

# МЕТОДЫ ПРЕДСТАВЛЕНИЯ И СРАВНЕНИЯ СЕМАНТИЧЕСКИ РАЗНЫХ КЛАССОВ ИЗОБРАЖЕНИЙ

**Г.А. Кухарев,**

доктор технических наук, профессор кафедры математического обеспечения ЭВМ Санкт-Петербургского государственного электротехнического университета, профессор факультета информатики Западно-поморского технологического университета, Щецин, Польша (West Pomeranian University of Technology, Szczecin, Poland)

**Е.И. Каменская,**

кандидат технических наук, инженер программного обеспечения компании Google, Цюрих, Швейцария

**Н.Л. Щеголева,**

кандидат технических наук, доцент кафедры математического обеспечения ЭВМ Санкт-Петербургского государственного электротехнического университета

E-mail: : kuga41@gmail.ru, ekamenskaya@google.com, stil\_hope@mail.ru

Адрес: г. Санкт-Петербург, улица Профессора Попова, 5

*В статье обсуждаются методы представления и сравнения семантически не связанных изображений с оценкой их меры подобия в исходном пространстве признаков, а также в пространстве канонических переменных (ПКП). Последнее достигается в рамках алгоритма 2D CCA/2D KLT для реализации проекции этих изображений в ПКП.*

**Ключевые слова:** сравнение семантически не связанных изображений, канонический корреляционный анализ, фазовая корреляция, индекс подобия.

## 1. Введение

В последнее десятилетие активно развивается область информатики, связанная с биометрическими технологиями распознавания и понимания людей [1 – 3]. В рамках этих технологий решаются как относительно простые задачи, например, поиск людей в больших базах по фотопор-

третам, идентификация и аутентификация людей, так и более сложные задачи – понимание намерений людей в процессе распознавания и интерпретации их эмоционального и психофизического состояния, а также задачи, охватываемые, проблемой человеко-компьютерного взаимодействия (human-computer interaction) или проблемой интерфейса мозг-компьютер (brain-computer interface).

Особенность этих задач состоит в наблюдении за человеком с разных, внешне не связанных между собой, сторон и использование информации этих наблюдений для интерпретации или понимания одних наблюдений через другие. Именно поэтому, в последнее время интересы специалистов по компьютерным технологиям и распознаванию образов обратились к математическим методам, позволяющим трансформировать два (не связанных между собой) набора данных из исходного пространства признаков (ИПП) в новое общее подпространство признаков, в котором исходные наборы данных сильно коррелируют. Эти методы строятся на основе проекций на латентные структуры (Partial Least Squares, Projection Latent Structures – PLS) и канонического корреляционного анализа (Canonical Correlation Analysis – CCA).

Метод CCA был представлен Хоттелингом в 30-х годах прошлого столетия и описывал соотношение между двумя **одномерными** рядами данных [7] с прогнозом его применения к многомерным данным. Однако сложность и объем вычислений в CCA привела к тому, что потребовалось семьдесят лет ожиданий на первое практическое применение CCA к наборам цифровых изображений. При использовании CCA в приложении к обработке цифровых изображений возникли две проблемы. Первая связана с многомерностью исходных данных – изображений, как объектов 2D, с третьим измерением, определяющим яркость и четвертым измерением, определяющим цветовые компоненты. Вторая связана с проблемой малой выборки (Small Sample Size) [20], когда размерность ИПП значительно превосходит число изображений в исходной выборке.

Поэтому публикации, связанные с широким применением методов PLS и CCA в задачах обработки изображений вообще и в задачах биометрии, в частности, появились только в последние 5 – 7 лет, когда сформировались новые подходы к реализации PLS и CCA на основе двумерных методов проекции.

Однако и здесь оказалось все не так просто. Можно с сожалением отметить, например, что работа [10], в которой впервые была представлена революционная идея реализации двумерного анализа главных компонент и двумерного преобразования Карунена-Лоэва в приложении к обработке изображений, осталась незамеченной большинством исследователей. Более того, исказилась история развития этого метода: ошибочно считается, что первый раз этот метод был представлен в работе [13], хотя речь в ней идет о проекции только по

одному направлению (а именно по строкам изображения), что не соответствует названию «**Two-dimensional PCA**».

Подчеркнем еще раз, что история метода двумерных проекций в приложении к обработке изображений началась с работы [10], продолжена в работах [1, 3, 11, 12], а для задач, решаемых на основе CCA и PLS, развита и показана в работах [14, 15]. Полная история развития методов представления и обработки изображений в подпространствах, определяемых на основе двумерных проекций в собственных базисах, получаемых в рамках PCA, LDA, PLS и CCA, показана в работе [15].

В работах [14, 15] рассмотрено решение нескольких прикладных задач с использованием бенчмарковых баз изображений лиц ORL, Face94, FERET, EQUINOX. Здесь были показаны: специфика обработки и взаимного распознавания изображений разных категорий (например, инфракрасных, термо- и видео- изображений), представляющих одно и то же лицо человека; особенности представления и обработки гендерно смешанных лиц (мужского и женского), состоящих в одном классе и методы их взаимного распознавания. При этом были рассмотрены случаи, когда все изображения в составах  $X$  и  $Y$  относились к одному и тому же глобальному классу – классу «лица людей». Возможно, это повлияло на полученные высокие результаты корреляции в подпространствах.

*Цель настоящей статьи* – представление методов сравнения семантически не связанных изображений с оценкой меры их подобия в исходном пространстве признаков, а также представление алгоритма 2D CCA/2D KLT для реализации проекции этих изображений в ПКП и оценка меры их подобия в ПКП.

## 2. Структура вычислений в CCA/KLT

В рамках CCA информация об объекте (форма, модель или иная группа признаков) переводится из ИПП в ПКП, и все дальнейшие операции осуществляются в этом пространстве. Рассмотрим структуру вычислений в CCA.

