

А.Л. ОЛЕЙНИК, Г.А. КУХАРЕВ
**АЛГОРИТМЫ ВЗАИМНОЙ РЕКОНСТРУКЦИИ
ИЗОБРАЖЕНИЙ ЛИЦ НА ОСНОВЕ МЕТОДОВ ПРОЕКЦИИ В
СОБСТВЕННЫЕ ПОДПРОСТРАНСТВА**

Олейник А.Л., Кухарев Г.А. Алгоритмы взаимной реконструкции изображений лиц на основе методов проекции в собственные подпространства.

Аннотация. Обсуждается проблема взаимной реконструкции изображений лиц в соответствующих друг другу парах. Эта проблема была поставлена в предыдущей статье авторов, а предложенные в ней решения обсуждались с приложением к задачам гетерогенного распознавания изображений лиц (Heterogeneous Face Recognition) и кросс-модального мультимедийного поиска (Cross-Modal Multimedia Retrieval). Эти решения основаны на одномерных и двумерных методах анализа главных компонент для двух исходных наборов изображений лиц и проекции их в независимые собственные подпространства, вычислении матриц взаимной трансформации между этими подпространствами и взаимной реконструкции изображений лиц на основе одномерного и двумерного преобразований Карунена — Лоэва.

В настоящей статье предлагаются новые подходы и решения, основанные исключительно на двумерных методах проекции в собственные подпространства и двух вариантах моделей регрессии — множественной линейной регрессии и регрессии частичных наименьших квадратов.

Приведены результаты экспериментов по взаимной реконструкции изображений лиц в парах скетч/фотография, в парах лиц с возрастными изменениями, а также в парах изображений лиц в формах 2D/3D. Для проведения экспериментов выбраны два варианта реализации предложенного подхода. Первый из них основан на двумерном анализе главных компонент и регрессии частичных наименьших квадратов, второй — на двумерном методе частичных наименьших квадратов и множественной линейной регрессии. Оба варианта показали приемлемые для практики результаты при решении задачи взаимной реконструкции изображений лиц. Кроме того, рассмотрен способ повышения качества реконструируемых изображений лиц при работе со смешанными базами. Он заключается в классификации на основе двумерного линейного дискриминантного анализа и построении регрессионной зависимости в рамках каждого класса.

Показано также, что в общем случае взаимная реконструкция изображений лиц достижима и в условиях, когда исходные изображения не входили в состав обучающих наборов изображений лиц.

Ключевые слова: изображение лица, скетч, фоторобот, взаимная реконструкция изображений лиц мультисенсорной природы, кросс-модальный мультимедийный поиск, метод главных компонент, метод частичных наименьших квадратов, двумерные проекционные методы, регрессия.

1. Введение. Одной из проблем поиска и распознавания изображений лиц (ИЛ) является то, что на входе соответствующих систем могут использоваться различные датчики ИЛ, различные методы предобработки ИЛ и различные формы их представления. В современной практике ИЛ могут быть представлены различными способами — в форме 2D изображений в видимом (VIS), тепловом (NIR) и инфракрасном свете (IR), в виде композиционных фотороботов, рисованных

скетчей и популяций из них; в форме карт глубины («range image», форма 2.5D), в форме контурных моделей области лица (Active Shape Model — ASM), моделей внешнего вида (Active Appearance Model — AAM), определяющих текстуру ИЛ, и, наконец, в форме моделей 3D.

При этом исходные данные на входе системы могут быть одновременно представлены несколькими наборами ИЛ, связанными «в пары», «тройки» или даже «группы из более 3 ИЛ». Примером последнего варианта является пять наборов ИЛ, представляющих лицо одного и того же человека, из которых первый набор содержит ИЛ(VIS), второй — ИЛ(NIR), а третий — ИЛ(IR), четвертый — ИЛ в форме скетчей и, наконец, пятый набор содержит ИЛ в форме 3D. В отечественной биометрии такие наборы относят к данным «мультисенсорной природы». В зарубежной литературе их относят к группе «гетерогенных данных».

С одной стороны, наличие разнообразных способов представления ИЛ существенно расширяет возможности и области применения систем поиска и распознавания людей. С другой стороны, такое разнообразие значительно усложняет структуру соответствующих систем распознавания ИЛ, алгоритмы их функционирования, интерпретацию результатов распознавания и саму реализацию таких систем.

Отметим, что в отечественной практике подобные системы поиска и распознавания ИЛ еще недостаточно исследованы и не представлены в научной литературе.

Исходя из этого, актуальной становится задача исследования подходов для построения методов поиска и распознавания ИЛ, предназначенных для одновременной обработки нескольких наборов исходных данных «мультисенсорной» природы. Примерами этих подходов могут быть широко обсуждаемые в последнее время в зарубежной технической литературе задачи «Heterogeneous Face Recognition and Matching», «Cross-Modal Face Matching», «Face Image Indexing and Retrieval», а также и более общие подходы для поиска информации, например «Cross-Modal Multimedia Retrieval».

Настоящая статья предлагает способы, которые призваны приблизить решение этих задач.

2. Обзор литературы. Класс задач, связанных с обработкой ИЛ различной физической природы, невероятно широк. По этой причине в научной литературе практически отсутствуют обзорные статьи, в полной мере раскрывающие эту тему. В некоторой степени она освещена в рамках класса задач, объединенных общим названием «Heterogeneous Face Recognition» (HFR). В него входит распознавание лиц по скетчам, ИЛ в форме 3D моделей, ИЛ в инфракрасном спектре и ИЛ в низком разрешении [1]. Основная особенность этих задач — наличие исходных

данных, состоящих из связанных пар, представленных двумя (или более) наборами ИЛ или другими числовыми матрицами.

С учетом того, что структура исходных данных — это два связанных между собой набора ИЛ, мы рассмотрим лишь решения, основанные на методах проекции, базовыми из которых являются *анализ главных компонент* (Principal Component Analysis, PCA), *линейный дискриминантный анализ* (Linear Discriminant Analysis, LDA), *канонический корреляционный анализ* (Canonical Correlation Analysis, CCA) и, наконец, *метод частичных наименьших квадратов* (Partial Least Squares, PLS). Их преимуществами являются универсальность, развитый математический аппарат, наличие реализующих эти методы численных алгоритмов. Это позволяет применять проекционные методы для решения широкого спектра задач обработки данных различной природы (которые иногда называют *мультисенсорными*).

Кроме того, в последние годы были предложены обобщения этих методов на двумерный случай [2, 3]: 2DPCA/2DKLT (Two-Dimensional Principal Component Analysis / Two-Dimensional Karhunen–Loève Transform), 2DLDA/2DKLT, 2DCCA/2DKLT, 2DPLS/2DKLT. При решении задач обработки изображений они обладают рядом преимуществ по сравнению с одномерными методами, основным из которых являются отсутствие *проблемы малой выборки* (Small Sample Size, SSS), существенное *сокращение вычислительных затрат* и возможность устойчивого решения задач на собственные значения в этих методах [3].

Ниже представим краткий обзор научной литературы по применению одномерных и двумерных методов проекции в собственные подпространства для решения задач обработки ИЛ различной физической природы.

Говоря о *сценариях и алгоритмах* обработки ИЛ различной физической природы, можно выделить два направления. Первое представляет конкретные решаемые задачи и соответствующие им сценарии (например, кросс-модальное распознавание и сверхразрешение). Второе направление охватывает конкретные способы представления исходных ИЛ — например, фотоизображения лиц и скетчи.

Проанализируем далее эти направления, выделив следующие задачи:

– *Индексирование и распознавание (matching)*. Индексирование предполагает взаимное сопоставление ИЛ, представленных разными способами, в рамках некоторого фиксированного набора, что позволяет эффективно решать задачи кросс-модального поиска. Распознавание включает сравнение и поиск произвольных ИЛ, представленных разными способами, что необходимо для идентификации

и верификации. Так, в работе [4] предлагается использовать метод ССА для объединения высокоуровневых (пол, возраст и др.) и низкоуровневых признаков (гистограммы ориентированных градиентов) в единое признаковое представление, пригодное для распознавания. В [5] авторы применяют метод PLS для проекции исходных ИЛ (фотографии в высоком и низком разрешении, скетчи) на промежуточные подпространства. Полученные проекции используются для решения задач индексирования и распознавания.

– *Трансформация (взаимная реконструкция)* [6, 7] предполагает переход между ИЛ, представленными различными способами. Так, например, из ИЛ(NIR) реконструируется ИЛ(VIS), которое может быть использовано для автоматического поиска в базе фотоизображений лиц. Такой подход позволяет использовать готовые (в том числе коммерческие) системы распознавания ИЛ, а реконструкцию ИЛ выделить в отдельную задачу. При этом результат реконструкции также может быть предъявлен человеку (например, свидетелю).

– *Сверхразрешение (Super Resolution, SR) и Face Hallucination (FH)* представляют собой два класса методов реконструкции изображений в высоком разрешении из изображений в низком разрешении. В рамках SR ставится задача получения одного изображения (не обязательно ИЛ) в высоком разрешении из набора изображений в низком разрешении. Методы SR чаще всего применяются в системах анализа видеопотока с камер низкого разрешения. С другой стороны, во многих случаях доступно только одно изображение в низком разрешении, а не их набор. В случае ИЛ для получения изображения в высоком разрешении может быть использована априорная информация о структуре лица, как это делается в рамках методов FH [8]. В работе [9] предложен подход на основе 2DCCA, выполняемом между ИЛ в низком и высоком разрешениях. Реализуемая далее реконструкция ИЛ производится посредством метода К ближайших соседей.

