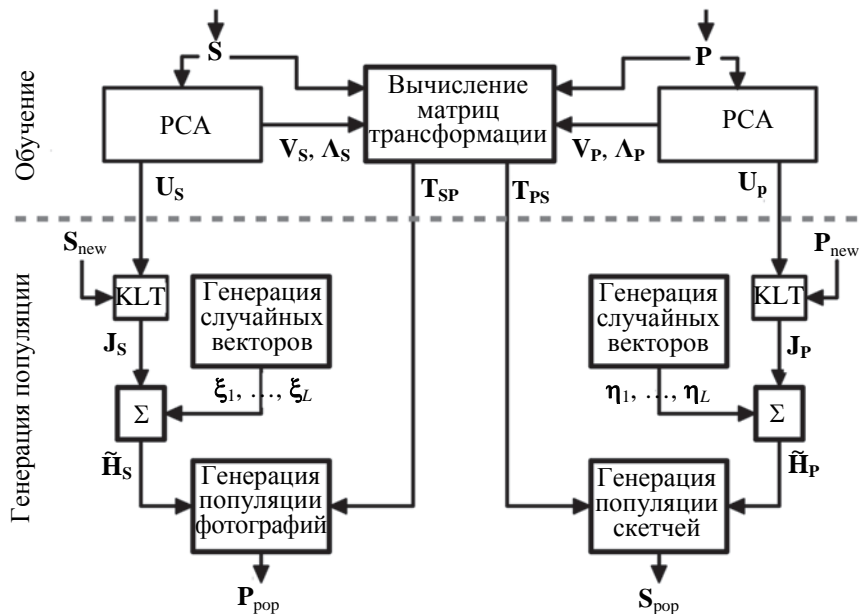


Рис. 4. Схема предложенного алгоритма трансформации и генерации популяций

Обучение	
<p><b>Вход:</b> <math>S, P</math>.  <b>Выход:</b> <math>U_S, T_{SP}</math>.  <b>Алгоритм:</b></p> <ol style="list-style-type: none"> <li>1. Вычислить матрицы <math>V_S</math> и <math>\Lambda_S</math> как решение задачи на собственные значения для матрицы Грама <math>S^T S</math>;</li> <li>2. <math>U_S \leftarrow S (V_S \Lambda_S^{-\frac{1}{2}})</math>;</li> <li>3. <math>T_{SP} \leftarrow P (V_S \Lambda_S^{-\frac{1}{2}})</math>.</li> </ol>	
<p style="text-align: center;"><b>Трансформация</b></p> <p><b>Вход:</b> <math>U_S, T_{SP}, S_{new}</math>.  <b>Выход:</b> фотография, полученная из <math>S_{new}</math>.  <b>Алгоритм:</b></p> <ol style="list-style-type: none"> <li>1. <math>J_S \leftarrow U_S^T S_{new}</math>;</li> <li>2. <math>P_{recon} = T_{SP} J_S</math>;</li> <li>3. Преобразовать <math>P_{recon}</math> в изображение размером <math>M \times N</math>.</li> </ol>	<p style="text-align: center;"><b>Генерация популяции</b></p> <p><b>Вход:</b> <math>U_S, T_{SP}, S_{new}, L</math>.  <b>Выход:</b> популяция из <math>L</math> фотографий.  <b>Алгоритм:</b></p> <ol style="list-style-type: none"> <li>1. <math>J_S \leftarrow U_S^T S_{new}</math>;</li> <li>2. Сгенерировать <math>L</math> случайных векторов <math>\xi_i</math>;</li> <li>3. <math>\tilde{H}_S \leftarrow [J_S + \xi_1, \dots, J_S + \xi_l, \dots, J_S + \xi_L]</math>;</li> <li>4. <math>P_{pop} \leftarrow T_{SP} \tilde{H}_S</math>;</li> <li>5. Преобразовать каждый из <math>L</math> столбцов матрицы <math>P_{pop}</math> в изображение размером <math>M \times N</math> путем поочередного заполнения столбцов изображения.</li> </ol>

Таблица 3. Алгоритмы построения модели перехода, трансформации скетча в фотографию и генерации популяций

Алгоритм генерации популяций может применяться для расширения существующих баз фотографий и скетчей. И в качестве  $S_{new}$  и  $P_{new}$  могут выступать также изображения из обучающей выборки. В таком случае сгенерированные популяции будут обладать достаточно высоким качеством, что позволит существенно улучшить репрезентативность базы фотографий и скетчей.

