
РАСПОЗНАВАНИЕ ОБРАЗОВ И ОБРАБОТКА ИЗОБРАЖЕНИЙ

УДК 004.932.2

МЕРЫ СХОДСТВА И МЕТРИКИ СРАВНЕНИЯ ФОРМ МОЗАИЧНЫХ ИЗОБРАЖЕНИЙ*

© 2014 г. Ю. В. Визильгер, С. Ю. Желтов

Москва, ФГУП “ГосНИИАС”

Поступила в редакцию 31.01.14 г.

Рассмотрена задача определения процедуры морфологического сравнения мозаичных форм, согласованной с морфологическим анализом Пытьева в части сравнения изображений с формами. Исследованы четыре различных способа построения такой процедуры сравнения форм на основе: поиска критических точек морфологического коэффициента корреляции; анализа свойств суперпозиции соответствующих им проекторов; оценки раствора угла между многообразиями; метрики Хаусдорфа между соответствующими многообразиями на единичной гиперсфере. Для всех подходов независимо получены меры сходства или различия мозаичных форм. При этом меры сходства, основанные на использовании первых двух подходов (коэффициенты морфологической корреляции форм), совпали, а два других подхода привели к формированию мер отличия форм, связанных с коэффициентом морфологической корреляции форм взаимно-однозначными зависимостями.

DOI: 10.7868/S0002338814040167

Введение. Любые задачи распознавания образов на изображениях так или иначе должны опираться на методы оценки их сходства или различия. При этом могут сравниваться либо непосредственно изображения, либо какие-то их описания или модели, полученные в ходе анализа изображений. Современные модельные (model-based) подходы к обнаружению и распознаванию объектов, как правило, предполагают их описание не векторами независимых признаков, а сложными структурными моделями, которые могут быть описаны, например, графами, либо грамматическими конструкциями. Если такое структурное описание реализуется непосредственно на уровне изображения, причем по такому описанию изображение может быть с той или иной точностью восстановлено, то говорят об описании формы или морфологическом описании изображения. Классическими примерами таких морфологических подходов являются широко известные морфологии Серра [1] и Пытьева [2]. Более широкий спектр возможных морфологических систем был описан в наших предыдущих работах в рамках формализма так называемой проективной морфологии [3–6].

Морфологический анализ позволяет решать различные задачи обработки видеоданных — от фильтрации и сегментации до выделения отличий. Однако использование морфологических описаний в задачах распознавания требует, прежде всего, определить некоторый набор метрик или мер сходства для сравнения изображений по форме (сравнения изображений с формой других изображений) или сравнения собственно форм изображений, описываемых в тех или иных морфологических системах.

Эта задача не столь тривиальна, поскольку морфологические модели не являются векторами независимых признаков. В частности, в морфологии Пытьева так называемые мозаичные модели описывают форму изображения как набор непересекающихся областей разбиения кадра. Как сравнивать такие разбиения между собой? Априори вообще неясно, образуют ли они метрическое пространство, и если да, то какими свойствами должна обладать соответствующая метрика, чтобы быть практически полезной. Поэтому в морфологии Пытьева решается другая проблема — предлагаются средства сравнения не форм с формами, а изображений с формами. Таким образом, несмотря на многолетнее плодотворное развитие методов морфологического анализа Пытьева, задачу морфологического сравнения мозаичных форм до сих пор нельзя считать удовлетворительно решенной.

* Работа выполнена при финансовой поддержке РФФИ (грант № 14-07-00914-а).

Данная работа посвящена разработке таких методов оценки сходства и различия мозаичных форм, которые решали бы задачу сравнения форм и при этом были полностью согласованы с пытьевской морфологией в части сравнения изображений с формами, являясь в этом смысле ее логическим продолжением и развитием. Ниже будут исследованы четыре различных способа решения данной задачи, которые, как будет показано в итоге, приводят к эквивалентным конструкциям оценки сходства/различия мозаичных форм, представленных компактными многообразиями в пространстве изображений.

1. Задача сравнения изображений и морфологический анализ Пытьева. Опишем кратко подход, предложенный Пытьевым. Под изображением будем понимать двумерную функцию распределения яркости на кадре

$$f(x, y): \Omega \rightarrow R, \quad \Omega \subset R^2,$$

где Ω – область кадра ограниченного размера, R – множество действительных чисел, R^2 – плоскость кадра; $x \in R, y \in R$ – соответственно горизонтальная и вертикальная декартовы координаты точек изображения. Кадр будем считать прямоугольным, хотя в общем случае форма кадра может быть произвольной.

Изображения рассматриваются как элементы гильбертова пространства $L^2(\Omega)$ функций с интегрируемым квадратом, и следовательно определены скалярное произведение

$$(f, g) = \iint_{\Omega} f(x, y)g(x, y)dx dy,$$

норма

$$\|f\| = \sqrt{(f, f)}$$

и евклидово расстояние между изображениями

$$d_E(f, g) = \|f - g\|.$$

Это расстояние характеризует различия между изображениями как функциями яркости. В частности, при совпадении изображений $d_E(f, f) = 0$. Однако оно не вполне соответствует нашим элементарным интуициям относительно визуального сравнения изображений. В частности, при линейных преобразованиях яркости изображений, расстояния между ними могут существенно меняться. Поэтому для того, чтобы сравнивать изображения инвариантно к линейным преобразованиям яркости, как правило, вводится нормированный коэффициент линейной корреляции вида

$$K(f, g) = \frac{(f, g)}{\|f\| \|g\|}.$$

При этом изображения также обычно центрируют, т.е. вычитают из них среднее постоянное значение яркости по кадру. В дальнейшем мы будем предполагать, что всюду, где специально не оговорено обратное, все рассматриваемые изображения центрированы по яркости.