Перечисленные выше задачи могут решаться применительно к ИЛ, представленным разными способами. В большинстве публикаций это *фотоизображение лица в видимом свете* — ИЛ(VIS) — и один (или несколько) из следующих способов представления:

– *Скетч* («набросок», «эскиз»). Скетчи могут быть нарисованы в присутствии человека или по его фотографии (Viewed Sketch), выполнены на компьютере и доработаны художником (Artist Sketch), составлены из готовых фрагментов (рты, носы, глаза и др.) художником (Composite Sketch) или криминалистом (Composite Forensic Sketch) [10]. Один из первых алгоритмов взаимной реконструкции ИЛ(VIS) и скетчей основан на PCA [6]. Данный подход является расширением

метода собственных лиц (*Eigenfaces*) на случай двух наборов изображений: фотографий и скетчей. Недавно этот алгоритм был улучшен и обобщен в работах [7, 11], где предложено использовать его не только для трансформации ИЛ, но и для генерации их популяций. Кроме того, в [7] впервые представлен метод трансформации на основе двумерного анализа главных компонент.

– *ИЛ в ближнем инфракрасном (ИК) спектре (NIR) и в тепловом ИК спектре (IR)*. В работе [12] предложен подход к распознаванию лиц по ИЛ(IR) на основе метода PLS, а в [13] ССА применен для реконструкции ИЛ(NIR) из ИЛ(VIS).

– *ИЛ в форме карты глубины («depth map», «range image», 2.5D) и трёхмерной модели (3D) включают информацию о глубине (т.е. z-координату)*. В общем случае задача восстановления этой информации по 2D изображению достаточно сложна. Подходы к ее решению включают, например, методы «shape from shading» и «structure from motion». В случае ИЛ обычно используют априорную информацию о структуре лица, представленную в форме среднего «range image» или некоторой статистической модели. Так, в работах [13, 14] карты глубины реконструируются из RGB-изображения с помощью ССА и линейной регрессии, выполненных между парами RGB-изображений и «range image» из обучающей выборки.

Кроме отмеченного выше, список способов представления ИЛ можно дополнить *ИЛ в низком разрешении* (что соответствует рассмотренным выше задачам SR и FH), а также *ИЛ с возрастными изменениями* [15]. Как правило, моделирование и анализ возрастных изменений сводятся к изучению зависимости тех или иных антропометрических параметров от возраста. С другой стороны, эта задача может быть представлена как анализ связанных наборов ИЛ (в качестве которых, например, могут выступать пары ИЛ с разницей в заданное количество лет). В настоящей статье обсуждается подход к решению этой задачи на основе проекционных методов.

Подводя итог краткому обзору и суммируя представленные результаты, отметим, что вариантов представления ИЛ существует достаточно много. Однако, в большинстве статей рассматривается какая-то одна пара, например, ИЛ(VIS) и ИЛ(IR). Отметим также, что в случаях использования *смешанных баз ИЛ* возникают дополнительные трудности, связанные с их неоднородностью. Возможным решением может быть разбиение таких баз на группы, в рамках которых создаются «локальные» модели взаимосвязи ИЛ, сгруппированных, например, по признакам сенсорной принадлежности разными способами [16].

При этом следует отметить, что именно методы проекции в подпространства применяются наиболее часто для решения задач совместной обработки ИЛ, представленных совершенно различными способами и при условии полного неподобия в парах. Кроме того, эти методы используются и для решения других задач, таких как Cross-Modal Multimedia Retrieval [11].

Кроме того, в последнее время появился ряд других подходов к решению рассматриваемого класса задач обработки ИЛ. К ним относятся, например, глубокие нейронные сети (Deep Neural Network, DNN), Марковские случайные поля (Markov Random Field, MRF) и скрытые Марковские модели (Embedded Hidden Markov Model, E-HMM). В некоторых задачах они могут превосходить проекционные методы по ряду характеристик, например, по качеству реконструкции. Однако такие решения зачастую узкоспециализированны, в то время как проекционные методы отличаются универсальностью. Это позволяет использовать их в совершенно различных сценариях при неизменном математическом и алгоритмическом описании. В настоящее время уже существуют попытки применения методов проекции в собственные подпространства в составе слоев глубоких нейронных сетей [17]. Этим, например, достигается и уменьшение исходного пространства признаков, и использование в слоях сети только наиболее важной информации, представляющей исходные и промежуточные данные. Именно поэтому исследование различных способов применения методов проекции в собственные подпространства представляет особый интерес. И каждое решение, полученное здесь, обладает особой практической ценностью и актуальностью.

Уместно отметить, что на сегодняшний день известны лишь единичные публикации по данной теме [9, 11, 18], в то время как полномасштабные исследования в научной литературе на сегодняшний день не представлены. Данная статья нацелена на устранение этого пробела и представляет подход к взаимной реконструкции ИЛ на основе двумерных проекционных методов.

3. Предлагаемый подход. На рисунке 1 представлена общая схема, описывающая предлагаемый подход к взаимной реконструкции ИЛ, представленных двумя наборами данных и обозначенных как «ИЛ 1» и «ИЛ 2». В роли «ИЛ 1» и «ИЛ 2» могут выступать ИЛ (VIS) + скетчи, ИЛ (IR) + ИЛ (VIS) или другие комбинации различных способов представления ИЛ.

Предлагаемый подход реализуется в два этапа. На первом этапе (верхняя часть рисунка 1) выполняется анализ исходных наборов ИЛ, включающий вычисление матриц проекции, трансформацию исходных данных в собственное подпространство с использованием двумерного преобразования Карунена-Лоэва (Two-Dimensional Karhunen–Loève transform, 2DKLT) и вычисление параметров регрессии для взаимной реконструкции ИЛ в собственном подпространстве. На втором этапе (нижняя часть рисунка 1) выполняется взаимная реконструкция (на рисунке 1: $ИЛ1 \rightarrow \widetilde{ИЛ2}$ и $ИЛ2 \rightarrow \widetilde{ИЛ1}$), реализуемая с помощью прямого 2DKLT, реконструкции в подпространстве и перехода в исходное пространство признаков (т.е. пространство ИЛ) с помощью обратного 2DKLT.

Результатом анализа исходных наборов ИЛ являются *матрицы проекции* для каждого из наборов и *матрицы регрессии*. Вычисление матриц проекции выполняется с помощью двумерных проекционных методов: 2DPCA, 2DCCA и 2DPLS. В результате для каждого исходного набора ИЛ вычисляются по две матрицы проекции: для строк и для столбцов. Процедура 2DKLT сводится к умножению каждого из исходных изображений на матрицы проекции (слева и справа). На основе полученных таким образом двумерных проекций вычисляются две матрицы регрессии (по одной для реконструкции в каждом направлении). Это может быть сделано, например, с помощью множественной линейной регрессии (Multiple Linear Regression, MLR) или регрессии PLS [19].

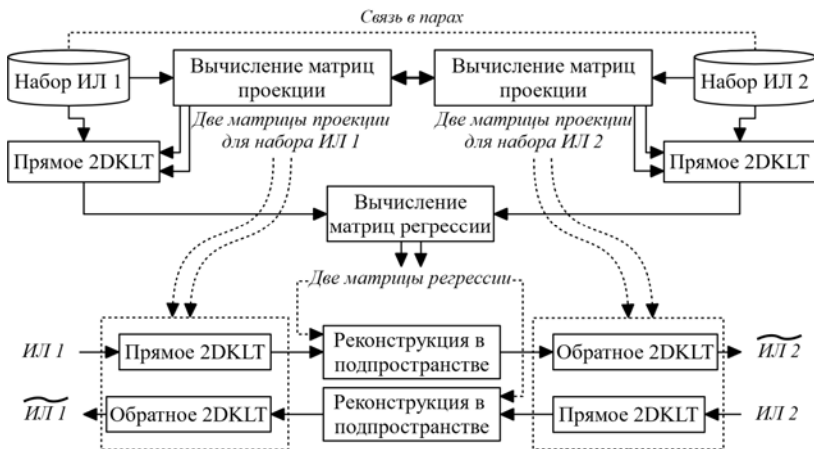


Рис. 1. Общая схема предлагаемого подхода к взаимной реконструкции ИЛ

Представленное выше решение является обобщением метода регрессии на главные компоненты (Principal Component Regression, PCR) [19], представляющего собой комбинацию PCA и MLR. Альтернативой PCR является регрессия PLS, сочетающая PLS с MLR. Использование двумерных проекционных методов позволяет эффективно решать задачи обработки, распознавания и преобразования изображений, так как в этом случае не требуется векторизация (т.е. конкатенация строк или столбцов изображений).

Отображение исходных наборов ИЛ в собственное подпространство с помощью двумерных проекционных методов преследует сразу несколько целей. Использование метода 2DCCA позволяет достичь высокой *взаимной* корреляции между проекциями исходных наборов ИЛ (в случае 2DPLS максимизируется ковариация). Более того, полученные в результате проекции признаки являются декоррелированными (*в рамках каждого из наборов ИЛ*), что обеспечивает устойчивость вычислений на этапе расчета матрицы регрессии. Кроме того, дополнительно может быть выполнена редукция размерности пространства признаков, что позволяет удалить из исходных данных шумовую составляющую и снизить вычислительные затраты.

Если ИЛ представлены тремя и более способами, предлагается группировать их в связанные пары и анализировать эти пары отдельно. В таком случае двумерный анализ главных компонент может оказаться предпочтительным, так как с его помощью достаточно вычислить матрицы проекции для каждого набора ИЛ, а не для каждой связанной пары наборов.

Двумерные проекционные методы. Обозначим два исходных набора ИЛ как $\{X_k\}$ и $\{Y_k\}$, $k=1\dots K$. Здесь X_k и Y_k — изображения размера $M \times N$. Заметим, что исходные данные, как правило, центрируют относительно средних изображений. Поэтому далее предполагается, что ИЛ отцентрированы.