На рис. 5 показан пример сгенерированной популяции и ее визуализация в трехмерном пространстве. На рис. 6 показаны примеры популяций для скетчей из обучающей и тестовой выборок. Можно отметить, что качество популяций для скетчей из обучающей выборки несколько выше, чем для тестовых скетчей. В целом к алгоритму генерации популяций применимы те же ограничения, что и к предложенному выше алгоритму трансформации.

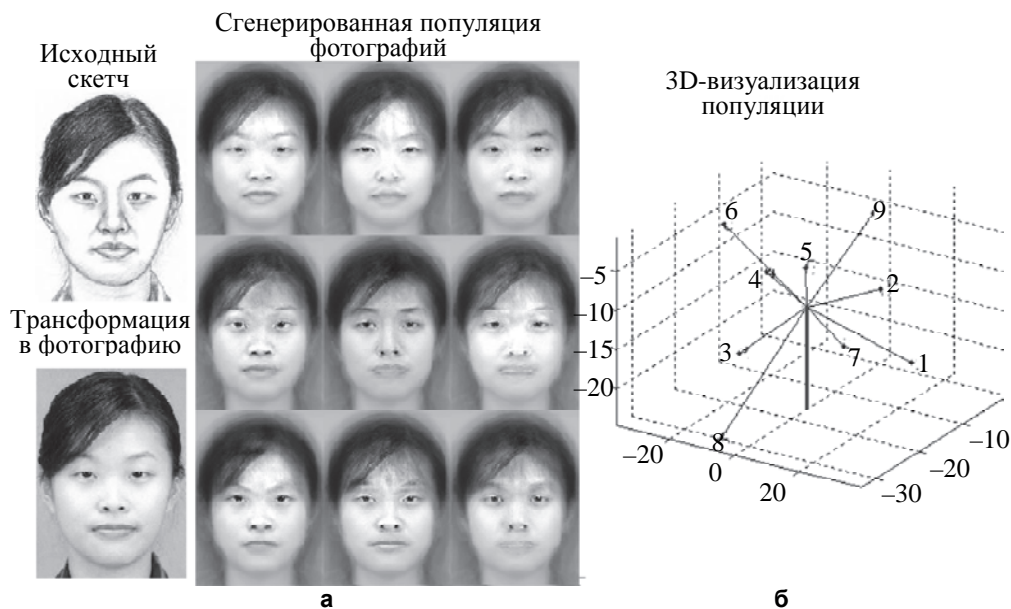


Рис. 5. Популяции фотографий, сгенерированные из скетча (а), и трехмерная визуализация проекций фотографий на три первые главные компоненты (б). Каждая фотография популяции соответствует одной точке на графике

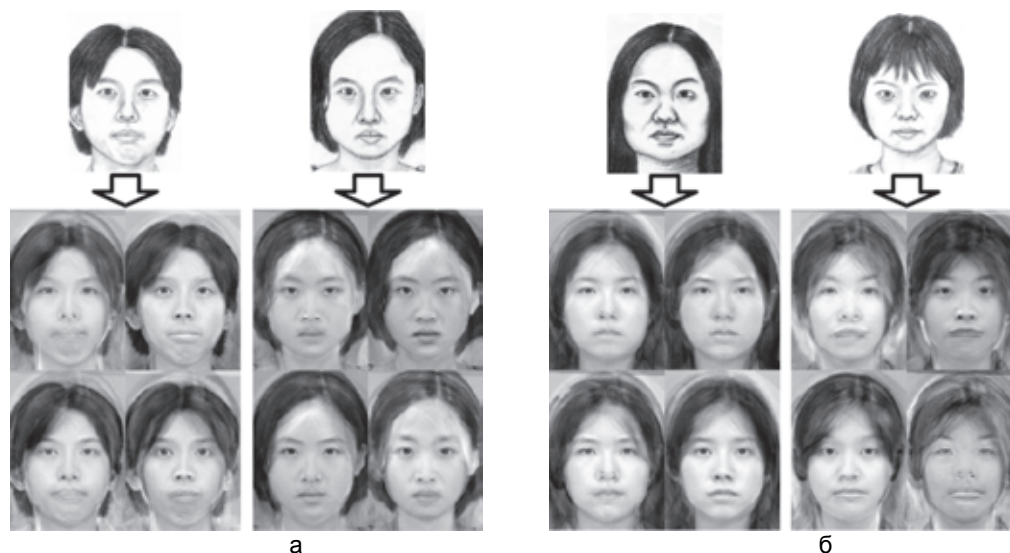


Рис. 6. Популяции фотографий, сгенерированные на основе скетчей из обучающей (а) и тестовой (б) выборки

Важно то, что изменения в пределах одной популяции достаточно значительны, чтобы сгенерированные изображения не были идентичными, но при этом сохраняли подобие с исходным скетчем. Это позволяет применять алгоритм для повышения надежности систем распознавания, работающих с фотографиями и скетчами.