В рамках простейшей морфологии Пытьева [2] изображения рассматриваются как кусочно-постоянные функции вида

$$f(x, y) = \sum_{i=1}^n f_i \chi_{F_i}(x, y), \tag{1.1}$$

где n – число областей разбиения \mathbf{F} кадра Ω на связные непересекающиеся области постоянной яркости, $\mathbf{F} = \{F_1, \dots, F_n\}$; $\mathbf{f} = (f_1, \dots, f_n)^T$ – вектор значений яркости, соответствующих каждой области разбиения; $\chi_{F_i}(x, y) \in \{0, 1\}$ – характеристическая функция i -й области яркости:

$$\chi_{F_i}(x, y) = \begin{cases} 1, & \text{если } (x, y) \in F_i; \\ 0 & \text{в противном случае.} \end{cases}$$

Такие изображения называются мозаичными. *Форма* мозаичного изображения f определяется как множество изображений, имеющих то же разбиение кадра \mathbf{F} , и представляет собой линейное подпространство $F \subseteq L^2(\Omega)$:

$$F = \left\{ f(x, y) = \sum_{i=1}^n f_i \chi_{F_i}(x, y), \mathbf{f} \in R^n \right\}.$$

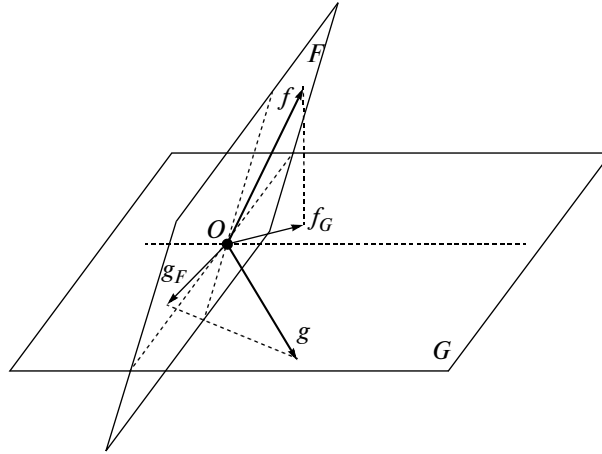


Рис. 1. Геометрическая иллюстрация морфологического проецирования

Для любого изображения $g(x, y) \in L^2(\Omega)$ может быть определена проекция на форму F :

$$g_F(x, y) = P_F g(x, y) = \sum_{i=1}^n g_{Fi} \chi_{Fi}(x, y),$$

$$g_{Fi} = \frac{(\chi_{Fi}, g)}{\|\chi_{Fi}\|^2}, \quad i = \overline{1, n}.$$

Здесь P_F — оператор проекции или проектор на F . Следует отметить, что с точки зрения морфологии Пытьева описание формы изображения как разбиения \mathbf{F} , линейного подпространства F и проектора P_F — эквивалентны и взаимозаменяемы.

Для дальнейших построений важно также, что пытьевские формы образуют алгебраическую структуру типа “решетка”, в которой для любых двух форм F и G можно указать форму более сложную $F \wedge G$ и менее сложную $F \vee G$. Более сложные формы получаются из менее сложных разбиением, а менее сложные из более сложных — слиянием областей.

Морфологическое сравнение изображений $f(x, y)$ и $g(x, y)$ в рамках морфологии Пытьева осуществляется при помощи морфологических квазирасстояний

$$d_M(g, F) = \|g - P_F g\|,$$

$$d_M(f, G) = \|f - P_G f\|$$

и нормированных морфологических коэффициентов корреляции Пытьева:

$$K_M(g, F) = \frac{\|P_F g\|}{\|g\|},$$

$$K_M(f, G) = \frac{\|P_G f\|}{\|f\|}.$$

Здесь в первом случае изображение g сравнивается с формой (множеством изображений) F , а во втором случае, наоборот, изображение f сравнивается с формой G . Легко убедиться, что в общем случае $K_M(g, F) \neq K_M(f, G)$.

Часто на практике используется не сам коэффициент K_M , а его квадрат K_M^2 , который имеет смысл статистического коэффициента детерминации, описывающего относительную долю дисперсии зависимой переменной $g(x, y)$, объясняемую моделью линейной регрессии F (1.1) по объясняющим переменным $\chi_{Fi}(x, y)$.

На рис. 1 показана геометрическая иллюстрация идей морфологического подхода (трехмерные векторы соответствуют изображениям, а плоскости — формам). Схема морфологического проецирования включает образы (изображения) f и g , формы F и G и проекции f_G и g_F . При этом

$$K(f, g) = \frac{(f, g)}{\|f\| \|g\|} = \cos(f \wedge g),$$

$$K_M(g, F) = \frac{\|P_F g\|}{\|g\|} = \cos(g \wedge g_F),$$

$$K_M(f, G) = \frac{\|P_G f\|}{\|f\|} = \cos(f \wedge f_G),$$

где $(f \wedge g)$ – угол между векторами f и g в рассматриваемом линейном пространстве.