Двумерные проекционные методы, подобно одномерным аналогам, позволяют выделять из исходных данных составляющие с наибольшей дисперсией, ковариационной или корреляционной связью. Отличительной особенностью двумерных методов является то, что при решении задач обработки изображений они обладают значительно большим быстродействием и вычислительной устойчивостью по сравнению с одномерными методами. Это достигается за счет замены одной задачи на собственные значения большой размерности MN двумя задачами меньшей размерности (соответственно, M и N). В таблице 1 представлено формальное описание этих методов.

Таблица 1. Двумерные проекционные методы

| Операция | 2DPCA/2DKLT | 2DPLS/2DKLT | 2DCCA/2DKLT |
|--|--|---|---|
| Вычисление матриц рассеяния по строкам | $C_{XX}^{(r)} = \sum_{k=1}^K X^{(k)} \left(X^{(k)} \right)^T, \quad C_{YY}^{(r)} = \sum_{k=1}^K Y^{(k)} \left(Y^{(k)} \right)^T$ | | |
| Вычисление матриц рассеяния по столбцам | $C_{XX}^{(c)} = \sum_{k=1}^K \left(X^{(k)} \right)^T X^{(k)}, \quad C_{YY}^{(c)} = \sum_{k=1}^K \left(Y^{(k)} \right)^T Y^{(k)}$ | | |
| Вычисление взаимных матриц рассеяния по строкам | — | $C_{XY}^{(r)} = \sum_{k=1}^K X^{(k)} \left(Y^{(k)} \right)^T$ $C_{YX}^{(r)} = \left(C_{XY}^{(r)} \right)^T$ | |
| Вычисление взаимных матриц рассеяния по столбцам | — | $C_{XY}^{(c)} = \sum_{k=1}^K \left(X^{(k)} \right)^T Y^{(k)}$ $C_{YX}^{(c)} = \left(C_{XY}^{(c)} \right)^T$ | |
| Вычисление полных матриц рассеяния | — | $S^{(tot1,r)} =$ $= C_{XY}^{(r)} C_{YX}^{(r)}$ $S^{(tot2,r)} =$ $= C_{YX}^{(r)} C_{XY}^{(r)}$ $S^{(tot1,c)} =$ $= C_{XY}^{(c)} C_{YX}^{(c)}$ $S^{(tot2,c)} =$ $= C_{YX}^{(c)} C_{XY}^{(c)}$ | $S^{(tot1,r)} =$ $= \left(C_{XX}^{(r)} \right)^{-1} C_{XY}^{(r)} \left(C_{YY}^{(r)} \right)^{-1} C_{YX}^{(r)}$ $S^{(tot2,r)} =$ $= \left(C_{YY}^{(r)} \right)^{-1} C_{YX}^{(r)} \left(C_{XX}^{(r)} \right)^{-1} C_{XY}^{(r)}$ $S^{(tot1,c)} =$ $= \left(C_{XX}^{(c)} \right)^{-1} C_{XY}^{(c)} \left(C_{YY}^{(c)} \right)^{-1} C_{YX}^{(c)}$ $S^{(tot2,c)} =$ $= \left(C_{YY}^{(c)} \right)^{-1} C_{YX}^{(c)} \left(C_{XX}^{(c)} \right)^{-1} C_{XY}^{(c)}$ |

Продолжение таблицы 1. Двумерные проекционные методы

| | | |
|--|--|---|
| <p>Решение задач на собственные значения</p> | $C_{XX}^{(r)} W_X^{(r)} = \Lambda_X^{(r)} W_X^{(r)}$ $C_{YY}^{(r)} W_Y^{(r)} = \Lambda_Y^{(r)} W_Y^{(r)}$ $C_{XX}^{(c)} W_X^{(c)} = \Lambda_X^{(c)} W_X^{(c)}$ $C_{YY}^{(c)} W_Y^{(c)} = \Lambda_Y^{(c)} W_Y^{(c)}$ | $S^{(tot1,r)} W_X^{(r)} = \Lambda_X^{(r)} W_X^{(r)}$ $S^{(tot2,r)} W_Y^{(r)} = \Lambda_Y^{(r)} W_Y^{(r)}$ $S^{(tot1,c)} W_X^{(c)} = \Lambda_X^{(c)} W_X^{(c)}$ $S^{(tot2,c)} W_Y^{(c)} = \Lambda_Y^{(c)} W_Y^{(c)}$ |
| | <p>Матрицы собственных векторов для строк $W_X^{(r)}$ и $W_Y^{(r)}$, а также матрицы собственных чисел для строк $\Lambda_X^{(r)}$ и $\Lambda_Y^{(r)}$ имеют размеры $M \times M$.</p> <p>Матрицы собственных векторов для столбцов $W_X^{(c)}$ и $W_Y^{(c)}$, а также матрицы собственных чисел для столбцов $\Lambda_X^{(c)}$ и $\Lambda_Y^{(c)}$ имеют размеры $N \times N$.</p> <p>Матрицы $W_X^{(r)}$, $W_Y^{(r)}$, $W_X^{(c)}$ и $W_Y^{(c)}$ в качестве столбцов содержат собственные векторы соответствующих матриц рассеяния.</p> <p>Матрицы $\Lambda_X^{(r)}$, $\Lambda_Y^{(r)}$, $\Lambda_X^{(c)}$ и $\Lambda_Y^{(c)}$ являются диагональными и содержат собственные числа, соответствующие столбцам матриц собственных векторов (в том же порядке).</p> | |
| <p>Выбор главных компонент (редукция размерности пространства признаков)</p> | <p>Выбор количества компонент по строкам и столбцам: $0 < d_r \leq M$, $0 < d_c \leq N$.</p> <p>Матрицы проекции $F_X^{(r)}$ и $F_Y^{(r)}$ (размера $M \times d_r$) составляются из первых d_r столбцов матриц $W_X^{(r)}$ и $W_Y^{(r)}$ (то есть из собственных векторов), упорядоченных по убыванию соответствующих им собственных чисел.</p> <p>Матрицы проекции $F_X^{(c)}$ и $F_Y^{(c)}$ (размера $N \times d_c$) составляются из первых d_c столбцов матриц $W_X^{(c)}$ и $W_Y^{(c)}$ (то есть из собственных векторов), упорядоченных по убыванию соответствующих им собственных чисел.</p> | |
| <p>2DKLT</p> | $U^{(k)} = \left(F_X^{(r)} \right)^T X^{(k)} F_X^{(c)}$ $V^{(k)} = \left(F_Y^{(r)} \right)^T Y^{(k)} F_Y^{(c)}$ <p>Двумерные проекции исходных данных $U^{(k)}$ и $V^{(k)}$ представляют собой матрицы размера $d_r \times d_c$.</p> | |

Полное описание двумерных проекционных методов, способов их реализации и анализ характеристик можно найти в монографии [3].

Отметим, что проекции исходных данных U_k и V_k , полученные методом 2DPCA/2DKLT, зависят только от «своих» наборов данных (соответственно, X_k и Y_k). Иными словами, 2DPCA выполняется два раза: для X_k и Y_k . В случае же 2DPLS/2DKLT и 2DCCA/2DKLT как U_k , так и V_k зависит от обоих наборов данных, то есть 2DPLS и 2DCCA выполняются один раз для двух наборов данных. Таким образом, если предполагается обработка большого количества связанных наборов данных, то метод 2DPCA/2DKLT позволит проводить их анализ независимо, без перебора всевозможных пар этих наборов данных. С другой стороны, методы 2DPLS/2DKLT и 2DCCA/2DKLT учитывают взаимосвязь между исходными данными, что снижает ошибку реконструкции при построении регрессионной зависимости и позволяет использовать меньше компонент, снижая таким образом вычислительные затраты.

Заметим также, что метод 2DCCA/2DKLT предполагает обращение матриц ковариации, сформированных по исходным данным. Поскольку такими данными являются изображения, то во многих случаях матрицы ковариации могут быть сингулярными, что приводит либо к невозможности их обращения, либо к неустойчивости дальнейших вычислений. Для того чтобы обойти эту проблему, в 2DCCA применяют методы регуляризации как матриц ковариации, так и общих матриц рассеяния. Подробное описание данного решения можно найти в [3].

Регрессионные модели между двумерными проекциями изображений. Анализ исходных наборов ИЛ завершается вычислением матриц регрессии для проекций, полученных с помощью 2DKLT. Существуют различные методы построения регрессионной зависимости. В настоящей статье представлены варианты базового подхода, основанные на множественной линейной регрессии (MLR) и регрессии PLS.

Метод MLR достаточно прост и эффективен. Если размерность пространства признаков после редукции $D = d_r \times d_c$ достаточно мала, то метод MLR не требует значительных вычислительных затрат, а полученное решение обладает вычислительной устойчивостью. Как правило, это может быть достигнуто при использовании 2DPLS/2DKLT и 2DCCA/2DKLT на этапе вычисления матриц проекции.

Как было отмечено выше, метод 2DPCA/2DKLT чаще всего предполагает использование большего количества компонент D , чем

2DPLS/2DKLT или 2DCCA/2DKLT. По этой причине метод MLR может оказаться неприменимым. В таком случае возникает необходимость в дальнейшем снижении размерности пространства признаков, что может быть достигнуто с помощью регрессии PLS. Пример совместного использования 2DPCA/2DKLT и регрессии PLS представлен ниже, в разделе экспериментальных исследований.

В настоящей статье использована реализация регрессии PLS, основанная на алгоритме NIPALS (Non-linear Iterative Partial Least Squares). По сути, это вариант степенного метода, представляющего собой итеративную процедуру решения задачи на собственные значения.

4. Экспериментальные исследования. Экспериментальные исследования проведены на трех базах ИЛ:

1. База фотографий лиц и соответствующих им скетчей, выполненных художниками: CUFS/CUFSF [20, 21];
2. ИЛ для заданной группы людей в различном возрасте [22];
3. Фотографии лиц и соответствующие им карты глубины («range image»): база Texas 3D [23-25].