Кроме того, возможно применение алгоритма для дополнения существующих баз новыми изображениями с целью повышения их репрезентативности. В этом случае вся обрабатываемая база выступает в роли обучающей выборки, поэтому качество новых изображений будет достаточно высоким.



**Двумерный метод трансформации изображений**

Одна из наиболее существенных проблем при обработке изображений с помощью PCA – это проблема малой выборки (Small Sample Size, SSS). Она заключается в том, что размерность исходного пространства признаков велика (т.е. равна  $MN$ ). В случаях, когда  $K \ll MN$ , проблему SSS можно обойти с помощью матрицы Грама (как это было сделано выше). Если же  $K$  велико, то PCA предполагает решение задачи на собственные значения для матриц больших размеров, что требует значительных вычислительных затрат. Размерность задачи можно искусственно снизить путем приведения исходных изображений к меньшему масштабу, что, однако, ведет к потере информации.

Принципиально иной подход заключается в применении процедур двумерного анализа главных компонент и двумерного преобразования Карунена–Лоэва (2DPCA/2DKLT). Идеи этого метода представлены в [17] и развиты в работах [18, 19], после чего аналогичный подход был использован для разработки методов двумерного канонического корреляционного анализа (2D CCA) и двумерного метода проекции на латентные структуры (2D PLS) [20]. Подробное описание этих алгоритмов, их характеристики и практические рекомендации по применению можно найти в [8].

Суть метода 2DPCA заключается в решении двух задач на собственные значения: для строк и столбцов изображений обучающей выборки. Таким образом, в каждом случае размерность исходного пространства признаков равна  $M$  или  $N$ , а объем выборки –  $KN$  и  $KM$  соответственно. В связи с этим задача на собственные значения решается напрямую для ковариационных матриц, имеющих размеры  $M \times M$  и  $N \times N$ . Таким образом, двумерные проекционные методы эффективно решают проблему малой выборки и повышают устойчивость решения [18].

В данном разделе представлен алгоритм трансформации изображений на основе 2DPCA/2DKLT. Аналогично представленному выше двумерному методу, он использует две матрицы трансформации. Для краткости изложения здесь рассмотрен случай трансформации скетчей в фотографии, так как обратная процедура выполняется аналогичным образом. В табл. 4 приведены дополнительные обозначения, использованные в данном разделе.

Обозначение	Описание
$\mathbf{p}_1, \dots, \mathbf{p}_K$	Исходная выборка $K$ фотографий размера $M \times N$
$\mathbf{s}_1, \dots, \mathbf{s}_K$	Исходная выборка $K$ скетчей размера $M \times N$
$\mathbf{P}_{col} = (\mathbf{p}_1 \dots \mathbf{p}_K)$	Матрица, составленная из $K$ фотографий (имеет размер $M \times NK$ )
$\mathbf{S}_{col} = (\mathbf{s}_1 \dots \mathbf{s}_K)$	Матрица, составленная из $K$ скетчей (имеет размер $M \times NK$ )
$\mathbf{P}_{row} = (\mathbf{p}_1^T \dots \mathbf{p}_K^T)$	Матрица, составленная из $K$ транспонированных фотографий (имеет размер $N \times MK$ )
$\mathbf{S}_{row} = (\mathbf{s}_1^T \dots \mathbf{s}_K^T)$	Матрица, составленная из $K$ транспонированных скетчей (имеет размер $N \times MK$ )
$\mathbf{C}_{S_{left}} = \mathbf{S}_{col} \mathbf{S}_{col}^T$	Ковариационная матрица порядка $M$
$\mathbf{C}_{S_{right}} = \mathbf{S}_{row} \mathbf{S}_{row}^T$	Ковариационная матрица порядка $N$
$\mathbf{U}_{S_{left}}, \mathbf{U}_{S_{right}}$	Матрицы собственных векторов для $\mathbf{C}_{S_{left}}$ и $\mathbf{C}_{S_{right}}$ порядков $M$ и $N$
$\mathbf{\Lambda}_{S_{left}}, \mathbf{\Lambda}_{S_{right}}$	Матрицы собственных чисел для $\mathbf{C}_{S_{left}}$ и $\mathbf{C}_{S_{right}}$ порядков $M$ и $N$
$\mathbf{V}_{S_{left}}, \mathbf{V}_{S_{right}}$	Матрицы собственных векторов для матриц Грама $\mathbf{S}_{col}^T \mathbf{S}_{col}$ и $\mathbf{S}_{row}^T \mathbf{S}_{row}$ (имеют размеры $NK \times M$ и $MK \times N$ )
$\mathbf{T}_{SP_{left}}, \mathbf{T}_{SP_{right}}$	Матрицы двумерной трансформации (имеют размеры $M \times M$ и $N \times N$ )
$\mathbf{s}_{new}$	Новый скетч размера $M \times N$ , подлежащий обработке
$\mathbf{J}_{sc}$	Двумерная проекция скетча на собственное пространство
$\tilde{\mathbf{p}}$	Результат трансформации нового скетча $\mathbf{s}_{new}$ в фотографию