Из рис. 1, в частности, сразу становится ясно, что морфологический коэффициент корреляции изображений никогда не может быть меньше их линейного нормированного коэффициента корреляции. Этот же рисунок позволяет ввести и основную идею сравнения мозаичных форм, которая будет далее исследоваться. Раз мозаичные формы F и G представляют собой линейные подпространства (гиперплоскости) в пространстве изображений, то между ними также имеется соответствующий “стереометрический угол” $F \wedge G$, который можно измерить и который соответственно должен характеризовать их сходство/различие. Идея достаточно очевидна и наглядна, однако при реализации в многомерных пространствах для достаточно сложных форм оценка таких “углов” может осуществляться разными способами, известными в математике.

Рассмотрим последовательно несколько таких подходов.

2. Морфологическое сравнение форм на основе поиска критических точек морфологического коэффициента корреляции. Известно, что морфологический коэффициент корреляции является критической точкой (экстремумом) нормированного линейного коэффициента корреляции. Именно это позволяет избавиться от учета особенностей яркости одного из изображений. Действительно, пусть имеется два кусочно-постоянных изображения типа (1.1) с различной формой, т.е. различной топологией разбиения кадра:

$$f(x, y) = \sum_{i=1}^n f_i \chi_{F_i}(x, y),$$

$$g(x, y) = \sum_{j=1}^m g_j \chi_{G_j}(x, y),$$

тогда

$$K(f, g) = \frac{\left(\sum_{i=1}^n f_i \chi_{F_i}(x, y), \sum_{j=1}^m g_j \chi_{G_j}(x, y) \right)}{\left(\sum_{i=1}^n f_i^2 \mu_i \right)^{1/2} \left(\sum_{j=1}^m g_j^2 \nu_j \right)^{1/2}} = \frac{\sum_{i=1}^n \sum_{j=1}^m f_i g_j (\chi_{F_i}(x, y), \chi_{G_j}(x, y))}{\left(\sum_{i=1}^n f_i^2 \mu_i \right)^{1/2} \left(\sum_{j=1}^m g_j^2 \nu_j \right)^{1/2}},$$

где $\mu_i = \|\chi_{F_i}(x, y)\|$ и $\nu_j = \|\chi_{G_j}(x, y)\|$ – площади областей разбиения для f и g соответственно.

В векторной форме это выражение примет вид

$$K^2 = \frac{(\mathbf{f}^T \mathbf{S} \mathbf{g})^2}{(\mathbf{f}^T \boldsymbol{\mu} \mathbf{f})(\mathbf{g}^T \mathbf{v} \mathbf{g})},$$

где матрицы представляют собой

$$\mathbf{S} = [s_{ij}]_{m \times n}, \quad s_{ij} = (\chi_{F_i}, \chi_{G_j}),$$

$$\boldsymbol{\mu} = \text{diag}(\mu_1, \dots, \mu_n), \quad \mathbf{v} = \text{diag}(\nu_1, \dots, \nu_m).$$

Введем два ортонормальных базиса $\mathbf{u} = (u_1, \dots, u_n)^T$ и $\mathbf{v} = (v_1, \dots, v_m)^T$:

$$u_i(x, y) = \frac{\chi_{F_i}(x, y)}{\mu_i^{1/2}}, \quad v_j(x, y) = \frac{\chi_{G_j}(x, y)}{\nu_j^{1/2}}.$$

Записывая изображения в этих базисах

$$f(x, y) = \sum_{i=1}^n f'_i u_i(x, y),$$

$$g(x, y) = \sum_{j=1}^m g'_j v_j(x, y)$$

и масштабируя обозначения для векторов яркостей, можно получить наиболее компактное выражение для коэффициента корреляции

$$K^2 = \frac{(\mathbf{f}^T \mathbf{M} \mathbf{g})^2}{(\mathbf{f}^T \mathbf{f})(\mathbf{g}^T \mathbf{g})}, \quad (2.1)$$

где

$$\mathbf{M} = [M_{ij}]_{m \times n}, \quad M_{ij} = s_{ij} \mu_i^{-1/2} \nu_j^{-1/2} - \text{нормированная площадь пересечения } F_i \text{ и } G_j,$$

$$\mathbf{f} = (f'_1, \dots, f'_n)^T, \quad f'_i = f \mu_i^{1/2},$$

$$\mathbf{g} = (g'_1, \dots, g'_m)^T, \quad g'_j = g \nu_j^{1/2}.$$

Из теории матриц хорошо известно, что $|K| \leq 1$, и при условии нормированной энергии изображений

$$\mathbf{f}^T \mathbf{f} = 1, \quad \mathbf{g}^T \mathbf{g} = 1$$

числитель выражения (2.1) имеет $n = \text{rang} \mathbf{M}$ критических значений, задаваемых собственными числами матрицы $\mathbf{M}^T \mathbf{M}$.