Реконструкция фотоизображений лиц по скетчам на основе 2DPCA/2DKLT и PLS-регрессии. Первый вариант представленного подхода основан на методах 2DPCA/2DKLT и PLS-регрессии; его схема приведена на рисунке 2.

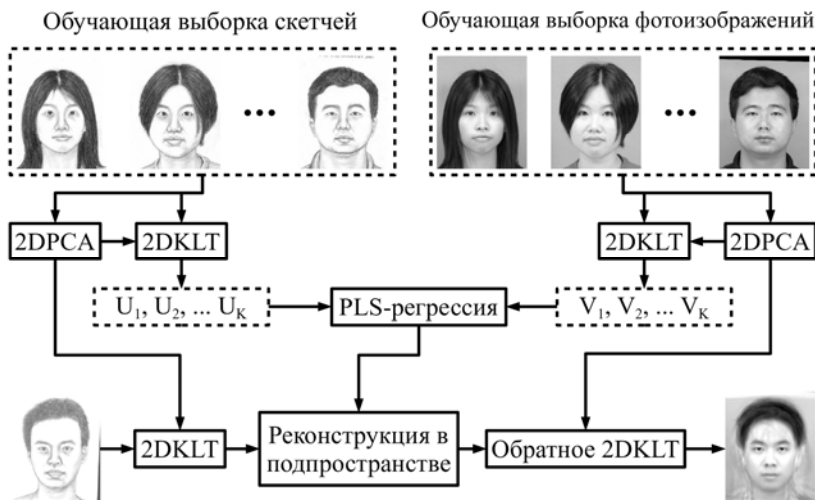


Рис. 2. Реконструкция фотоизображений по скетчам на основе метода 2DPCA/2DKLT и PLS-регрессии

В рамках экспериментов на базах CUFS/CUFSF были использованы следующие параметры:

1. Размеры ИЛ: $M = 250$, $N = 200$. змерность исходного пространства признаков $M \times N = 50000$;

2. Количество элементов по строкам и столбцам: $d_r = 60$, $d_c = 50$. Размерность промежуточного пространства признаков $D = d_r \times d_c = 3000$;

3. Количество компонент в регрессии PLS: 17.

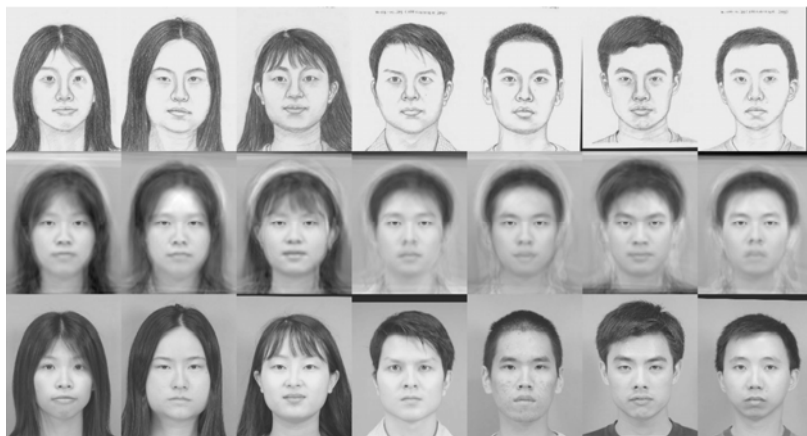
Таким образом, размерность пространства признаков поэтапно снижается с 50000 до 3000 и далее до 17. Попытка применения PLS-регрессии без использования 2DPCA/2DKLT приводит к неоправданно высоким вычислительным затратам и неустойчивости решения регрессионной задачи. Кроме того, при добавлении нового набора ИЛ для тех же людей (например, IR), собственные базисы, построенные с помощью 2DPCA/2DKLT, могут быть использованы повторно.

На рисунке 3 показаны примеры реконструкции фотографий по соответствующим скетчам на обучающей и тестовой выборках. Можно отметить приемлемое качество реконструкции как для обучающей, так и для тестовой выборки. Ухудшение качества реконструкции на тестовой выборке можно объяснить уникальными особенностями каждого скетча, зачастую не подобного исходной фотографии.

На рисунке 4 представлены результаты реконструкции фотоизображений по скетчам в низком разрешении с добавлением шума. Для этого размер скетчей был уменьшен до 9% от исходного (т.е. более чем в десять раз) и к ним был добавлен равномерный шум со значениями из диапазона $\pm 5\%$ от максимальной яркости изображения. Результаты, представленные на рисунке 4, свидетельствуют об устойчивости предложенного подхода к шумам и понижению разрешения исходного изображения. Здесь важно то, что обучение проводилось только на ИЛ в высоком разрешении. Таким образом, можно говорить о применимости предложенного решения к задачам сверхразрешения (SR) и Face Hallucination (FH).

Моделирование возрастных изменений с помощью 2DPCA/2DKLT и PLS-регрессии. Набор ИЛ людей различных возрастов включает ИЛ в возрасте 20, 30, 40, 50, 60 лет. Исходные данные

выбраны из статьи [22]. Предлагаемый нами подход применен для «состаривания» ИЛ человека на 10 лет.



a)



б)

Рис. 3. Реконструкция фотографий лиц по скетчам, выполненная на тестовой (б) выборке. В первом ряду приведены исходные скетчи, во втором — реконструированные из них фотографии, в третьем — исходные фотографии



Рис. 4. Реконструкция фотографий лиц по зашумлённым скетчам в низком разрешении, выполненная на обучающей (а) и тестовой (б) выборках. В первом ряду приведены исходные скетчи, во втором — реконструированные из них фотографии, в третьем — исходные фотографии

Для этого для каждого человека соответствующие ИЛ были сгруппированы в пары по (30, 40), (40, 50), (50, 60) лет (как показано на рисунке 5) и использованы в качестве обучающей выборки. ИЛ в возрасте 20 лет играют роль тестовой выборки. Полученные таким образом наборы данных могут быть обработаны в соответствии с описанным выше подходом.

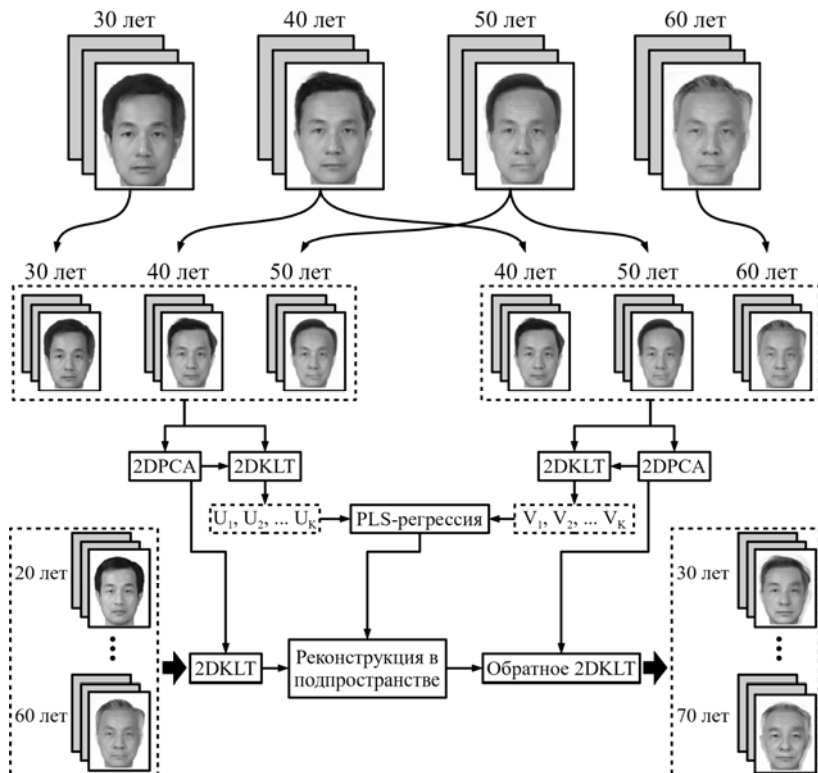


Рис. 5. Сведение задачи моделирования возрастных изменений к задаче реконструкции ИЛ в связанных парах

В рамках этих экспериментов были использованы следующие параметры:

1. Размеры ИЛ: $M = 338$, $N = 264$. Размерность исходного пространства признаков $M \times N = 89232$;
2. Количество элементов по строкам и столбцам: $d_r = 60$, $d_c = 50$. Размерность промежуточного пространства признаков $D = d_r \times d_c = 3000$;
3. Количество компонент в регрессии PLS: 8.

На рисунке 6 показаны примеры реконструкции ИЛ старшего возраста из ИЛ младшего возраста. Несмотря на то что качество реконструкции на тестовой выборке несколько ниже, чем на обучающей, можно отметить характерные возрастные изменения, вносимые в реконструированные изображения. Для повышения качества результирующих ИЛ необходима более объемная и репрезентативная выборка ИЛ.