Таблица 4. Дополнительные обозначения, используемые для описания двумерного алгоритма трансформации

**Замечание.** Как и в одномерном случае, предполагается, что исходные выборки фотографий и скетчей  $\mathbf{p}_k$  и  $\mathbf{s}_k$  центрированы относительно средних изображений.

2DPCA предполагает вычисление двух ковариационных матриц:

$$\mathbf{C}_{S_{left}} = \mathbf{S}_{col} \mathbf{S}_{col}^T = \sum_{k=1}^K \mathbf{s}_k \mathbf{s}_k^T,$$

$$\mathbf{C}_{S_{right}} = \mathbf{S}_{row} \mathbf{S}_{row}^T = \sum_{k=1}^K \mathbf{s}_k^T \mathbf{s}_k.$$

Решения задач на собственные значения для этих матриц имеют следующий вид:

$$\mathbf{C}_{S_{left}} = \mathbf{U}_{S_{left}} \mathbf{\Lambda}_{S_{left}} \mathbf{U}_{S_{left}}^T,$$

$$\mathbf{C}_{S_{right}} = \mathbf{U}_{S_{right}} \mathbf{\Lambda}_{S_{right}} \mathbf{U}_{S_{right}}^T.$$

В свою очередь, матрицы  $V_{S_{left}}$ ,  $V_{S_{right}}$  могут быть выражены в следующей форме:

$$V_{S_{left}} = S_{col}^T U_{S_{left}} \Lambda_{S_{left}}^{-\frac{1}{2}}, \tag{2}$$

$$V_{S_{right}} = S_{row}^T U_{S_{right}} \Lambda_{S_{right}}^{-\frac{1}{2}}. \tag{3}$$

Аналогично тому, как это было сделано в одномерном случае, производится замена матрицы  $S_{col}$  на  $P_{col}$  и  $S_{row}$  на  $P_{row}$ . Тогда определение матриц трансформации выглядит следующим образом:

$$T_{SP_{left}} \stackrel{\text{def}}{=} P_{col} \left( V_{S_{left}} \Lambda_{S_{left}}^{-\frac{1}{2}} \right) \stackrel{(2)}{=} P_{col} S_{col}^T U_{S_{left}} \Lambda_{S_{left}}^{-1} = \left( \sum_{k=1}^K p_k s_k^T \right) U_{S_{left}} \Lambda_{S_{left}}^{-1},$$

$$T_{SP_{right}} \stackrel{\text{def}}{=} P_{row} \left( V_{S_{right}} \Lambda_{S_{right}}^{-\frac{1}{2}} \right) \stackrel{(3)}{=} P_{row} S_{row}^T U_{S_{right}} \Lambda_{S_{right}}^{-1} = \left( \sum_{k=1}^K p_k^T s_k \right) U_{S_{right}} \Lambda_{S_{right}}^{-1}.$$

Вычисление матриц трансформации завершает процесс обработки обучающей выборки.