Ортонормальные базисы \mathbf{u} и \mathbf{v} порождают два линейных подпространства, которые в морфологической теории Пытьева называются формами F и G изображений f и g соответственно:

$$F = \text{span}\{u_1, \dots, u_n\}, \quad G = \text{span}\{v_1, \dots, v_m\}.$$

Обозначим через P_F и P_G операторы ортогонального проецирования на F и G соответственно. В случае мозаичных изображений

$$K_M^2 = \frac{(P_F g, P_F g)}{\|g\|^2} = (P_F g^*, P_F g^*),$$

где $g^* = \frac{g}{\|g\|}$ – центрированное и нормированное изображение g . Отсюда, представляя квадрат нормы проекции как скалярное произведение проекции на саму себя и учитывая ортонормированность базиса \mathbf{u} , получим

$$\begin{aligned} K_M^2 &= \left(\sum_{i=1}^n (g(x, y), u_i(x, y)) u_i(x, y), \sum_{i=1}^n (g(x, y), u_i(x, y)) u_i(x, y) \right) = \\ &= \sum_{i=1}^n (g(x, y), u_i(x, y))^2 \|u_i(x, y)\|^2 = \sum_{i=1}^n (g(x, y), u_i(x, y))^2. \end{aligned}$$

Затем подставим выражение для $g(x, y)$ как мозаичной функции

$$\sum_{i=1}^n (g, u_i(x, y))^2 = \sum_{i=1}^n \left(\sum_{j=1}^m (g_j v_j(x, y), u_i(x, y)) \right)^2 = \sum_{i=1}^n \left(\sum_{j=1}^m g_j (v_j(x, y), u_i(x, y)) \right)^2 = \sum_{i=1}^n \left(\sum_{j=1}^m g_j M_{ij} \right)^2$$

и выразим квадрат морфологического коэффициента корреляции в матричной форме

$$K_M^2 = \sum_{i=1}^n \left(\sum_{j=1}^m g_j M_{ij} \right)^2 = \mathbf{g}^T \mathbf{M}^T \mathbf{M} \mathbf{g}.$$

Учитывая свойства оператора проецирования, запишем условие ортогональности

$$(f, g - P_F g) = 0,$$

откуда

$$(f, g) = (f, P_F g) \leq \|f\| \|P_F g\|,$$

следовательно

$$K = \frac{(f, g)}{\|f\| \|g\|} \leq \frac{\|f\| \|P_F g\|}{\|f\| \|g\|} = \frac{\|P_F g\|}{\|g\|} = K_M. \quad (2.2)$$

Из (2.2) можно получить достаточное условие равенства обычного и морфологического коэффициентов корреляции, а именно требование коллинеарности векторов f и $P_F g$, т.е. для случая мозаичных изображений:

$$P_F g = \alpha f \Rightarrow (f, g) = (f, P_F g) = (f, \alpha f) = \alpha \|f\|^2 = \|f\| \|P_F g\| \Rightarrow K = K_M$$

или покомпонентно

$$(g, u_i) u_i = \alpha u_i, \quad i = \overline{1, n} \Rightarrow (g, u_i) = \alpha, \quad i = \overline{1, n}.$$

Дальнейшее стремление отказаться от использования яркостной информации при сравнении мозаичных изображений приводит к задаче поиска критических (экстремальных) точек морфологического коэффициента корреляции:

$$\min_{g \in G} K_M(g, F) = \min_{g \in G} \frac{\|P_F g\|}{\|g\|} = \min_{g^* \in G, \|g^*\|=1} \|P_F g^*\|.$$

В силу неотрицательности морфологического коэффициента корреляции и монотонно возрастающего характера квадратичной функции можно перейти к эквивалентной задаче отыскания

$$\min_{g \in G} K_M^2(g, F) = \min_{g^* \in G, \|g^*\|=1} \|P_F g^*\|^2. \quad (2.3)$$

Эта задача может быть решена методом множителей Лагранжа. Пусть

$$g^*(x, y) = \sum_{j=1}^m \beta_j v_j(x, y),$$

тогда, вводя вектор $\beta = (\beta_1, \dots, \beta_m)^T$ и учитывая свойства ортонормированного базиса v , условие нормировки изображения g^* можно записать в виде

$$\|g^*\|^2 = \left(\sum_{j=1}^m \beta_j v_j, \sum_{j=1}^m \beta_j v_j \right) = \beta^T \beta = 1.$$

При этом

$$\begin{aligned} \|P_F g^*\|^2 &= \left(\sum_{i=1}^n (g^*, u_i(x, y)) u_i(x, y), \sum_{i=1}^n (g^*, u_i(x, y)) u_i(x, y) \right) = \\ &= \sum_{i=1}^n (g^*, u_i(x, y))^2 = \sum_{i=1}^n \left(\sum_{j=1}^m \beta_j v_j(x, y), u_i(x, y) \right)^2 = \sum_{i=1}^n \left(\sum_{j=1}^m \beta_j M_{ij} \right)^2 = \beta^T M^T M \beta. \end{aligned}$$

С учетом этого задача условной оптимизации (2.3) принимает вид

$$\min_{\beta \in \{\beta^T \beta = 1\}} \beta^T M^T M \beta,$$

откуда, вводя множитель λ , получаем функцию Лагранжа

$$L(\beta, \lambda) = \beta^T M^T M \beta + \lambda(1 - \beta^T \beta). \quad (2.4)$$

Дифференцируя (2.4) по β и приравнявая векторную производную нулю, получим

$$2\mathbf{M}^T\mathbf{M}\beta - 2\lambda\beta = 0,$$

откуда сразу следует, что

$$\mathbf{M}^T\mathbf{M}\beta = \lambda\beta, \quad (2.5)$$

т.е. искомый экстремум выражения (2.3), а значит, и морфологического коэффициента корреляции достигается при векторах β_i , которые являются собственными векторами матрицы $\mathbf{M}^T\mathbf{M}$, а соответствующие им значения множителя Лагранжа совпадают с собственными числами λ_i этой матрицы.