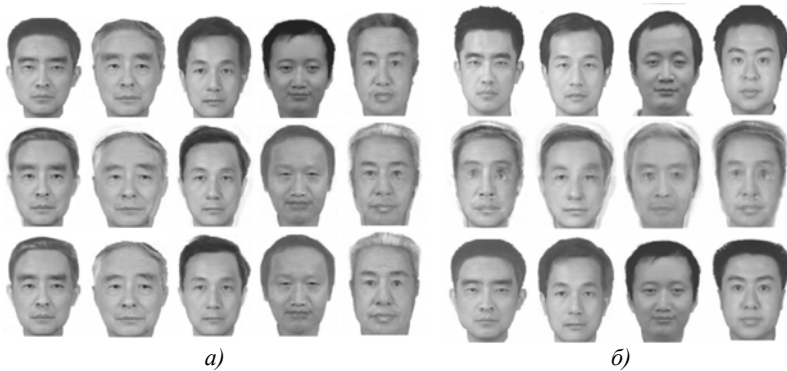


Рис. 6. Реконструкция ИЛ старшего возраста по ИЛ младшего возраста, выполненная на обучающей (а) и тестовой (б) выборках. В первом ряду приведены ИЛ младшего возраста, во втором — реконструированные из них ИЛ старшего возраста, в третьем — исходные ИЛ старшего возраста

Реконструкция фотоизображений лиц по скетчам на основе 2DPLS/2DKLT и MLR. Здесь представлены эксперименты по взаимной реконструкции фотоизображений лиц и скетчей, а также фотоизображений лиц и карт глубины («range image») с помощью комбинации методов 2DPLS/2DKLT и MLR.

Реконструкция выполнялась в двух режимах:

- «в классе» — регрессия по столбцам двумерных проекций выполняется отдельно для каждой пары ИЛ (рисунок 7);
- «по базе» — регрессия по столбцам двумерных проекций ИЛ выполняется для всей обучающей выборки (рисунок 8).

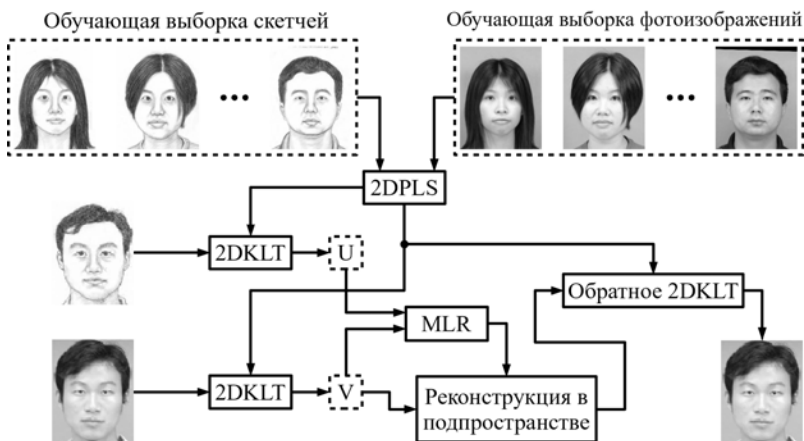


Рис. 7. Реконструкция фотоизображений лиц по скетчам «в классе»

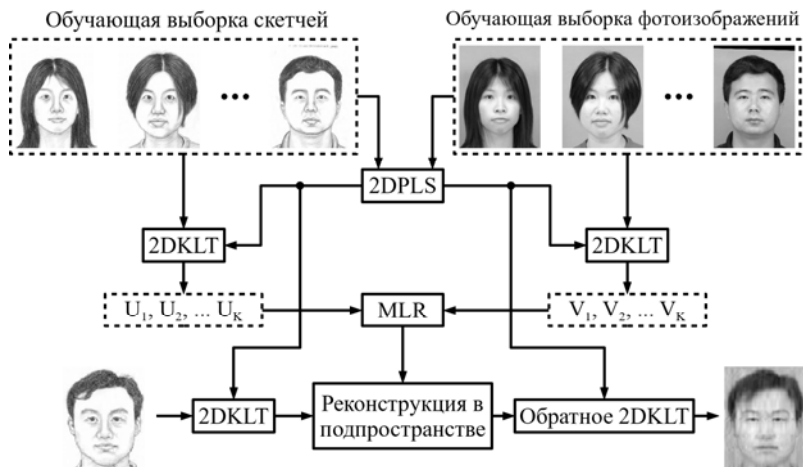


Рис. 8. Реконструкция фотоизображений лиц по скетчам «по базе»

На рисунке 9 представлены примеры изображений исходной базы CUFS, полученные с помощью 2DPLS наборы собственных чисел, а также фазовая корреляция в собственном подпространстве между переменными U и V , полученными с помощью 2DKLT. Достаточно высокое значение, достигаемое в области пика фазовой корреляции, говорит об относительно высоком подобии исходных данных в собственном подпространстве признаков.

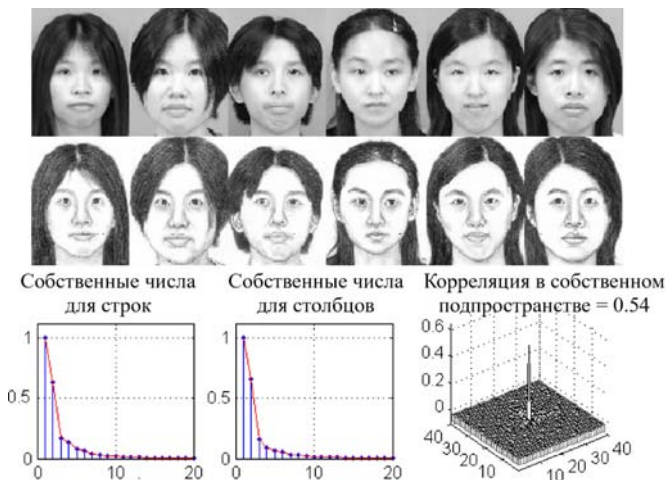


Рис. 9. Примеры обучающей выборки (фото и скетчи) из базы CUFS и результаты их обработки в рамках метода 2DPLS/2DKLT: собственные числа и фазовая корреляция в подпространстве

На рисунке 10 представлены результаты реконструкции тестовых ИЛ «в классе» и «по базе». Отметим, что реконструкция «в классе» всегда выполняется точнее, чем «по базе», и ограничена лишь точностью реконструкции на этапе обратного 2DKLT.

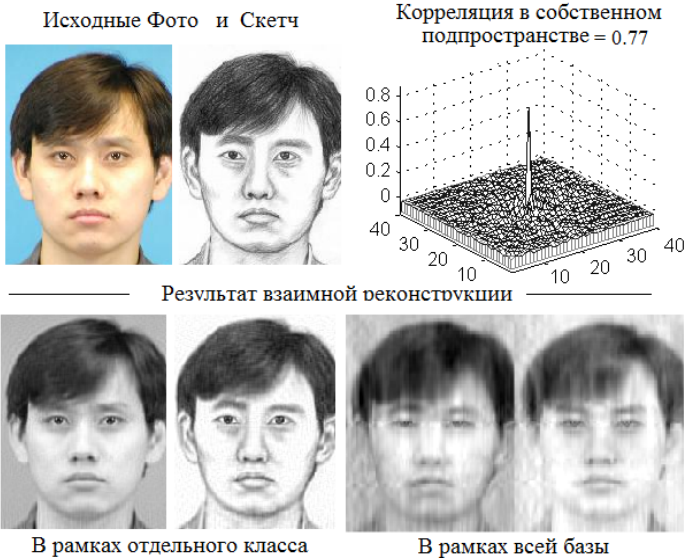


Рис. 10. Примеры реконструкции ИЛ «в классе» (слева) и «по базе» (справа)

На рисунке 11 представлены результаты взаимной реконструкции фотографий и карт глубины («range image»), выполненной на базе Texas 3D.

Эксперимент показывает, что предложенный метод позволяет решать задачу взаимной реконструкции ИЛ различной физической природы, что позволяет сделать вывод о его универсальности.

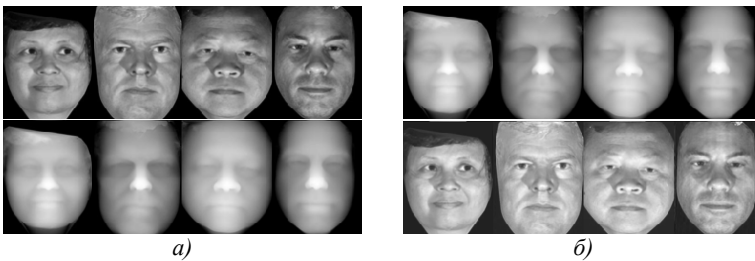


Рис. 11. Реконструкция ИЛ «в классе» для базы Texas 3D, содержащей фотографии лиц и соответствующие им карты глубины («range image»). Первый ряд — исходные ИЛ, второй ряд — результат реконструкции

На рисунке 12 показана взаимная реконструкция фотографий и скетчей, входящих в обучающую и тестовую базы.

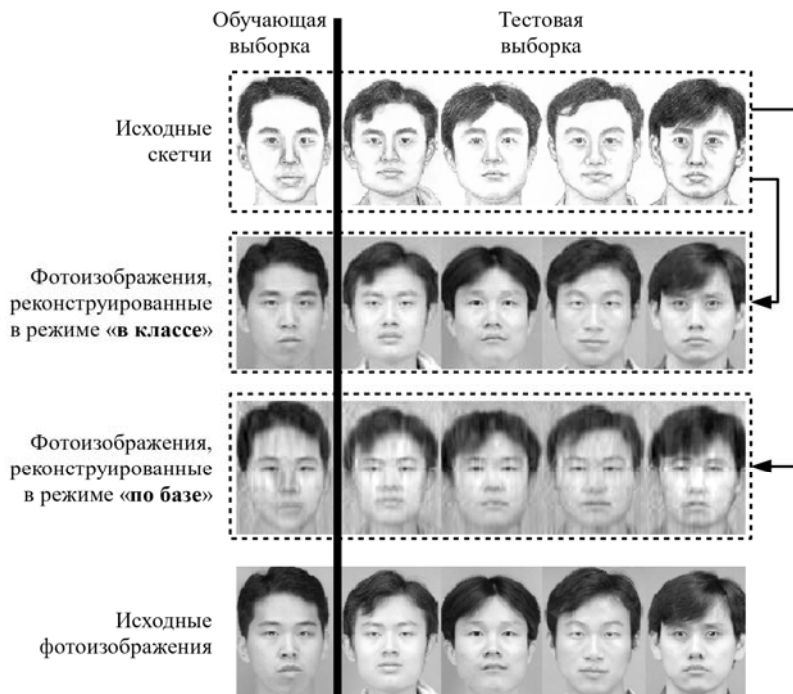


Рис. 12. Реконструкция ИЛ «в классе» и «по базе», выполненная на обучающей и тестовой выборках

Из представленных выше экспериментов видно, что режим реконструкции «в классе» позволяет получить наилучшие результаты. Это достигается за счет того, что в роли «класса» (в рамках которого строится модель и выполняется реконструкция) выступает единственная пара ИЛ.