Как и в одномерном случае, трансформация нового скетча  $s_{new}$  начинается с проекции на собственное пространство. Однако здесь используется двумерное преобразование Карунена–Лозва (2DKLT):

$$J_{s_{rc}} = U_{S_{left}}^T s_{new} U_{S_{right}}.$$

Преобразование в фотографию выполняется посредством умножения на матрицы трансформации:

$$\tilde{p} = T_{SP_{left}} J_{s_{rc}} T_{SP_{right}}^T.$$

В табл. 5 представлены алгоритмы, реализующие описанный двумерный метод. Заметим, что в отличие от одномерного алгоритма, PCA реализован через ковариационные матрицы, а не матрицы Грама. Это возможно в силу отсутствия проблемы малой выборки. Несложно видеть, что все вычисления выполняются над матрицами небольших размеров ( $M \times M$  или  $N \times N$ ). В этом случае двумерный алгоритм трансформации не требует ни значительных вычислительных затрат, ни больших объемов памяти, а решения задач на собственные значения обладают устойчивостью.

Обучение	Трансформация
<b>Вход:</b> $p_1, \dots, p_K, s_1, \dots, s_K$ . <b>Выход:</b> $U_{S_{right}}, U_{S_{left}}, T_{SP_{right}}, T_{SP_{left}}$ . <b>Алгоритм:</b> 1. $C_{S_{left}} \leftarrow \sum_{k=1}^K s_k s_k^T, C_{S_{right}} \leftarrow \sum_{k=1}^K s_k^T s_k$ ; 2. Вычислить $U_{S_{left}}, \Lambda_{S_{left}}$ и $U_{S_{right}}, \Lambda_{S_{right}}$ как решения задачи на собственные значения для $C_{S_{left}}$ и $C_{S_{right}}$ ; 3. $T_{SP_{left}} \leftarrow \left( \sum_{k=1}^K p_k s_k^T \right) U_{S_{left}} \Lambda_{S_{left}}^{-1}$ , $T_{SP_{right}} \leftarrow \left( \sum_{k=1}^K p_k^T s_k \right) U_{S_{right}} \Lambda_{S_{right}}^{-1}$ .	<b>Вход:</b> $U_{S_{right}}, U_{S_{left}}, T_{SP_{right}}, T_{SP_{left}}, s_{new}$ . <b>Выход:</b> $\tilde{p}$ . <b>Алгоритм:</b> 1. $J_{s_{rc}} \leftarrow U_{S_{left}}^T s_{new} U_{S_{right}}$ ; 2. $\tilde{p} = T_{SP_{left}} J_{s_{rc}} T_{SP_{right}}^T$ .

Таблица 5. Алгоритмы, реализующие двумерный метод трансформации изображений



Рис. 7. Трансформация скетчей в фотографии с помощью двумерного алгоритма. Ряд 1 – исходные скетчи, ряд 2 – результат трансформации. Преобразование выполнялось для скетчей из обучающей (а) и тестовой (б) выборок, а также для синтезированных скетчей [13] (в)



Рис. 8. Трансформация скетчей из базы AR в фотографии с помощью двумерного алгоритма.  
Ряд 1 – исходные скетчи, ряд 2 – результат трансформации

На рис. 7 показаны результаты трансформации скетчей из различных наборов изображений в фотографии. В отличие от одномерного варианта, данный алгоритм не позволяет получать безошибочную реконструкцию на обучающем множестве, но при этом качество трансформации на тестовых изображениях не ухудшается. Даже при преобразовании синтезированных скетчей (таких скетчей не было в обучающей выборке) результаты остаются хорошими. Кроме того, так как данный алгоритм строит модель перехода не по целым изображениям, а по строкам и столбцам, он менее чувствителен к сдвигам и изменениям масштаба, чем одномерный вариант.

На рис. 8 представлены результаты трансформации, выполненной на базе скетчей AR [9]. Модель была построена по парам «фотография–скетч» из базы «СUHK student data set» [9], а трансформированные скетчи не входили в обучающую выборку. Примечателен тот факт, что на всех реконструированных фотографиях отсутствуют очки. Это является следствием того, что ни одно изображение из обучающей выборки не содержит очков, поэтому очки не отобразились в собственном базисе. Во многих областях применения такой эффект может быть полезен. В противном случае он может быть устранен путем дополнения обучающей выборки изображениями лиц с очками.

Таким образом, двумерный алгоритм хорошо подходит для трансформации новых изображений, даже если в обучающей базе не было подобных фотографий или скетчей.