Пусть нам заданы два набора исходных данных  $X$  и  $Y$ , состоящие из  $K$  векторов размером  $DIM \times 1$  каждый так, что:

$$X = [X^{(1)} X^{(2)} \dots X^{(K)}] \text{ и } Y = [Y^{(1)} Y^{(2)} \dots Y^{(K)}]. \quad (1)$$

Целью CCA является нахождение двух матриц проекции, трансформирующих исходные

данные в пространство канонических переменных  $X^{(k)} \rightarrow U^{(k)}$  и  $Y^{(k)} \rightarrow V^{(k)}$  так, чтобы выполнялось условие:  $\|U^{(k)} - V^{(k)}\| \rightarrow \min$ . Исходные данные  $X^{(k)}$  и  $Y^{(k)}$  могут быть не связаны между собой корреляцией, в то время как канонические переменные  $U^{(k)}$  и  $V^{(k)}$ , полученные в результате ССА, связываются между собой устойчивой корреляцией, максимум которой достигается при решении двух совместных задач на собственные значения [7]:

$$\begin{cases} (C_{xx}^{-1}C_{xy}C_{yy}^{-1}C_{yx})W_x = W_x\Lambda_x \\ (C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy})W_y = W_y\Lambda_y, \end{cases} \quad (2)$$

где:  $C_{xx}, C_{yy}, C_{xy}, C_{yx}$  — матрицы ковариации порядка DIM, причем

$C_{xx} = XX^T, C_{yy} = YY^T, C_{xy} = XY^T, C_{yx} = C_{xy}^T$ ;  $\Lambda_x, \Lambda_y$  — диагональные матрицы собственных значений;  $W_x$  и  $W_y$  — матрицы собственных векторов (матрицы проекций).

Если представить результаты ССА следующим образом:

$$U = [U^{(1)}U^{(2)} \dots U^{(K)}] \text{ и } V = [V^{(1)}V^{(2)} \dots V^{(K)}], \quad (3)$$

то условие  $\|U^{(k)} - V^{(k)}\| \rightarrow \min$  для  $\forall k$ , заменяется условием  $\|U - V\| \rightarrow \min$ , которое достигается при максимизации отношения:

$$\rho^2 = \frac{w_x^T C_{xy} w_y}{\sqrt{(w_x^T C_{xx} w_x)(w_y^T C_{yy} w_y)}} = \max, \quad (4)$$

что связано с поиском векторов  $w_x$  и  $w_y$ , входящих в матрицы  $W_x$  и  $W_y$ .

Обозначение ССА/KLT подчеркивает тот факт, что реализация собственно канонического корреляционного анализа (ССА) является отдельным этапом, на котором определяются матрицы проекции  $W_x$  и  $W_y$  (или собственный базис). Заметим, что первый этап может выполнять самостоятельную роль — анализ исходных данных.

Второй этап — собственно проекция исходных данных в пространство канонических переменных (KLT подчеркивает реализацию проекции в собственном базисе, определенном матрицами  $W_x$  и  $W_y$ ) выполняется следующим образом:

$$U = W_x^T X, \quad V = W_y^T Y. \quad (5)$$

В приложении к обработке изображений в биометрии входной информацией для ССА могут быть некоторые признаки исходных изображений (например, яркостные или спектральные в базисе Фурье или базисе собственного преобразования),

антропометрические точки на области лица или координаты (параметры) эластичной модели лица). В отличие от исходных признаков, канонические переменные  $U$  и  $V$  всегда относятся к категории спектральных признаков в собственном базисе, а число признаков  $d$  определяется выбором главных компонент на первом этапе (или числом выбранных базовых функций в матрицах  $W_x$  и  $W_y$  на втором этапе). В любом случае, размерность  $d$  всегда меньше размерности DIM (то есть  $d \ll DIM$ ), что обычно связывают с редукцией размерности пространства признаков.

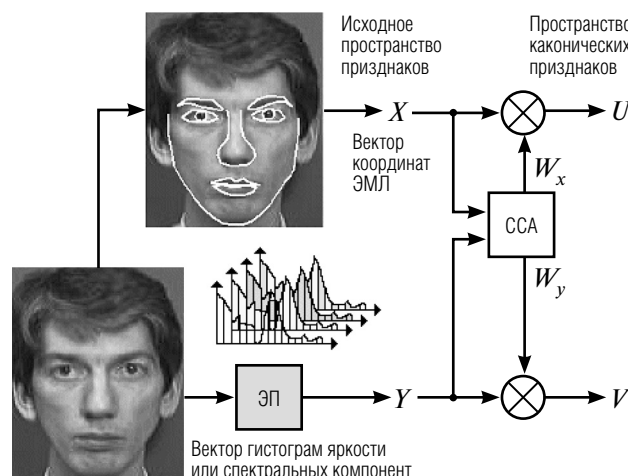


Рис. 1. Пример системы, основанной на ССА

На рис. 1 показан пример системы распознавания изображений лиц на основе ССА, где входными данными являются параметры ЭМЛ и гистограммы яркости изображений или параметры ЭМЛ и спектр изображения лица. Система реализует идею «multirepresentation» — представление одного и того же лица различными составами признаков, которые в общем случае не коррелируют между собой.

Эта система представляет собой типичный пример ранних идей применения ССА в приложении к обработке изображений (как двумерных данных), в которой отразился этап векторного представления исходных изображений (ИПП). При этом часто перед векторизацией исходных изображений, предварительно уменьшался их размер и только после этого выполнялся ССА или сначала выполнялась векторизация исходных изображений, затем редукция размерности ИПП с использованием PCA/KLT или LDA/KLT и, только потом, ССА. Аналогичный путь прошло и применение PLS, поэтому вместо блока ССА на рис. 1, можно было бы поставить блок PLS.