Для практического использования можно предложить следующий подход. Исходные наборы ИЛ разбиваются на классы по некоторому признаку (фенотипические признаки, пол и др.), после чего для каждого из классов строится отдельная регрессионная модель. На этапе реконструкции сначала выполняется классификация входного ИЛ, а затем выполняется собственно реконструкция с использованием соответствующей матрицы регрессии. Для классификации ИЛ может быть использован метод 2DLDA/2DKLT [3] и некоторый алгоритм классификации (например, алгоритм ближайшего соседа). На рисунке 13 представлен один из вариантов такого подхода.

Заметим, что для реализации такого подхода необходима выборка, снабженная метками классов. Если метки классов отсутствуют, можно применить некоторый алгоритм кластеризации. Такие подходы существуют и представлены в научной литературе [16]. Их применение к решаемой задаче является предметом дальнейших исследований.

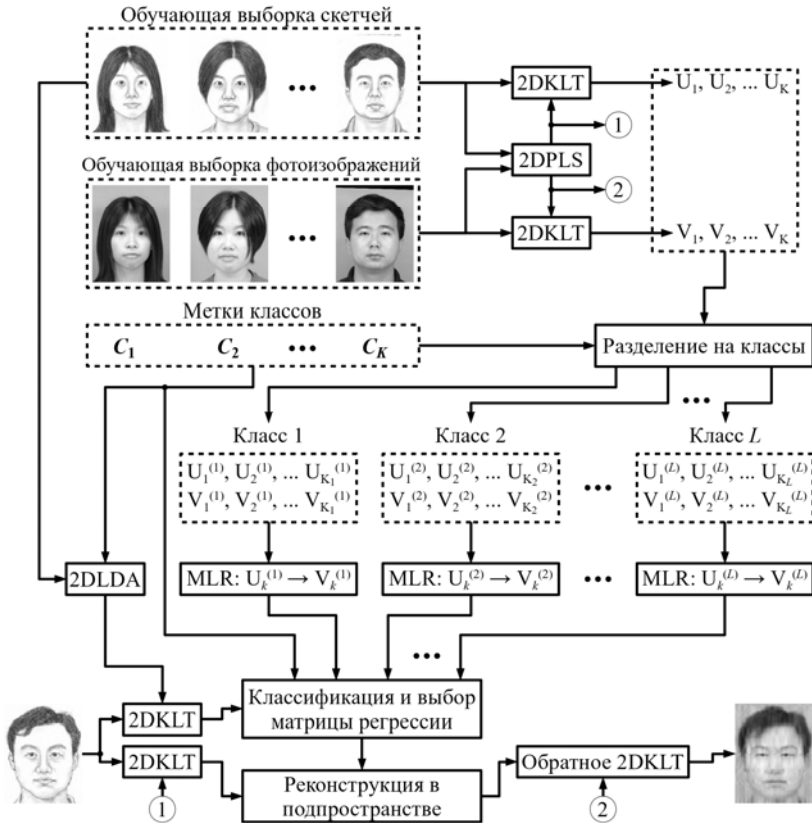


Рис. 13. Реконструкция ИЛ на основе 2DPLS/2DKLT, 2DLDA/2DKLT и MLR

5. Оценка качества реконструкции. Количественная оценка качества реконструкции ИЛ — задача нетривиальная. Поэтому довольно часто в научной литературе вместо прямой оценки качества приводят результаты по точности распознавания ИЛ по реконструированным изображениям. Несмотря на то, что такой способ позволяет получить конкретные количественные результаты, они имеют смысл только для ограниченного набора сценариев. При этом использование других методов распознавания может привести к иным результатам.

Другой подход использует сравнительную оценку результатов реконструкции ИЛ на основе субъективного сравнения собственных результатов с известными из литературных источников по решаемой проблеме. Этот подход применяется в рамках одинаковых сценариев реконструкции ИЛ и относительно близких методов реконструкции.

В настоящей статье рассмотрены новые (оригинальные) сценарии реконструкции ИЛ, а представленные решения основаны на двумерных методах проекции изображений в собственные подпространства. При этом какая-либо связь — корреляция, семантическое подобие или сенсорная природа — между исходными наборами может отсутствовать (несмотря на то, что в собственных пространствах возникает устойчивая корреляционная или ковариационная связь). Аналоги этих сценариев и решений для них, насколько известно авторам, в литературе не описаны, а следовательно, провести сравнительный анализ качества реконструкции ИЛ не представляется возможным.

В настоящей статье качество реконструкции оценивается на основе метрик подобия, не привязанных ни к системам распознавания, ни к известным результатам. Примерами таких метрик являются *фазовая корреляция* и *индекс структурного подобия SSIM (Structural SIMilarity)*.

Фазовая корреляция, в отличие от амплитудной, чувствительна к текстуре изображения, а основные различия между изображениями «мультисенсорной» природы заключаются именно в текстуре. Высота пика фазовой корреляции и определяет степень сходства двух ИЛ.

Индекс SSIM первоначально был предложен для оценки качества фото- и видеоизображений путем сравнения различных версий изображения. Однако в научной литературе индекс SSIM используется также и для сравнения различных ИЛ [26]. Таким образом, в рамках задачи взаимной реконструкции ИЛ использование индекса SSIM как меры подобия исходных и реконструированных ИЛ оправдано.

На рисунке 14 представлены фазовая корреляция и значения индекса SSIM для идентичных фотографий и фотографий лиц разных людей. Видно, что для одного и того же изображения фазовая корреляция имеет единственный пик, достигающий единицы; индекс SSIM также равен единице. При сравнении изображений лиц разных людей индекс SSIM составляет 0,53, а главный пик фазовой корреляции практически не выражен на фоне «шумовой составляющей».

На рисунке 15а аналогичным способом выполнено сравнение фотографии и скетчей (в оригинальном и в низком разрешении) из базы CUF5, соответствующих одному и тому же человеку. На рисунке 15б показано сравнение той же исходной фотографии с фотографиями, реконструированными из соответствующих скетчей рисунка 15а. Ре-

конструкция выполнена на основе 2DPCA/2DKLT и PLS-регрессии. Как фазовая корреляция, так и индекс SSIM указывают на то, что реконструкция значительно повышает подобие между ИЛ.

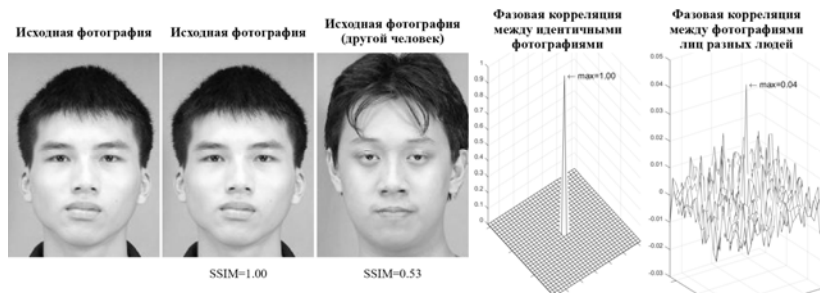


Рис. 14. Значения индекса SSIM и фазовая корреляция для идентичных фотографий и для фотографий лиц разных людей

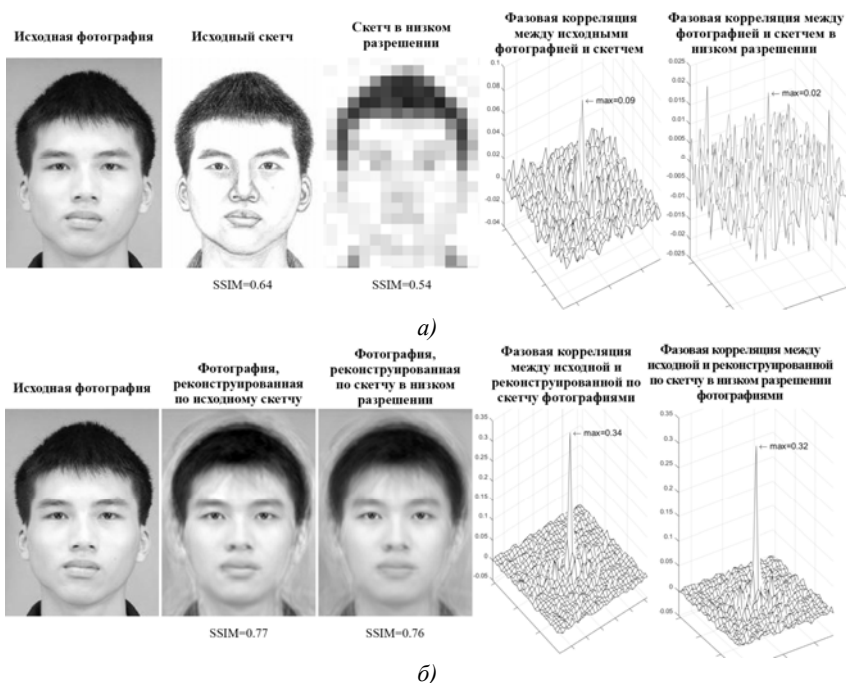


Рис. 15. Значения индекса SSIM и фазовая корреляция для исходной фотографии и соответствующих ей скетчей в оригинальном и в низком разрешении (а) и для исходной и реконструированных фотографий (б)

На рисунке 16 показаны результаты сравнения реконструированных фотографий с фотографией другого человека. Можно видеть, что реконструированные фотографии *не* подобны фотографиям другого человека.