### Заключение

В работе рассмотрена задача взаимной трансформации связанных пар изображений и генерации их популяций. Подробно рассмотрен базовый метод, выявлены его недостатки. На основании анализа этих недостатков предложены три алгоритма, каждый из которых имеет свои слабые и сильные стороны, которые определяют его область применения.

1. *Алгоритм на основе матриц трансформации.* Отличается тем, что не использует на этапе трансформации обучающую выборку. Хорошо работает на обучающей выборке, содержащей произвольные изображения, даже в условиях низкого разрешения и шумов. Для обработки новых изображений важно использовать объемную базу обучения и обеспечить отсутствие значительных сдвигов и изменений масштаба. Кроме того, в последнем случае следует учитывать предположение о структурном подобии изображений.
2. *Алгоритм генерации популяций.* Может быть применен для улучшения репрезентативности существующих баз данных и повышения надежности распознавания новых изображений, не входящих в базу.
3. *Алгоритм трансформации на основе 2DPCA/2DKLT.* Хорошо подходит для преобразования изображений, не входящих в обучающую выборку. При этом работает даже в условиях наличия сдвигов и изменений масштаба.

Все описанные алгоритмы рассмотрены на примере задачи взаимного преобразования и сравнения изображений лиц, так как этот пример широко представлен на практике и позволяет использовать существующие базы фотографий лиц и скетчей.

Отметим, что при реконструкции новых фотографий из скетчей сохраняются все антропометрические параметры лиц. Исходя из этого, можно далее преобразовать эти фотографии в трехмерную форму, что практически невозможно выполнить непосредственно из скетчей.

За счет унифицированного математического описания каждый из алгоритмов может быть применен для произвольных пар связанных изображений. При этом важно учитывать ограничения каждого из алгоритмов и те условия, при которых его применение корректно. Таким образом, область применения

предложенных алгоритмов не ограничивается HFR, а включает в себя множество приложений, таких как сверхразрешение (Super-Resolution Imaging), кодирование изображений, поиск в условиях низкокачественного изображения-запроса и другие задачи CMMR, связанные с обработкой изображений различной физической природы.

Дальнейшее развитие алгоритмов взаимной трансформации изображений в связанных парах возможно за счет применения двумерных методов канонического корреляционного анализа (2D CCA) и проекций на латентные структуры (2D PLS).

## Литература

1. Kato T. Database architecture for content-based image retrieval // *Proceedings of SPIE*. 1992. V. 1662. P. 112–123.
2. Wang K., Yin Q., Wang W., Wu S., Wang L. A comprehensive survey on cross-modal retrieval // *ArXiv Preprint*. ArXiv160706215. 2016. 20 p.
3. Liao S., Yi D., Lei Z., Qin R., Li S.Z. Heterogeneous face recognition from local structures of normalized appearance // *Lecture Notes in Computer Science*. 2009. V. 5558. P. 209–218. doi: 10.1007/978-3-642-01793-3\_22
4. Guo G. Heterogeneous face recognition: an emerging topic in biometrics // *Intel Technology Journal*. 2014. V. 18. N 4. P. 80–97.
5. Kukharev G.A., Shchegoleva N.L., Kamenskaya E.I. Representation and comparison methods for semantically different images // *Pattern Recognition and Image Analysis*. 2014. V. 24. N 4. P. 518–529. doi: 10.1134/S1054661814040105
6. Tang X., Wang X. Face sketch recognition // *IEEE Transactions on Circuits and Systems for Video Technology*. 2004. V. 14. N 1. P. 50–57. doi: 10.1109/TCSVT.2003.818353
7. Kukharev G., Oleinik A. Face photo-sketch transformation and population generation // *Lecture Notes in Computer Science*. 2016. V. 9972. P. 329–340. doi: 10.1007/978-3-319-46418-3\_29
8. Кухарев Г.А., Каменская Е.И., Матвеев Ю.Н., Щеголева Н.Л. Методы обработки и распознавания изображений лиц в задачах биометрии / под ред. М.В. Хитрова. СПб.: Политехника, 2013. 388 с.
9. CUHK Face Sketch Database [Электронный ресурс]. Режим доступа: <http://mmlab.ie.cuhk.edu.hk/facesketch.html>, свободный. Яз. англ. (дата обращения 16.04.2016).
10. CUHK Face Sketch FERET Database [Электронный ресурс]. Режим доступа: <http://mmlab.ie.cuhk.edu.hk/cufsf/>, свободный. Яз. англ. (дата обращения 16.04.2016).
11. Chen H., Liu Z., Rose C., Xu Y., Shum H.Y., Salesin D. Example-based composite sketching of human portraits // *Proc. 3<sup>rd</sup> Int. Symposium on Non-Photorealistic Animation and Rendering*. Annecy, France, 2004. P. 95–102.
12. Иванова А. Собака инстинктивно подражает человеку. Собаки, похожие на хозяев. Фото [Электронный ресурс]. 2010. URL: [http://zdravkom.ru/nauka\\_i\\_obrazovanie/sobaki\\_kopiruyt\\_hozyaev](http://zdravkom.ru/nauka_i_obrazovanie/sobaki_kopiruyt_hozyaev) (дата доступа: 20.11.2016).
13. Щеголева Н.Л., Кухарев Г.А. Способ автоматической генерации скетчей и система для его осуществления. Патент РФ № 2541132. Бюл. 2015. №4.
14. Zhang Y., Ellyson S., Zone A., Gangam P., Sullins J., McCullough C., Canavan S., Yin L. Recognizing face sketches by a large number of human subjects: a perception-based study for facial distinctiveness // *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition and Workshops (FG 2011)*. Santa Barbara, USA, 2011. P. 707–712. doi: 10.1109/FG.2011.5771335
15. Kukharev G.A., Matveev Y.N., Shchegoleva N.L. New solutions for face photo retrieval based on sketches // *Pattern Recognition and Image Analysis*. 2016. V. 26. N 1. P. 165–175. doi: 10.1134/S1054661816010144
16. Кухарев Г.А., Матвеев Ю.Н., Форчманьски П. Поиск людей по фотороботам: методы, системы и практические решения // *Научно-технический вестник информационных технологий, механики и оптики*. 2015. Т. 15. № 4. С. 640–653. doi: 10.17586/2226-1494-2015-15-4-640-653
17. Tsapatsoulis N., Alexopoulos V., Kollias S. A vector based approximation of KLT and its application to face recognition //