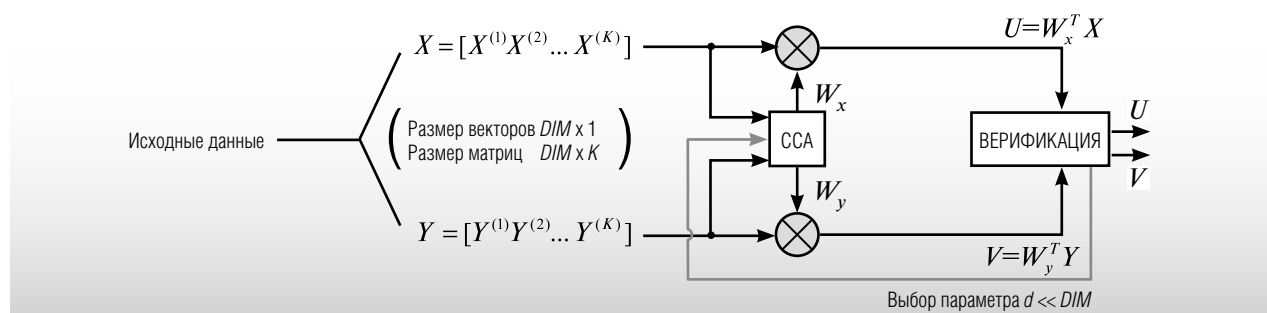


Рис. 2. Исходные данные и вычисления в CCA/KLT

Исходные данные, структуру вычислений в CCA, а также их результаты можно схематически представить так, как показано на рис. 2, где блоки «⊗» реализуют процедуру проекции исходных данных в новое пространство переменных.

Вычисления в CCA заканчиваются выбором параметра  $d$  редукции размерности пространства признаков (причем  $d < DIM$ ), нахождением коэффициентов корреляции  $\rho_k = \sqrt{\lambda_k} \forall k$  и проверкой качества взаимного соответствия переменных в новом пространстве признаков. Эти операции реализуются в блоке «Верификация».

Вычислительная сложность в CCA/KLT определяется размерностью  $DIM$ , а возможность устойчивого решения задач на собственные значения связана с условием:

$$DIM < K \quad (6)$$

Выполнение этого условия особенно важно в приложении к обработке изображений, поскольку размерность  $DIM$  связана с размерами  $M \times N$  изображения (или общим числом его пикселей  $MN$ ). Размерность исходного признакового пространства в этом случае составляет  $DIM = MN$  и обычно  $DIM \gg K$ , что известно как проблема малой выборки. Именно поэтому, методология CCA в приложении к обработке изображений в биометрии и была основана на двух подходах, связанных, с одной стороны, с обеспечением выполнения условия (6), а с другой стороны, с возможностью использования базового метода [7], изначально ориентированного на обработку одномерных данных, представляющих каждый образ.

Представим метод реализации CCA, не требующий предварительной процедуры «DownSampling», и не требующий формального выполнения условия (6). Этот метод развивает идеи двумерной проекции [10].

### 3. Алгоритмы 2D CCA/2D KLT

Пусть заданы два набора исходных данных, состоящие из  $K$  изображений размером  $M \times N$  каждое:

$$X = [X^{(1)} X^{(2)} \dots X^{(K)}] \text{ и } Y = [Y^{(1)} Y^{(2)} \dots Y^{(K)}], \quad (7)$$

где  $X^{(k)}$  и  $Y^{(k)}$  — пары изображений, при этом  $MN \gg K$ .

Определим средние образы:

$$\bar{X} = \frac{1}{K} \sum_{k=1}^K X^{(k)} \text{ и } \bar{Y} = \frac{1}{K} \sum_{k=1}^K Y^{(k)} \quad \forall k=1, 2, \dots, K. \quad (8)$$

Положим  $\tilde{X}^{(k)} = (X^{(k)} - \bar{X})$  и  $\tilde{Y}^{(k)} = (Y^{(k)} - \bar{Y})$ , тогда вместо (8) получим наборы данных, центрированных относительно средних значений:

$$\tilde{X} = [\tilde{X}^{(1)} \tilde{X}^{(2)} \dots \tilde{X}^{(K)}]; \tilde{Y} = [\tilde{Y}^{(1)} \tilde{Y}^{(2)} \dots \tilde{Y}^{(K)}], \quad (9)$$

для которых вычислим по четыре матрицы ковариации:

$$C_{xx}^{(r)} = (\tilde{X}) \tilde{X}^T; C_{yy}^{(r)} = (\tilde{Y}) \tilde{Y}^T; C_{xy}^{(r)} = \tilde{X} (\tilde{Y})^T; C_{yx}^{(r)} = (C_{xy}^{(r)})^T; \quad (10a)$$

$$C_{xx}^{(c)} = (\tilde{X})^T \tilde{X}; C_{yy}^{(c)} = (\tilde{Y})^T \tilde{Y}; C_{xy}^{(c)} = (\tilde{X})^T \tilde{Y}; C_{yx}^{(c)} = (C_{xy}^{(c)})^T. \quad (10b)$$

В выражении (10) знаки « $r$ » и « $c$ » подчеркивают факт вычисления матриц ковариации относительно строк (*row*) и столбцов (*column*) исходных изображений.

Вычислим (с использованием матриц (10)), по две общие (total) матрицы рассеяния:

$$S^{(total1,r)} = [C_{xx}^{(r)}]^{-1} C_{xy}^{(r)} [C_{yy}^{(r)}]^{-1} C_{yx}^{(r)}, \quad (11a)$$

$$S^{(total2,r)} = [C_{yy}^{(r)}]^{-1} C_{yx}^{(r)} [C_{xx}^{(r)}]^{-1} C_{xy}^{(r)}.$$

$$S^{(total1,c)} = [C_{xx}^{(c)}]^{-1} C_{xy}^{(c)} [C_{yy}^{(c)}]^{-1} C_{yx}^{(c)} \quad (11b)$$

$$S^{(total2,c)} = [C_{yy}^{(c)}]^{-1} C_{yx}^{(c)} [C_{xx}^{(c)}]^{-1} C_{xy}^{(c)}$$

В рамках двумерных методов, целью 2D CCA является определение четырех матриц проекции  $W_{x1}$ ,  $W_{x2}$  и  $W_{y1}$ ,  $W_{y2}$  трансформирующих исходные