Матрица вида $\mathbf{M}^T\mathbf{M}$ обладает, как известно, свойствами симметричности и диагонализруемости, а все ее собственные числа вещественны и неотрицательны, что следует из условия

$$\beta^T\mathbf{M}^T\mathbf{M}\beta = (\mathbf{M}\beta, \mathbf{M}\beta) \geq 0.$$

Из (2.5) вытекает формула

$$\det \mathbf{M}^T\mathbf{M} = \prod_{i=1}^n \lambda_{F_i},$$

где λ_{F_i} являются решениями характеристического уравнения

$$\det(\mathbf{M}^T\mathbf{M} - \lambda\mathbf{I}) = 0,$$

соответствующими собственным векторам β_i . При этом наименьшему собственному числу

$$\lambda_{F \min} = \min_{i=1, n} \lambda_{F_i}$$

соответствует собственный вектор β_{\min} .

Учитывая, что, согласно (2.5), экстремум K_M^2 достигается при векторах β_i , которые являются собственными векторами матрицы $\mathbf{M}^T\mathbf{M}$, подставляя β_{\min} в (2.3) и используя свойства собственного вектора, а также равенство его нормы 1, окончательно получаем

$$\min_{g \in G} K_M^2(g, F) = \min_{g^* \in G, \|g^*\|=1} \|P_F g^*\|^2 = \beta_{\min}^T \mathbf{M}^T \mathbf{M} \beta_{\min} = \beta_{\min}^T \lambda_{F \min} \beta_{\min} = \lambda_{F \min} \beta_{\min}^T \beta_{\min} = \lambda_{F \min}.$$

Поскольку собственные числа вещественны и неотрицательны, можно определить морфологический коэффициент корреляции форм G и F как

$$K_{MM}(G, F) = \min_{g \in G} K_M(g, F) = \lambda_{F \min}^{1/2}, \quad (2.6)$$

т.е. корень из минимального собственного числа матрицы $\mathbf{M}^T\mathbf{M}$. С практической точки зрения важно то, что $K_{MM}(G, F)$ вычисляется с использованием информации, которая полностью сосредоточена в матрице \mathbf{M} , определяемой разбиениями кадра \mathbf{F} и \mathbf{G} и никак не зависящей от распределений яркости обоих изображений. Таким образом, мера сходства форм K_{MM} является полностью инвариантной к любым взаимнооднозначным яркостным преобразованиям обоих сравниваемых изображений.

Заметим, что поскольку, как мы показали выше,

$$K_M(g, F) = \max_{f \in F} K(g, f),$$

то из (2.6) непосредственно следует

$$K_{MM}(G, F) = \min_{g \in G} K_M(g, F) = \min_{g \in G} \max_{f \in F} K(g, f).$$

Другими словами, сравнение форм осуществляется здесь путем решения минимаксной задачи оценивания степени близости по форме двух мозаичных изображений, причем базовой мерой близости двух изображений является именно нормированный линейный коэффициент корреляции. Как известно, минимаксное оценивание используется в тех случаях, когда информация о вероятностном распределении неизвестных параметров (факторов) отсутствует или недостовер-

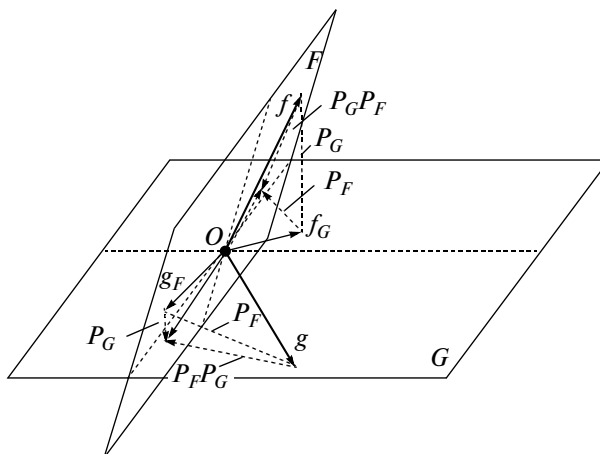


Рис. 2. Геометрическая иллюстрация суперпозиции проекторов на сравниваемые формы

на. В данном случае такими неизвестными параметрами формы являются распределения значений яркости $\{f_1, \dots, f_n\}$ и $\{g_1, \dots, g_m\}$. Если мы введем какие-либо предположения о характере их распределения, то получим иную статистическую оценку корреляции форм F и G . Такие оценки, в частности, были предложены авторами в работе [7] на основе среднеквадратического усреднения значений морфологического коэффициента корреляции $K_M(g, F)$ по ансамблю изображений $g \in G$. Такой путь построения меры сходства форм представляется достаточно перспективным, но в данной работе мы ограничимся рассмотрением чисто геометрических соображений без привлечения какой-либо априорной вероятностной информации.