## References

1. Kato T. Database architecture for content-based image retrieval. *Proceedings of SPIE*, 1992, vol. 1662, pp. 112–123.
2. Wang K., Yin Q., Wang W., Wu S., Wang L. A comprehensive survey on cross-modal retrieval. *ArXiv Preprint*, 2016, ArXiv160706215, 20 p.
3. Liao S., Yi D., Lei Z., Qin R., Li S.Z. Heterogeneous face recognition from local structures of normalized appearance. *Lecture Notes in Computer Science*, 2009, vol. 5558, pp. 209–218. doi: 10.1007/978-3-642-01793-3\_22
4. Guo G. Heterogeneous face recognition: an emerging topic in biometrics. *Intel Technology Journal*, 2014, vol. 18, no. 4, pp. 80–97.
5. Kukharev G.A., Shchegoleva N.L., Kamenskaya E.I. Representation and comparison methods for semantically different images. *Pattern Recognition and Image Analysis*, 2014, vol. 24, no. 4, pp. 518–529. doi: 10.1134/S1054661814040105
6. Tang X., Wang X. Face sketch recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, vol. 14, no. 1, pp. 50–57. doi: 10.1109/TCSVT.2003.818353
7. Kukharev G., Oleinik A. Face photo-sketch transformation and population generation. *Lecture Notes in Computer Science*, 2016, vol. 9972, pp. 329–340. doi: 10.1007/978-3-319-46418-3\_29
8. Kukharev G.A., Kamenskaya E.I., Matveev Y.N., Shchegoleva N.L. *Metody Obrabotki i Raspoznavaniya Izobrazhenii Lis v Zadachakh Biometrii* [Methods for Face Image Processing and Recognition in Biometric Applications] Ed. M.V. Khitrov. St. Petersburg, Politekhnik Publ., 2013, 388 p.
9. *CUHK Face Sketch Database*. Available at: <http://mmlab.ie.cuhk.edu.hk/facesketch.html> (accessed 11.08.2014).
10. *CUHK Face Sketch FERET Database*. Available at: <http://mmlab.ie.cuhk.edu.hk/cufsf/> (accessed 11.08.2014).
11. Chen H., Liu Z., Rose C., Xu Y., Shum H.Y., Salesin D. Example-based composite sketching of human portraits. *Proc. 3<sup>rd</sup> Int. Symposium on Non-Photorealistic Animation and Rendering*. Annecy, France, 2004, pp. 95–102.
12. Ivanova A. *Dog instinctively imitates humans. Dogs, like owners*. Available at: [http://zdravkom.ru/nauka\\_i\\_obrazovanie/sobaki\\_kopiruyt\\_hozyaev](http://zdravkom.ru/nauka_i_obrazovanie/sobaki_kopiruyt_hozyaev) (accessed 20.11.2016).
13. Shchegoleva N.L., Kukharev G.A. *Automatic Sketch Generation Method and System Therefor*. Patent RU2541132, 2015.
14. Zhang Y., Ellyson S., Zone A., Gangam P., Sullins J., McCullough C., Canavan S., Yin L. Recognizing face sketches by a large number of human subjects: a perception-based study for facial distinctiveness. *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition and Workshops, FG 2011*. Santa Barbara, USA, 2011, pp. 707–712. doi: 10.1109/FG.2011.5771335
15. Kukharev G.A., Matveev Y.N., Shchegoleva N.L. New solutions for face photo retrieval based on sketches. *Pattern Recognition and Image Analysis*, 2016, vol. 26, no. 1, pp. 165–175. doi: 10.1134/S1054661816010144
16. Kukharev G.A., Matveev Yu.N., Forczmanski P. People retrieval by means of composite pictures – methods, systems and practical decisions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol.15, no. 4, pp. 640–653. (In Russian). doi: 10.17586/2226-1494-2015-15-4-640-653
17. Tsapatsoulis N., Alexopoulos V., Kollias S. A vector based