данные в пространство переменных  $X^{(k)} \rightarrow U^{(k)}$  и  $Y^{(k)} \rightarrow V^{(k)}$ , что достигается решением четырех задач на собственные значения:

$$\begin{aligned} S^{(total1,r)} W_{x_1}^{(r)} &= \Lambda_{x_1}^{(r)} W_{x_1}^{(r)}; & S^{(total1,c)} W_{x_2}^{(c)} &= \Lambda_{x_2}^{(c)} W_{x_2}^{(c)}; \\ S^{(total2,r)} W_{y_1}^{(r)} &= \Lambda_{y_1}^{(r)} W_{y_1}^{(r)}; & S^{(total2,c)} W_{y_2}^{(c)} &= \Lambda_{y_2}^{(c)} W_{y_2}^{(c)}, \end{aligned} \quad (12)$$

где  $\Lambda_{x_1}, \Lambda_{y_1}, \Lambda_{x_2}, \Lambda_{y_2}$  – диагональные матрицы собственных значений для данных по строкам (1) и столбцам (2) соответственно;  $W_{x_1}, W_{x_2}$  и  $W_{y_1}, W_{y_2}$  – матрицы проекции.

Трансформация исходных данных в новое пространство реализуется как двумерное преобразование Карунена-Лоэва в форме:

$$U^{(k)} = W_{x_1}^T \bar{X}^{(k)} W_{y_1} \text{ и } V^{(k)} = W_{x_2}^T \bar{Y}^{(k)} W_{y_2}, \forall k, \quad (13)$$

где  $U^{(k)}, V^{(k)}$  – матрицы, представляющие исходные изображения в новом пространстве признаков.

Определим главные компоненты. Ими будут те из них, которые имеют наибольшие значения. При этом основная изменчивость матриц ковариации может быть «описана» всего  $d \ll K$  наибольшими собственными значениями, т.е. если собственные значения упорядочены по убыванию, то при любых  $d_1 \geq 1$  и  $d_2 \geq 1$  всегда выполняются следующие условия:

$$\sum_{i=1}^{d_1} (\lambda_i^{(r)})^2 \gg \sum_{i=d_1+1}^M (\lambda_i^{(r)})^2 \text{ и } \sum_{i=1}^{d_2} (\lambda_i^{(c)})^2 \gg \sum_{i=d_2+1}^N (\lambda_i^{(c)})^2 \quad (14)$$

где  $\lambda_i^{(r)}$  и  $\lambda_i^{(c)}$  – собственные значения;  $d_1 \ll M$ ;  $d_2 \ll N$  и  $d_1 \neq d_2$  в общем случае.

Нижняя граница параметра  $d$  может быть выбрана экспериментально на этапе верификации результата выполненного анализа по критерию (4).

Для редукции размерности процедуру (13) необходимо модифицировать так, чтобы в проекции участвовали только те собственные векторы, которые соответствуют  $d$  главным компонентам. Для этого, из матриц  $W_{x_1}, W_{y_1}$  выберем  $d$  строк, соответствующих  $d$  наибольшим собственным значениям, и на их основе сформируем матрицы редукции  $F_{x1}$  и  $F_{y1}$ . А из матриц  $W_{x_2}, W_{y_2}$  выберем  $d$  столбцов, соответствующих  $d$  наибольшим собственным значениям, и на их основе сформируем матрицы  $F_{x2}$  и  $F_{y2}$ .

Для каждого исходного изображения выполним «усеченное» двумерное преобразование Карунена-Лоэва, которое может быть представлено в следующей форме:

$$\hat{U}^{(k)} = F_{x1} \bar{X}^{(k)} F_{x2}, \hat{V}^{(k)} = F_{y1} \bar{Y}^{(k)} F_{y2} \forall k. \quad (15)$$

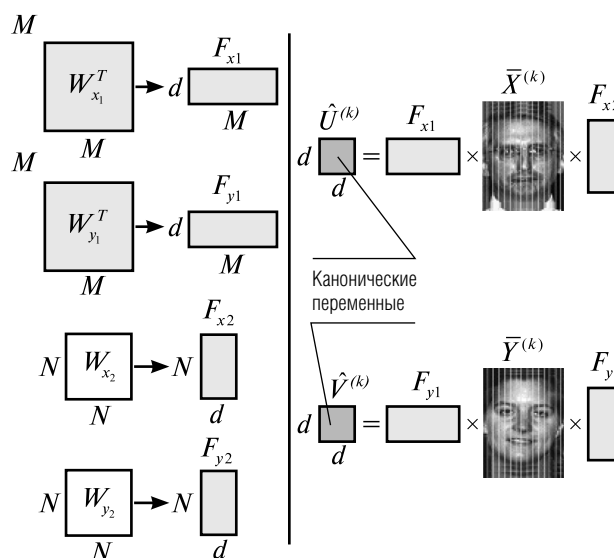


Рис. 3. Матрицы проекции и их использование

Здесь знак « $\hat{\phantom{x}}$ » определяет отличие результата от (13).

Размер матриц-результатов в (15) равен  $d \times d$  или  $d_1 \times d_2$ . Результирующие матрицы (наборы) переменных определяемые как

$$\hat{U} = [\hat{U}^{(1)} \hat{U}^{(2)} \dots \hat{U}^{(K)}] \text{ и } \hat{V} = [\hat{V}^{(1)} \hat{V}^{(2)} \dots \hat{V}^{(K)}] \quad (16)$$

будут иметь размеры  $d \times d \times K$  или  $d_1 \times d_2 \times K$ .

Реализация этапа формирования матриц проекции и выполнение проекции (15) отражена на рис. 3: формирование матриц  $F_{x1}, F_{x2}, F_{y1}$  и  $F_{y2}$  (слева) и схема вычисления переменных (справа). И, как видно, именно на этом этапе и реализуется двумерное преобразование Карунена-Лоэва (2D KLT – Two Dimensional Karhunen-Loeva Transform).

#### 4. Характеристика метода 2D CCA/2D KLT

Метод 2D CCA/2D KLT основан на представлении исходных изображений как совокупности строк и столбцов и реализации CCA по этим совокупностям.