Однако в силу асимметрии морфологического коэффициента Пытьева наряду с

$$K_{MM}(G, F) = \min_{g \in G} \max_{f \in F} K(g, f)$$

следует рассмотреть также и

$$K_{MM}(F, G) = \min_{f \in F} \max_{g \in G} K(f, g),$$

причем вычисление $K_{MM}(F, G)$ аналогичным образом сводится к поиску собственных чисел матрицы MM^T .

На основе этих двух асимметричных мер сходства симметричная мера сходства для сравнения форм может быть окончательно определена как взаимный морфологический коэффициент корреляции форм

$$K_{MM}^*(F, G) = \min\{K_{MM}(G, F), K_{MM}(F, G)\}, \tag{2.7}$$

который как раз и оценивает косинус искомого нами “угла” между формами F и G как линейными подпространствами векторного пространства изображений.

3. Морфологическое сравнение форм на основе анализа свойств суперпозиции соответствующих им проекторов. Возможен и другой, путь получения этого результата, основанный, скорее, на наглядных рассуждениях геометрического толка.

Рассмотрим понятия независимости форм изображений и ранга корреляции форм изображений, введенные Пытьевым [2]. Под независимыми формами понимаются такие две формы F и G , для которых любое центрированное изображение $f \in F$ имеет нулевую проекцию на G и наоборот. В рамках нашей геометрической иллюстрации (рис. 1) это подразумевает, что подпространства (плоскости) F и G ортогональны.

Понятие ранга корреляции связано с исследованием отображений $P_G P_F$ и $P_F P_G$, представляющих собой суперпозицию проекторов, соответствующих исследуемым формам (рис. 2).

Результатом действия оператора $P_G P_F$ на образ $f \in F$ является $P_G P_F f = f_{GF} \in F$. При повторном применении оператора $P_G P_F$ получим $P_G P_F P_G P_F f = f_{GFGF} \in F$. При этом в общем случае, очевидно, $f \neq f_{GF}$ и $f_{GF} \neq f_{GFGF}$. Однако имеются важные частные случаи, когда между f и f_{GF} существует более

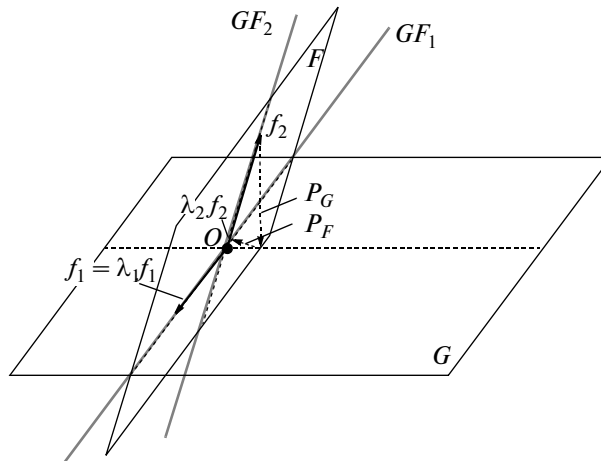


Рис. 3. Геометрическая иллюстрация собственных векторов (собственных образов) и собственных подпространств (собственных подформ) оператора $P_G P_F$

определенная связь. Это случаи собственных векторов преобразования $P_G P_F$ (рис. 3). При этом, как следует из определения собственных векторов,

$$P_G P_F f = \lambda_{FG} f,$$

где λ_{FG} – собственное значение, а f – собственный вектор оператора $P_G P_F$.

Аналогичным образом можно определить и собственные элементы λ_{GF} преобразования $P_F P_G$ (рис. 4). В нашем примере (геометрической иллюстрации) формы изображаются плоскостями, поэтому имеется только два собственных числа: $\lambda_{FG1}=1$ и $\lambda_{FG2} < 1$.

Первое собственное число в обоих случаях равно 1 и соответствует тривиальному собственному подпространству $GF_1 = FG_1 = F \vee G$, все элементы которого, естественно, сохраняются как при проецировании на F , так и при проецировании на G . Второе собственное число соответствует нетривиальному собственному подпространству, элементы которого при проецировании укорачиваются в λ_2 раз. При этом, как видно из рис. 3, данное собственное число λ_{FG} определяет искомый стереометрический угол ($F \wedge G$) между плоскостями F и G :

$$\lambda_{FG} = \frac{\|f_{GF}\|}{\|f\|} = \left(\frac{\|f_{GF}\|}{\|f_G\|} \right) \left(\frac{\|f_G\|}{\|f\|} \right) = \cos(F \wedge G) \cos(F \wedge G) = \cos^2(F \wedge G).$$

Рассмотрим простой пример (рис. 5). Переменные x, y, z обозначают здесь значения яркости областей централизованного мозаичного изображения (с нулевым средним). Для пары форм, по-