Отметим, что даже при использовании скетча в низком разрешении индекс SSIM и фазовая корреляция ведут себя аналогично тому, как это происходит в случае скетчей в оригинальном разрешении. Здесь важно отметить то, что для построения матриц проекции и регрессии в наших решениях *изображения в низком разрешении на этапе обучения не использовались*.

На основании приведенных рисунков можно сделать вывод о согласованности визуальной оценки подобия изображений и результатов, полученных на основании индекса SSIM и фазовой корреляции.

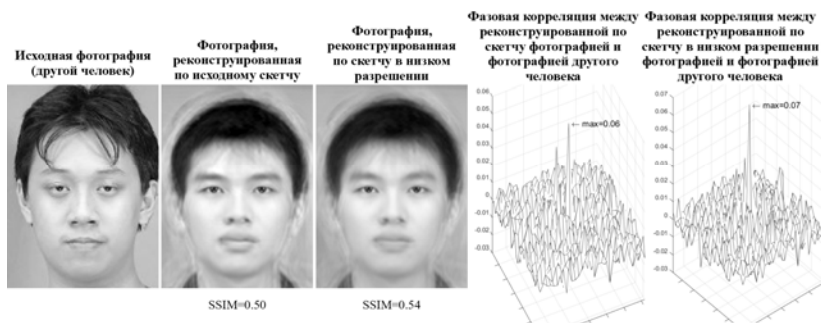


Рис. 16. Значения индекса SSIM и фазовая корреляция для реконструированных фотографий и исходной фотографии другого («постороннего») человека из базы CUFS

В таблице 2 приведены результаты сравнения исходных фотографий со скетчами и реконструированными из них фотографиями для метода на основе 2DPCA/2DKLT и PLS-регрессии, а также метода на основе 2DPLS/2DKLT и MLR. Показаны средние значения индекса SSIM (mean) и его среднеквадратического отклонения (std), полученные по базе CUFS. Кроме того, для результатов, полученных при сравнении исходных фотографий и скетчей, указано максимальное (то есть наилучшее, «наиболее удачное») значение (max). Для значений индекса SSIM, полученных при сравнении исходных и реконструированных из скетчей фотографий, указано минимальное («наименее удачное») значение (min).

Таблица 2. Результаты оценки качества реконструкции на основе индекса SSIM

| | Сравнение исходных фотографий и скетчей | | | Сравнение исходных и реконструированных из скетчей фотографий | | |
|--|---|------|------|---|------|------|
| | mean | std | max | mean | std | min |
| 2DPCA/2DKLT + регрессия PLS | 0.48 | 0.07 | 0.65 | 0.81 | 0.04 | 0.71 |
| 2DPCA/2DKLT + регрессия PLS (скетчи в низком разрешении) | 0.48 | 0.05 | 0.59 | 0.73 | 0.03 | 0.64 |
| 2DPLS/2DKLT + MLR | 0.48 | 0.07 | 0.65 | 0.77 | 0.03 | 0.69 |

На основании рисунков 14-16 и таблицы 2 можно сделать следующие выводы:

- визуальная (субъективная) оценка подобия изображений и результаты, полученные на основании индекса SSIM и фазовой корреляции (объективная оценка), согласуются друг с другом и подтверждают тот факт, что реконструкция достижима;
- фазовая корреляция и индекс SSIM позволяют отличить ИЛ различных людей;
- качество реконструкции остается приемлемым при реконструкции фотографий из зашумлённых скетчей в низком разрешении, даже в том случае, когда для построения матриц проекции и регрессии изображения в низком разрешении не используются на этапе обучения. Это существенно отличает представленный в статье сценарий от рассматриваемых в научной литературе;
- эксперименты на базе CUFS показали, что подобие реконструированных и исходных фотографий по индексу SSIM существенно выше, чем в случае исходных фотографий и скетчей. При использовании фазовой корреляции этот эффект выражен в большей степени, так как в случае исходных фотографий и скетчей фазовая корреляция между ними полностью отсутствует.

6. Заключение. В настоящей работе предложен подход к взаимной реконструкции ИЛ мультисенсорной природы. Этот подход реализуется на основе ряда представленных выше двумерных проекционных методов и регрессионных моделей. Подробно рассмотрены и описаны двумерные проекционные методы, выделены их основные свойства, даны рекомендации по применению.

Для подтверждения практической значимости предложенного подхода проведены экспериментальные исследования на различных базах ИЛ, включающих фотографии лиц, скетчи, карты глубины («range images»). Для проведения экспериментов выбраны два варианта реализации предложенного подхода. Первый из них основан на 2DPCA/2DKLT и PLS-регрессии, второй — на 2DPLS/2DKLT и MLR. Оба варианта показали хорошие результаты при решении задачи реконструкции.

Проведена визуальная (субъективная) и количественная (объективная) оценки качества реконструкции. Для количественной оценки использованы фазовая корреляция и индекс структурного подобия. Оба показателя указывают на достаточное качество реконструкции. Более того, результаты остаются хорошими даже в том случае, когда входные изображения имеют низкое разрешение, несмотря на то, что для построения регрессионных моделей использовались только изображения в высоком разрешении. Это выгодно отличает предложенное решение от ряда методов сверхразрешения, которые обычно применяют в таких ситуациях.

Кроме того, были получены результаты по моделированию возрастных изменений лица человека с помощью методов 2DPCA/2DKLT и регрессии PLS. То, что предложенный подход позволяет решать данную задачу, указывает на то, что область его применения не ограничивается взаимной реконструкцией различных ИЛ и может включать, например, генерацию популяций ИЛ и деидентификацию. Также вероятно, что предложенное решение может быть использовано не только для обработки и преобразования изображений, но и в других задачах анализа связанных данных, таких как, например, разработка многомодальных человеко-машинных интерфейсов.

Литература

1. *Ouyang S. et al.* A survey on heterogeneous face recognition: Sketch, infra-red, 3D and low-resolution // ArXiv Prepr. ArXiv14095114. 2014.
 2. *Kukharev G., Kamenskaya E.* Application of two-dimensional canonical correlation analysis for face image processing and recognition // Pattern Recognition and Image Analysis. 2010. vol. 20. no. 2. pp. 210–219.
 3. *Кухарев Г.А., Каменская Е.И., Матвеев Ю.Н., Щеголева Н.Л.* Методы обработки и распознавания изображений лиц в задачах биометрии / под ред. Хитрова М.В. // СПб: Политехника. 2013. 388 с.
 4. *Ouyang S. et al.* Cross-modal face matching: Beyond viewed sketches // 12th Asian Conference on Computer Vision (ACCV). 2014. pp. 210–225.
 5. *Sharma A., Jacobs D.W.* Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2011. pp. 593–600.
 6. *Tang X., Wang X.* Face sketch recognition // IEEE Trans. Circuits Syst. Video Technol. 2004. vol. 14. no. 1. pp. 50–57.
 7. *Kukharev G., Oleinik A.* Face Photo-Sketch Transformation and Population Generation // International Conference on Computer Vision and Graphics (ICCVG 2016). 2016. LNCS 9972. pp. 329–340.
 8. *Baker S., Kanade T.* Hallucinating faces // Fourth IEEE International Conference on Automatic Face and Gesture Recognition. 2000. pp. 83–88.
 9. *An L., Bhanu B.* Face image super-resolution using 2D CCA // Signal Process. 2014. vol. 103. pp. 184–194.
 10. *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.
 11. *Кухарев Г.А., Матвеев Ю.Н., Олейник А.Л.* Алгоритмы взаимной трансформации изображений для систем обработки и поиска визуальной информации // Научно-
- 70 Труды СПИИРАН. 2018. Вып. 2(57). ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
www.proceedings.spiiras.nw.ru

- технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 1. С. 62–74.
12. Xie Z. Partial least squares regression on DCT domain for infrared face recognition // Twelfth International Conference on Photonics and Imaging in Biology and Medicine (PIBM 2014). International Society for Optics and Photonics. 2014. vol. 9230. id 92301I. pp. 6.
 13. Reiter M. et al. 3D and Infrared Face Reconstruction from RGB data using Canonical Correlation Analysis // 18th International Conference on Pattern Recognition (ICPR 2006). 2006. vol. 1. pp. 425–428.
 14. Reiter M. et al. Estimation of face depth maps from color textures using canonical correlation analysis. Computer Vision Winter Workshop. 2006. pp. 1–5.
 15. Ramanathan N., Chellappa R., Biswas S. Computational methods for modeling facial aging: A survey // J. Vis. Lang. Comput. 2009. vol. 20. no. 3. pp. 131–144.
 16. Liang J. et al. Group-invariant cross-modal subspace learning // Proc. IJCAI. 2016. pp. 1739–1745.
 17. Tian L., Fan C., Ming Y. Multiple scales combined principle component analysis deep learning network for face recognition // J. Electron. Imaging. 2016. vol. 25. no. 2. pp. 16.
 18. Kim H., Fyfe C., Ko H. Feature locations in images // International Conference on Intelligent Data Engineering and Automated Learning. 2008. pp. 459–463.
 19. Эсбенсен К. Анализ многомерных данных // Черноголовка: ИПХФ РАН. 2005. 160 с.
 20. CUFS dataset. URL: <http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html> (дата обращения: 25.04.2015).
 21. CUFS dataset. URL: <http://mmlab.ie.cuhk.edu.hk/archive/cuofs/> (дата обращения: 16.04.2016).
 22. Suo J. et al. A compositional and dynamic model for face aging // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010. vol. 32. no. 3. pp. 385–401.
 23. Gupta S. et al. Texas 3D Face Recognition Database. URL: <http://live.ece.utexas.edu/research/texas3dfr/index.htm> (дата обращения: 23.04.2017).
 24. Gupta S. et al. Texas 3D face recognition database // IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI). 2010. pp. 97–100.
 25. Gupta S., Markey M.K., Bovik A.C. Anthropometric 3D Face Recognition // Int. J. Comput. Vis. 2010. vol. 90. no. 3. pp. 331–349.
 26. Sun Y., Tistarelli M., Maltoni D. Structural similarity based image quality map for face recognition across plastic surgery // 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS). 2013. pp. 1–8.