Метод не требует предварительного уменьшения размеров исходных изображений, не является итерационным и прямо реализуется по двум направлениям координат – а именно по строкам и столбцам исходного изображения, что определено в (13) и (15).

Максимальный порядок матриц ковариации в методе 2D CCA/2D KLT равен  $DIM = \max\{M, N\}$ ,

что предопределяет практическую возможность решения задачи на собственные значения и стабильность этого решения даже для изображений очень больших размеров.

Кроме того, проблема SSS в этих методах также обходится стороной, поскольку вместо каждого изображения размером  $M \times N$  фактически используется  $N$  изображений размером  $M \times 1$  и  $M$  изображений размером  $N \times 1$ . При таком представлении исходных данных, всегда выполняется условие:  $DIM = \max\{M, N\} < (KM)$  или  $DIM = \max\{M, N\} < (KN)$ , если  $K \geq 2$ .

### 5. Исходная база изображений для экспериментов

Рассмотрим случай, когда исходные изображения относятся к разным глобальным классам. Примером являются изображения базы «Люди и Собаки» («People and Dogs»), содержащей пары изображений — портреты хозяина и его собаки, подобранные так, чтобы показать, что между портретами существует некоторое подобие (см. рис. 4) [16].

При помощи данной базы изображений можно выяснить два принципиальных вопроса:

- 1) насколько наша визуальная оценка подобия этих портретов в исходной их форме, сравнима с формальными оценками подобия этих портретов;
- 2) как изменяется подобие этих портретов в ПКП.

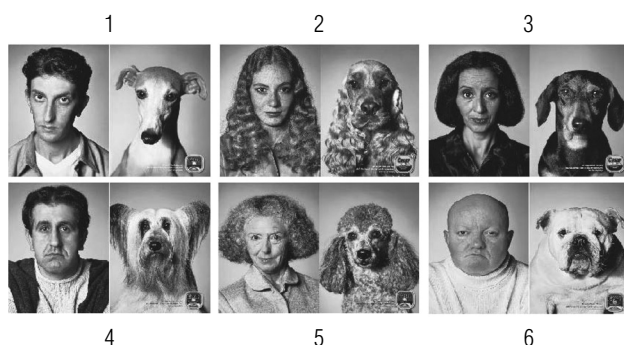


Рис. 4. Примеры пар изображений базы «Люди и собаки» [16]

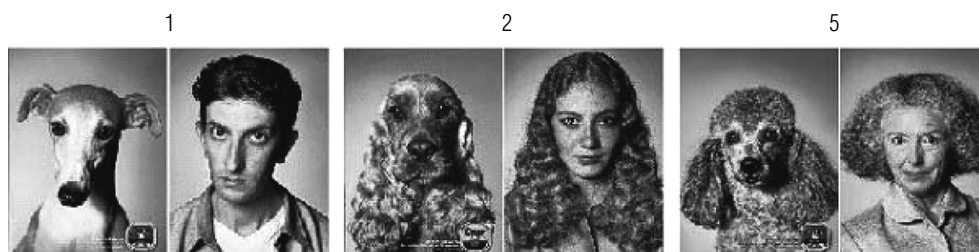


Рис. 5. Выбранные пары изображений

### 6. Сравнение в исходном пространстве признаков

Пары изображений на рис. 4, имеют некоторое внешнее сходство между хозяином и его собакой, которое определяется следующими факторами: близким по форме выражением лица хозяина и «морды» собаки; одинаковым ракурсом двух портретов; близкой по форме прической хозяина и экстерьером собаки; цветовой гаммой — одинаковый цвет волос хозяина и окраски шерсти собак, соответствующий цвет одежды.

Очевидно, что наиболее просто можно проверить отмеченное сходство по цветовой гамме. Формально его можно оценить, сравнивая соответствующие яркостные гистограммы исходных цветных изображений.

Покажем это для 1, 2 и 5 пар исходных изображений, поскольку они имеют наиболее выразительное сходство (см. рис. 5).

Вычислим для каждого изображения цветовую яркостную гистограмму [19] и фазовую корреляцию для каждой пары изображений [1] (см. рис. 6).

Сравнивая гистограммы изображений каждой пары можно «на глаз увидеть их сходство», которое формально оценено через фазовую корреляцию. Пик этой корреляции превышает значение 0.75, при относительно малых ее боковых значениях, что свидетельствует о высокой степени подобия между изображениями внутри каждой пары. Максимум корреляции ( $\approx 0.9$ ) достигается для пары № 1, изображения которой имеют наиболее выразительное сходство.

Тем самым показано, что, по крайней мере, субъективная и формальная оценки фактора цветовой гаммы указывают на существование подобия в двух изображениях, семантически не связанных между собой.

На рис. 7 и 8 представлены изображения базы «Люди и Собаки» и результаты фазовой корреляции самих исходных изображений. Здесь: 1, 2 — портрет человека и собаки; 3 — автокорреляции для портрета человека; 4 — взаимная фазовая корреляция между портретом человека и собаки.

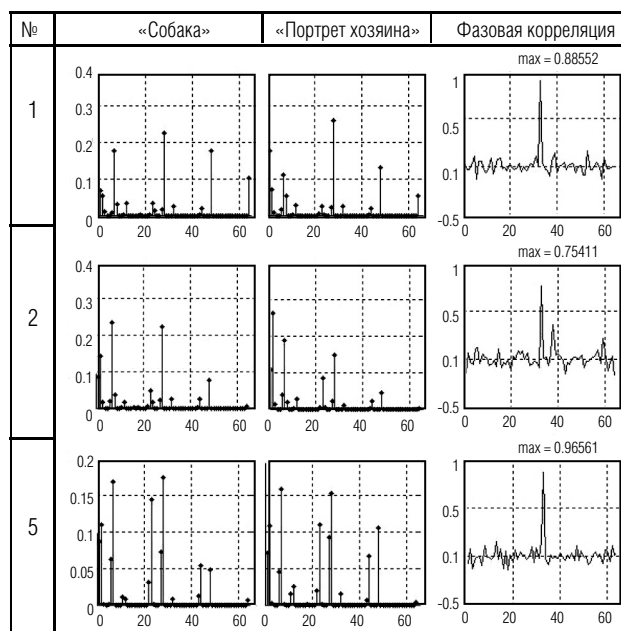


Рис. 6. Яркостные гистограммы исходных цветных изображений

Фазовая автокорреляция представлена ярко выраженным импульсом со значением равным 1. Взаимная корреляция при этом практически равна 0. Таким образом, между парами исходных изображений нет никакой корреляции (!), в то время как наша субъективная оценка и предыдущий эксперимент свидетельствует об обратном – нам кажется, что портреты подобны между собой! И трудно даже представить, какими формальными способами это можно подтвердить и оценить.