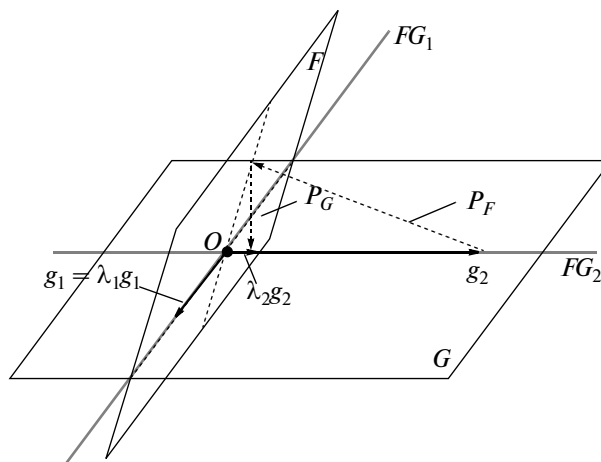


Рис. 4. Геометрическая иллюстрация собственных векторов (собственных образов) и собственных подпространств (собственных подформ) оператора $P_F P_G$

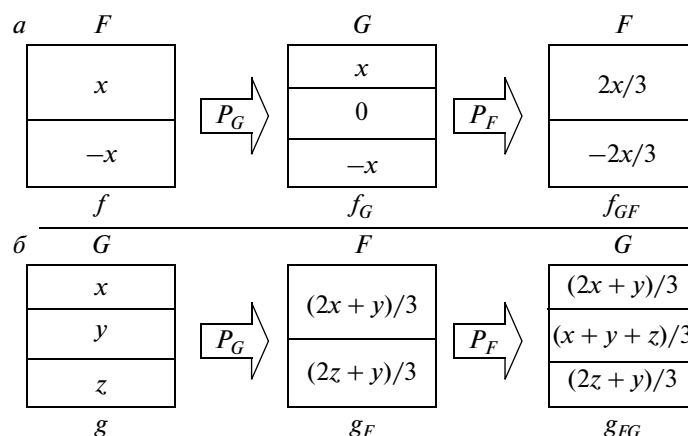


Рис. 5. Пример суперпозиции проекторов для простых форм-разбиений: $\lambda_{GF} = \|f_{GF}\|/\|f\| = 2/3$

казанных на рис. 5, собственное число $\lambda_{FG} = \lambda_{GF} = 2/3$. Данный пример позволяет также наглядно представить и собственное подпространство исследуемых операторов. Так, в случае исследования оператора $P_F P_G$ (рис. 5, б) собственное подпространство, соответствующее $\lambda_{GF} = 2/3$, состоит из изображений, удовлетворяющих условию $y = (x + z)/2$, что с учетом центрирования по яркости ведет к форме вида $\{x, y = 0, z = -x\}$, как и на рис. 5, а. Поэтому λ_{FG} и λ_{GF} здесь равны.

Однако следующий пример (рис. 6) показывает, что не все так просто в мире мозаичных форм. На рис. 6, б и с показаны два различных решения, соответствующие двум различным нетривиальным собственным числам: $\lambda_1 = 2/3$ и $\lambda_2 = 1/4$.

Именно этот эффект описывается понятием ранга корреляции [2]: рангом корреляции формы F с формой G называется размерность пространства значений оператора $P_G P_F$. Если ранг корреляции больше 1, то существует несколько различных нетривиальных собственных чисел. Поэтому в качестве оценки “стереометрического угла” между формами следует выбирать минимальное собственное число:

$$\cos^2(F \wedge G) = \min_{i=1,n} \lambda_{GF_i}$$

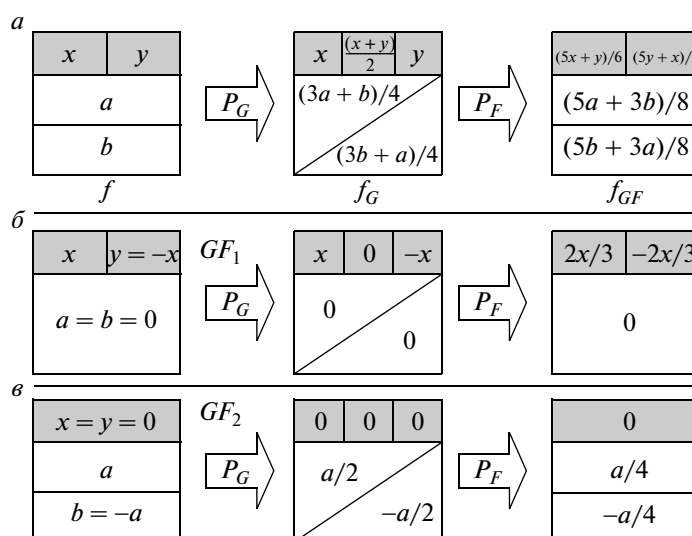


Рис. 6. Пример суперпозиции проекторов для более сложных форм-разбиений

Если теперь использовать для описания мозаичных изображений обозначения, введенные в предыдущем разделе, то из выражения

$$P_F P_G g = \lambda g$$

непосредственно следует уравнение

$$\mathbf{M}^T \mathbf{M} \mathbf{b} = \lambda \mathbf{b},$$

что полностью совпадает с выражением, полученным выше, т.е.

$$K_{MM}(G, F) = \lambda_{F \min}^{1/2} = \lambda_{GF \min}^{1/2}.$$

Таким образом, меры сходства мозаичных форм, основанные на вычислении собственных чисел суперпозиции операторов проецирования, полностью совпадают с мерами сходства, полученными, исходя из соображений поиска критических точек коэффициента морфологической корреляции.