- Proc. 9<sup>th</sup> European Signal Processing Conference EUSIPCO-98, Island of Rhodes, Greece, 1998. V. 1581. P. 1–4.
18. Shchegoleva N.L., Kukharev G.A. Application of two-dimensional principal component analysis for recognition of face images // *Pattern Recognition and Image Analysis*, 2010. V. 20. N 4. P. 513–527. doi: 10.1134/S1054661810040127
  19. Kukharev G., Forczmanski P. Data dimensionality reduction for face recognition // *Machine Graphics and Vision*, 2004. V. 13. N 1-2. P. 99–121.
  20. Kukharev G., Tujaka A., Forczmanski P. Face recognition using two-dimensional CCA and PLS // *International Journal of Biometrics*, 2011. V. 3. N 4. P. 300–321. doi: 10.1504/IJBM.2011.042814
  - approximation of KLT and its application to face recognition. *Proc. 9<sup>th</sup> European Signal Processing Conference EUSIPCO-98*. Island of Rhodes, Greece, 1998, vol. 1581, pp. 1–4.
  18. Shchegoleva N.L., Kukharev G.A. Application of two-dimensional principal component analysis for recognition of face images. *Pattern Recognition and Image Analysis*, 2010, vol. 20, no. 4, pp. 513–527. doi: 10.1134/S1054661810040127
  19. Kukharev G., Forczmanski P. Data dimensionality reduction for face recognition. *Machine Graphics and Vision*, 2004, vol. 13, no. 1-2, pp. 99–121.
  20. Kukharev G., Tujaka A., Forczmanski P. Face recognition using two-dimensional CCA and PLS. *International Journal of Biometrics*, 2011, vol. 3, no. 4, pp. 300–321. doi: 10.1504/IJBM.2011.042814

### Авторы

**Кухарев Георгий Александрович** – доктор технических наук, профессор, профессор, Западнпоморский технологический университет в Щецине, Щецин, 70-310, Польша, gkukharev@wi.zut.edu.pl

**Матвеев Юрий Николаевич** – доктор технических наук, главный научный сотрудник, ООО «ЦРТ-инновации», Санкт-Петербург, 196084, Российская Федерация; заведующий кафедрой, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; matveev@speechpro.com

**Олейник Андрей Леонидович** – аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, andrey\_oleynik@niuitmo.ru

### Authors

**Georgy A. Kukharev** – D.Sc., Full Professor, West Pomeranian University of Technology Szczecin, Szczecin, 70-310, Poland, gkukharev@wi.zut.edu.pl

**Yuri N. Matveev** – D.Sc., Chief Scientific Officer, “STC-Innovation”, Saint Petersburg, 196084, Russian Federation; Head of the Chair, ITMO University, Saint Petersburg, 197101, Russian Federation, matveev@speechpro.com

**Andrei L. Oleinik** – postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation, andrey\_oleynik@niuitmo.ru