На нашу субъективную оценку влияют факторы цвета, форма и текстура изображений, поэтому сравним их по индексу структурного подобия (Structural SIMilarity Index – SSIM ( $I_{SSIM}$ ) [17,18]), который позволяет оценить степень подобия (искажения) двух изображений как комбинацию трех факторов: яркостных изменений, изменений контраста и потери корреляции [18]:

$$I_{SSIM}(X, Y) = [L(X, Y)]^\alpha [C(X, Y)]^\beta [R(X, Y)]^\gamma, \quad (17)$$

где:  $X, Y$  – пара исходных изображений;  $\alpha > 0, \beta > 0, \gamma > 0$  определяют значимость каждого фактора в (17).

Поскольку исходные изображения «в парах» по субъективным оценкам не значительно отличаются друг от друга, положим  $\alpha = \beta = \gamma = 1$ . Тогда изменения яркости  $L$ , контраста  $C$  и корреляции  $R$  можно выразить следующим образом:

$$L(X, Y) = (2\mu_x\mu_y + C1) / (\mu_x^2 + \mu_y^2 + C1), \quad (18)$$

$$C(X, Y) = (2\sigma_x\sigma_y + C2) / (\sigma_x^2 + \sigma_y^2 + C2), \quad (19)$$

$$R(X, Y) = (\sigma_{xy} + C3) / (\sigma_x\sigma_y + C3), \quad (20)$$

где  $\mu_x$  и  $\mu_y$  – среднее значение каждого изображения в паре;  $\sigma_x$  и  $\sigma_y$  – стандартное отклонение;  $\sigma_x^2$  и  $\sigma_y^2$  – дисперсии;  $\sigma_{xy}$  – ковариация изображений  $X$  и  $Y$ .

При обработке изображений может возникнуть ситуация, когда знаменатели в (18) – (20) становятся близки к нулю. Использование констант  $C1, C2$  и  $C3$  позволяет это учесть [18]. Для изображений шкалы GRAY с максимальным уровнем серого цвета имеем  $Lmax = 255$ , поэтому, если  $C1 = (K1L)2$ ,

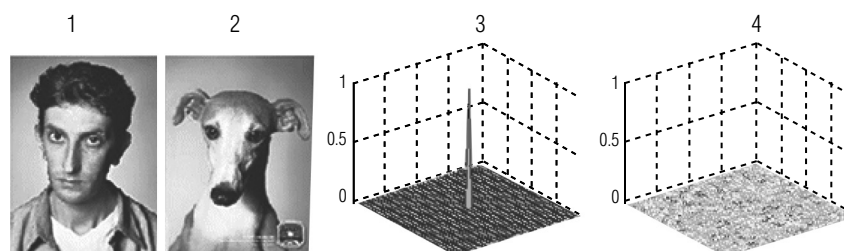


Рис. 7. Исходные изображения (пара 1)

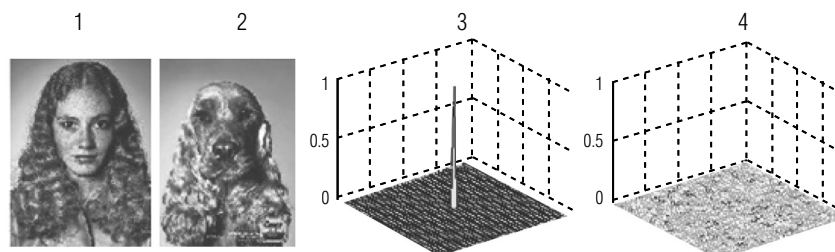


Рис. 8. Исходные изображения (пара 2)

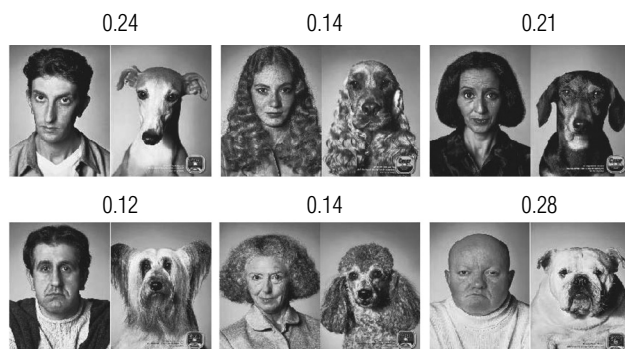


Рис. 9. Значения SSIM для исходных пар изображений

$C2 = K2L/2$ ,  $C3 = C2/2$  [17] и  $K1 = 0.01$  и  $K2 = 0.03$ , то значения констант  $C1$ ,  $C2$  и  $C3$  могут быть легко вычислены.

Результат оценки SSIM для исходных пар изображений показаны на рис. 9.

Если положить  $C1 = C2 = 0$ , то (17) сводится к универсальному индексу качества (Universal Quality Index – UQI [17]), который является предшественником индекса SSIM. Тогда (20), при отсутствующем  $C3$ , представляет собой амплитудную линейную корреляцию двух изображений [1]. Из оценок на рис. 9 видно, что индекс SSIM также не выявил заметного сходства внутри пар изображений «хозяин-собака».