Олейник Андрей Леонидович — аспирант кафедры речевых информационных систем, ФГАОУ ВО "Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики" (Университет ИТМО), инженер кафедры речевых информационных систем, ФГАОУ ВО "Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики" (Университет ИТМО). Область научных интересов: машинное обучение, цифровая обработка изображений, лицевая биометрия, распознавание образов. Число научных публикаций — 15. aoleinik@corp.ifmo.ru; ул. Красуцкого, 4, Санкт-Петербург, 196084; р.т.: +7(812)325-88-48.

Кухарев Георгий Александрович — д-р техн. наук, профессор, профессор кафедры мультимедийных систем, Западнопоморский технологический университет, профессор кафедры математического обеспечения и применения ЭВМ, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина) (СПбГЭТУ «ЛЭТИ»). Область научных интересов: цифровая обработка изображений, распознавание образов, лицевая биометрия. Число научных публикаций — 255. kuga41@gmail.com; Зольнерская ул., 49, Щецин, 71-210, Польша; р.т.: +48 91-449-56-60.

Поддержка исследований. Исследование проводится при частичной финансовой поддержке Правительства Российской Федерации (грант № 074-U01).

A.L. OLEINIK, G.A. KUKHAREV
**ALGORITHMS FOR FACE IMAGE MUTUAL RECONSTRUCTION
BY MEANS OF TWO-DIMENSIONAL PROJECTION METHODS**

Oleinik A.L., Kukharev G.A. Algorithms for Face Image Mutual Reconstruction by Means of Two-Dimensional Projection Methods.

Abstract. In this paper, we consider the problem of mutual reconstruction of face image pairs. We addressed this problem in our previous article, where the proposed solutions were discussed in connection with Heterogeneous Face Recognition and Cross-Modal Multimedia Retrieval problems. Those solutions are based on one-dimensional and two-dimensional Principal Component Analysis performed over two original face images followed by their projection on independent eigenspaces, estimation of a transformation matrix and mutual reconstruction of the face image by means of one-dimensional and two-dimensional Karhunen-Loève Transform.

In this article, we propose new approaches and solutions, which are based solely on the two-dimensional eigenspace projection methods, and two regression models — Multiple Linear Regression and Partial Least Squares regression.

We present the experiments on mutual reconstruction of face images in sketch/photo pairs, in pairs of face images with age-related changes, and in pairs of 2D/3D face images. In order to conduct the experiments, we selected two variants of the proposed approach. First one is based on two-dimensional Principal Component Analysis and Partial Least Squares regression, and the second one is based on two-dimensional Partial Least Squares and Multiple Linear Regression. Both variants showed acceptable performance for practical applications involving the mutual reconstruction of face images. Furthermore, we consider the method to improve the quality of reconstructed face images in the case of mixed datasets. This method involves classification of the dataset by means of two-dimensional Linear Discriminant Analysis and fitting of a separate regression model for each class.

In addition, we show that generally, mutual reconstruction of face images is also achievable in conditions when original images are not a part of training sets of face images.

Keywords: face image, sketch, facial composite, mutual reconstruction of multisensory face images, cross-modal multimedia retrieval, principal component analysis, partial least squares, two-dimensional projections, regression.

Oleinik Andrei Leonidovich — Ph.D. student of the department of speech information systems, ITMO University (Saint Petersburg National Research University of Information Technologies, Mechanics and Optics), engineer of the department of speech information systems, ITMO University (Saint Petersburg National Research University of Information Technologies, Mechanics and Optics). Research interests: machine learning, digital image processing, face biometrics, pattern recognition. The number of publications — 15. aoleinik@corp.ifmo.ru; 4, Krassutskogo str., St. Petersburg, 196084, Russia; office phone: +7(812)325-88-48.

Kukharev Georgy Aleksandrovich — Dr. Sci., professor, professor of multimedia systems department, West Pomeranian University of Technology, professor of software engineering and computer applications department, Saint Petersburg Electrotechnical University "LETI". Research interests: Digital image processing, pattern recognition, face biometrics. The number of publications — 255. kuga41@gmail.com; 49, Zolnierska St., Szczecin, 71-210, Poland; office phone: +48 91-449-56-60.

Acknowledgements. This research is supported by the Government of the Russian Federation (grant 074-U01).

References

1. Ouyang S. et al. A survey on heterogeneous face recognition: Sketch, infra-red, 3D and low-resolution. ArXiv Prepr. ArXiv14095114. 2014.
2. Kukharev G., Kamenskaya E. Application of two-dimensional canonical correlation analysis for face image processing and recognition. *Pattern Recognition and Image Analysis*. 2010. vol. 20. no. 2. pp. 210–219.
3. Kukharev G.A., Kamenskaya E.I., Matveev Yu.N., Shchegoleva N.L. *Metody obrabotki i raspoznavaniya izobrazhenij lic v zadachah biometrii. Pod red. Hitrova M.V.* [Methods of facial images processing and recognition in biometrics. Edited by Khitrov M.V.]. SPb.: Politehnika. 2013. 388 p. (In Russ.).
4. Ouyang S. et al. Cross-modal face matching: Beyond viewed sketches. 12th Asian Conference on Computer Vision (ACCV). 2014. pp. 210–225.
5. Sharma A., Jacobs D.W. Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2011. pp. 593–600.
6. Tang X., Wang X. Face sketch recognition. *IEEE Trans. Circuits Syst. Video Technol.* 2004. vol. 14. no. 1. pp. 50–57.
7. Kukharev G., Oleinik A. Face Photo-Sketch Transformation and Population Generation. International Conference on Computer Vision and Graphics (ICCVG 2016). 2016. LNCS 9972. pp. 329–340.
8. Baker S., Kanade T. Hallucinating faces. Fourth IEEE International Conference on Automatic Face and Gesture Recognition. 2000. pp. 83–88.
9. An L., Bhanu B. Face image super-resolution using 2D CCA. *Signal Process.* 2014. vol. 103. pp. 184–194.
10. Kukharev G.A., Matveev Yu.N., Forczmański P. [People retrieval by means of composite pictures – methods, systems and practical decisions]. *Nauchno-Tekhnicheskii Vestnik Informatsionnykh Tekhnologii, Mekhaniki i Optiki – Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2015. vol. 15. no. 4. pp. 640–653. (In Russ.).
11. Kukharev G.A., Matveev Yu.N., Oleinik A.L. [Mutual image transformation algorithms for visual information processing and retrieval]. *Nauchno-Tekhnicheskii Vestnik Informatsionnykh Tekhnologii, Mekhaniki i Optiki – Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2017. vol. 17. no. 1. pp. 62–74. (In Russ.).
12. Xie Z. Partial least squares regression on DCT domain for infrared face recognition. Twelfth International Conference on Photonics and Imaging in Biology and Medicine (PIBM 2014). International Society for Optics and Photonics. 2014. vol. 9230. id 92301I. pp. 6.
13. Reiter M. et al. 3D and Infrared Face Reconstruction from RGB data using Canonical Correlation Analysis. 18th International Conference on Pattern Recognition (ICPR 2006). 2006. vol. 1. pp. 425–428.
14. Reiter M. et al. Estimation of face depth maps from color textures using canonical correlation analysis. Computer Vision Winter Workshop. 2006. pp. 1–5.
15. Ramanathan N., Chellappa R., Biswas S. Computational methods for modeling facial aging: A survey. *J. Vis. Lang. Comput.* 2009. vol. 20. no. 3. pp. 131–144.
16. Liang J. et al. Group-invariant cross-modal subspace learning. Proc. IJCAI. 2016. pp. 1739–1745.
17. Tian L., Fan C., Ming Y. Multiple scales combined principle component analysis deep learning network for face recognition. *J. Electron. Imaging*. 2016. vol. 25. no. 2. pp. 16.
18. Kim H., Fyfe C., Ko H. Feature locations in images. International Conference on Intelligent Data Engineering and Automated Learning. Springer. 2008. pp. 459–463.
19. Esbensen K. *Analiz mnogomernykh dannykh* [Multivariate Data Analysis]. Chernogolovka: IPCP RAS. 2005. 160 p. (In Russ.).

20. CUFS dataset. Available at: <http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html> (accessed: 25.04.2015).
21. CUFSF dataset. Available at: <http://mmlab.ie.cuhk.edu.hk/archive/cufsf/> (accessed: 16.04.2016).
22. Suo J. et al. A compositional and dynamic model for face aging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010. vol. 32. no. 3. pp. 385–401.
23. Gupta S. et al. Texas 3D Face Recognition Database. Available at: <http://live.ece.utexas.edu/research/texas3dfr/index.htm> (accessed: 23.04.2017).
24. Gupta S. et al. Texas 3D face recognition database. IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI). 2010. pp. 97–100.
25. Gupta S., Markey M.K., Bovik A.C. Anthropometric 3D Face Recognition. *Int. J. Comput. Vis.* 2010. vol. 90. no. 3. pp. 331–349.
26. Sun Y., Tistarelli M., Maltoni D. Structural similarity based image quality map for face recognition across plastic surgery. 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS). 2013. pp. 1–8.