## 7. Представление семантически разных классов изображений в ПКП

Теперь сформулируем новую цель – проверка наличия корреляции между изображениями «хозяин-собака» в ПКП на основании двумерных проекций (13) и (15).

На рис. 10 приведены пары исходных данных из набора  $X$  (набор портретов хозяина) и  $Y$  (набор портретов собаки) для одного и того же класса и разных классов, а также коэффициенты взаимной фазовой

корреляции между ними. И, как видно, в ПКП фазовая корреляция «в своем классе» определена на уровне  $\approx 0.6$ , что существенно выше фазовой корреляции для изображений разных классов (здесь  $\max \approx 0.15$ ) и существенно выше соответствующих значений SSIM.

Рассмотрим рис. 11, где представлены новые результаты анализа подобия образов в ПКП. Здесь в левом верхнем углу показан график зависимости  $U = f(V)$ , далее – две пары исходных изображений и, наконец, – канонические переменные  $U$  и  $V$ , соответствующие исходным изображениям.

Легко заметить, что три канонические переменные из четырех содержат одинаковые значения (яркостные элементы) и, следовательно, даже внешне похожи друг на друга.

В нижнем ряду рис. 11 слева направо показаны следующие виды фазовой корреляции:

- между всеми переменными  $U$  и  $V$ ;
- между переменными, соответствующими изображениям из одного класса;
- между переменными, соответствующими изображениям из разных классов.

Из представленных результатов видно, что в ПКП между переменными  $U$  и  $V$  появилась корреляция. При этом корреляция между переменными из одного класса значительно выше значения 0,5. Корреляцию в другом случае можно считать несущественной, а отношение этих корреляций составляет 4 раза.

Графическое представление связи между переменными в новом пространстве признаков для ССА, полученными по (15), показано на рис. 12.

Заметим, что точечные области представления канонических переменных (слева) – как функции  $U$  от  $V$  связаны практически линейной зависимостью (прямая линия, проходящая через середину точечной области). И, кроме того, с ростом значений  $V$

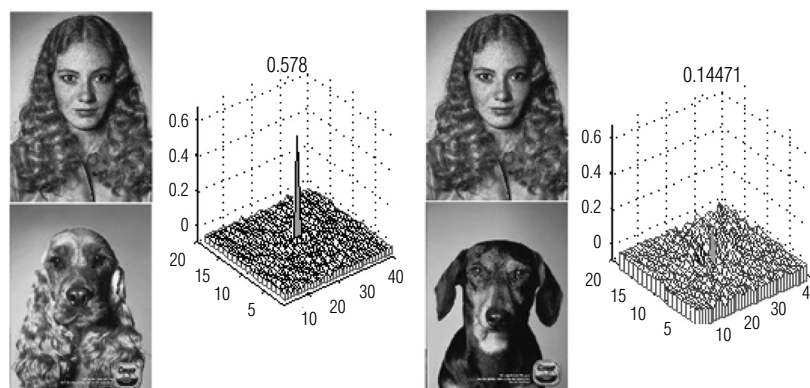


Рис. 10. Исходные данные, вид и коэффициенты фазовой корреляции



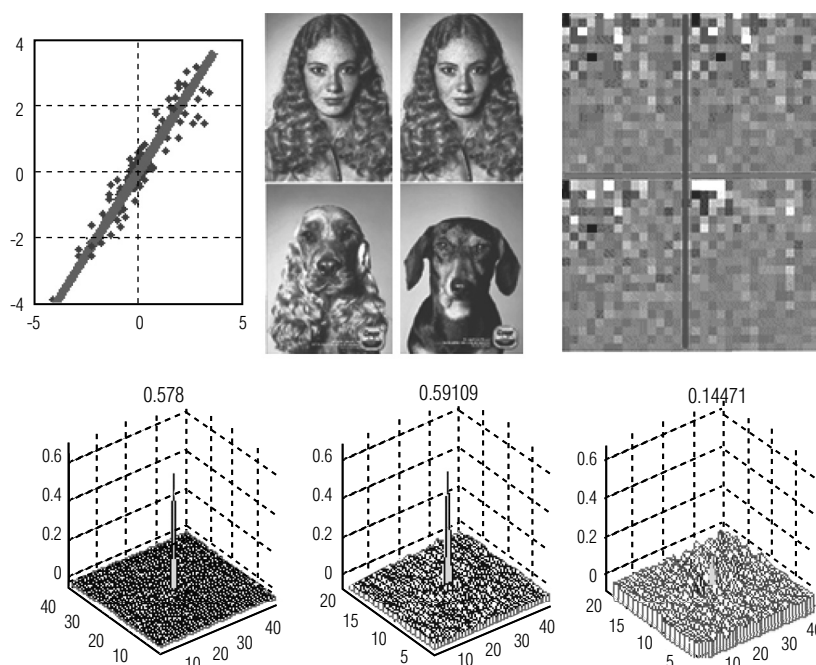


Рис. 11. Исходные данные и результаты их преобразований

растут и значения  $U$ , что соответствует критерию (4).

Другой вариант соотношений переменных  $U$  и  $V$  показан на рис. 12 справа — кривые  $U^{(k)}$  и  $V^{(k)}$  практически синхронно совпадают (коррелируют между собой) и находятся «в фазе».

Высокая корреляция переменных  $U$  и  $V$  в ПКП, их линейная зависимость друг от друга, «синхронизм в поведении» создают предпосылки для взаимного распознавания и модельного отображения одних переменных по другим. И эти задачи легко реализуется в ПКП, что было показано, например, в работах [13, 14].

Аналогичные результаты были получены и для других пар исходных образов, хотя мы исходили из того, что соответствующие пары исходных образов (хозяин-собака) принадлежали к разным глобальным (или семантически разным) классам и факти-

чески их подобие было доказано только на уровне цветовой гаммы.

Главный итог выполненного анализа состоит в следующем:

1) подтверждается факт установления корреляции между образами собаки и ее хозяина в ПКП, в то время как никакими другими способами это подтвердить не удалось;

2) используя п. 1, можно в рамках методов 2D ССА решать задачи индексирования одних изображений через другие (поиск, распознавание, модельное отображение одних образов в другие, реконструкция образов);

В целом же, показанный пример использования разработанных алгоритмов 2D ССА/2D KLT показал, что они могут быть широко использованы в задачах распознавания и классификации образов и сокращения избыточности их представления. Последнее связано с тем, что выбираемое значение параметра  $d \ll MN$  — размерности ИПП.

## 8. Заключение

В статье обсуждались методы представления и сравнения семантически не связанных изображений с оценкой их меры подобия в исходном пространстве признаков, а также представление алгоритма 2D ССА/2D KLT для реализации проекции

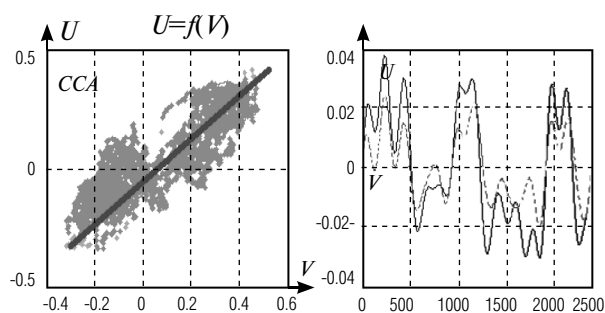


Рис. 12. Соотношение переменных в новом пространстве признаков

этих изображений в ПКП и оценка их меры подобия в ПКП.

Для сравнения изображений в исходном пространстве признаков использованы гистограммы цветных изображений и их фазовая корреляция, двумерная фазовая корреляция исходных изображений, а также индекс структурного подобия. Однако только частично удалось доказать подобие, соответствующее субъективным ощущениям для выбранных примеров.

При этом показано, что «неподобное в исходном пространстве признаков» может быть подобным в ПКП. Проекция в ПКП реализуется в рамках метода 2D CCA/2D KLT, специально предназначенного для обработки двух составов изображений и детально представленного в статье. И этот результат показал, что методы 2D CCA/2D KLT могут быть широко использованы в задачах поиска, распознавания и классификации образов, а также сокращения избыточности их представления независимо от их семантической взаимосвязи. ■

### Литература

1. Kukharev G. Biometric Systems: Methods and Means of People Identification. — Sankt-Petersburg: Politehnika, 2001.
2. Encyclopedia of Biometrics / Li Stan Z. (editor). — Springer Science+Business Media, 2009.
3. Kukharev G., Kuzminski A. Biometric Technology: Face Recognition Methods. — Szczecin: WI PS, 2003.
4. Bie Tjil De, Nello C., Roman R. Eigenproblems in Pattern Recognition // Handbook of Geometric Computing: Applications in Pattern Recognition, Computer Vision, Neuralcomputing, and Robotics. — Heidelberg: Springer-Verlag, 2005, P. 1-39.
5. Pearson K. On lines and planes of closest fit to systems of points in space // Edinburgh and Dublin Philosophical Magazineand Journal of Sciences. — 1901. — V. 6, № 2. — P. 559-572.
6. Hotelling H. Analysis of complex variables into principal components // Journal of Educational Psychology. — 1933. — №24. — P. 417-441.
7. Hotelling H. Relations between two sets of variates // Biometryka. — 1936. — №28. — P. 321-377.
8. Sirovich L., Kirby M. Low-dimensional Procedure for the Characterization of Human Faces // Journal of the Optical Society of America, A - Optics, Image Science and Vision. — 1987. — V. 4, №3. — P. 519-524.
9. Turk M., Pentland A. Eigenfaces for Recognition // Journal of Cognitive Neuroscience. — 1991. — V. 3, №1. — P. 71-86.
10. Tsapatsoulis N., Alexopoulos V., Kollias S. A Vector Based Approximation of KLT and Its Application to Face Recognition. — Proc. IX European Signal Processing Conf. EUSIPCO'98, Greece. — 1998. — V. III. — P. 1581-1584.
11. Kukharev G., Forczmański P. Data Dimensionality Reduction for Face Recognition // Journal Machine Graphics & Vision. — 2004. — V.13 №1/2. — P. 99-121.
12. Kukharev G., Forczmanski P. Face Recognition by Means of Two-Dimensional Direct Linear Discriminant Analysis // PRIP'05, Minsk, Republic of Belarus. — 2005. — V. 2. — P. 63-67.
13. Yang J., Zhang D., Frangi A.F., Yang J-Y. Two-dimensional PCA: A new approach to appearance-based face representation and recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2004. — №26(1). — P. 131-137.
14. Kukharev G., Kamenskaya E. Application of Two\_Dimensional Canonical Correlation Analysis for Face Image Processing and Recognition // Pattern Recognition and Image Analysis. — 2010. — V. 20, №2. — P. 210-219.
15. Kukharev G., Tujaka A., Forczmański P. Face Recognition using Two-dimensional CCA and PLS // International Journal of Biometrics. — 2011. — №3. — P. 300-321.
16. [www.popular-pics.com/Funny\\_People\\_And\\_Dog\\_Similarity\\_Pictures\\_\\_1](http://www.popular-pics.com/Funny_People_And_Dog_Similarity_Pictures__1)
17. Wang Z., Bovik A.C. A universal image quality index // IEEE Signal Process. Letters. — 2002. — V. 9, №3. — P. 81-84.
18. Dosselmann R., Yang X.D. A comprehensive assessment of the structural similarity index // SIViP. — 2011. — №5. — P. 81-91.
19. Smith John R., Chang Shih-Fu. Tools and Techniques for Color Image Retrieval // In IS&T/SPIE Symposium on Electronic Imaging: Science and Technology (EI'96) - Storage and Retrieval for Image and Video Databases IV. — 1996. — V. 2670, San Jose, CA, February 1996. — P. 426-437.
20. Мокеев В.В., Томилов С.В. О решении проблемы выборки малого размера при использовании линейного дискриминантного анализа в задачах распознавания лиц // Бизнес-информатика. — 2013. — №1(23). — С. 37-42.