4. Морфологическое сравнение форм на основе оценки раствора угла между многообразиями. Оба предыдущих подхода к сравнению мозаичных форм основывались на идее обобщения морфологического коэффициента корреляции Пытьева. Воспользуемся иным математическим аппаратом определения угловых мер между подпространствами гильбертова пространства. Рассмотрим известный подход [8], связанный с определением раствора $\Theta(M_1, M_2)$ между многообразиями M_1 и M_2 :

$$\Theta(M_1, M_2) = \max \left\{ \sup_{f \in M_2, \|f\|=1} \|(I - P_1)f\|, \sup_{g \in M_1, \|g\|=1} \|(I - P_2)g\| \right\}. \quad (4.1)$$

Здесь P_1 и P_2 – операторы проецирования на M_1 и M_2 соответственно, I – тождественный оператор, а величину раствора $\Theta(M_1, M_2)$ можно рассматривать как синус “угла” θ между многообразиями M_1 и M_2 , имея в виду, что чем больше величина раствора, тем больше указанный “угол”. Попробуем вычислить минимальную величину косинуса этого “угла”. Учитывая, что

$$\cos^2 \theta_{\min} = 1 - \sin^2 \theta_{\min} = 1 - \Theta(M_1, M_2)^2,$$

$$\|(I - P_1)f\|^2 + \|P_1 f\|^2 = \|f\|^2,$$

$$\|f\|^2 = 1,$$

подставляем (4.1) и получаем

$$\cos^2 \theta_{\min} = \min \left\{ \inf_{f \in M_2, \|f\|=1} \|P_1 f\|, \inf_{g \in M_1, \|g\|=1} \|P_2 g\| \right\}. \quad (4.2)$$

Сравнивая (4.2) с (2.3), (2.6) и (2.7), легко убедиться, что при $F = M_1$ и $G = M_2$ $\cos^2 \theta_{\min}$ есть ничто иное как квадрат симметричного морфологического коэффициента корреляции форм $K_{MM}^*(F, G)$, введенного выше. Он представляет собой комбинацию (минимум) критических точек операторов проецирования и вычисляется по формуле

$$K_{MM}^{*2}(F, G) = \cos^2 \theta_{\min} = \min \left\{ \min_{i=\overline{1, n}} \lambda_i(\mathbf{M}^T \mathbf{M}), \min_{j=\overline{1, m}} \lambda_j(\mathbf{M} \mathbf{M}^T) \right\}.$$

В частности, в простейшем частном случае, когда $\mathbf{M} = \mathbf{M}^T$,

$$K_{MM}^*(F, G) = \cos \theta_{\min} = \min_{i=\overline{1, n}} \lambda_i(\mathbf{M}).$$

Таким образом, между введенным выше понятием морфологического коэффициента корреляции мозаичных форм и понятием раствора угла между мозаичными формами как линейными многообразиями существует взаимно-однозначное соответствие.

5. Морфологическое сравнение форм на основе метрики Хаусдорфа между соответствующими многообразиями на единичной гиперсфере. Рассмотрим еще один способ сопоставления форм, основанный на переходе в пространство всех центрированных нормированных изображений (т.е. на проецировании пространства изображений на единичную гиперсферу).

Как известно, нормированный коэффициент линейной корреляции

$$K(f, g) = \frac{(f - f_0, g - g_0)}{\|f - f_0\| \|g - g_0\|}$$

есть скалярное произведение нормированных центрированных изображений:

$$f^*(x, y) = \frac{f(x, y) - f_0}{\|f(x, y) - f_0\|},$$

$$g^*(x, y) = \frac{g(x, y) - g_0}{\|g(x, y) - g_0\|},$$

$$K(f, g) = (f^*, g^*).$$

Рассмотрим проекцию F^* центрированной формы

$$F = \left\{ f(x, y) = \sum_{i=1}^n f_i \chi_{F_i}(x, y) : \sum_{i=1}^n f_i \mu_i = 0, f \in R^n \right\}$$

на единичную гиперсферу:

$$F^* = \left\{ f^*(x, y) \in F : \|f^*\|^2 = 1, \sum_{i=1}^n f_i \mu_i = 0 \right\}.$$

Множество F^* будем называть нормированной центрированной формой.

Легко показать, что поскольку для любой пары f и $f^* = \frac{f}{\|f\|}$ соответствующих изображений $f \in F \Leftrightarrow f^* \in F^*$, соответствие форм F и F^* является взаимно однозначным. Поэтому метрика, введенная на нормированных центрированных формах, может быть однозначно отображена на исходные формы. Между тем для точек гиперсферы (центрированных нормированных изображений) естественным образом вводится геодезическая метрика, причем геодезическое расстояние на единичной гиперсфере равно длине кратчайшей дуги единичной окружности, соединяющей два центрированных нормированных изображения f^* и g^* , а эта дуга, в свою очередь, равна углу между соответствующими центрированными изображениями f и g , характеризующему их нормированным коэффициентом корреляции:

$$d(f^*, g^*) = \arccos(f^*, g^*) = \arccos K(f, g).$$

Примем эту метрику за базовую и построим на основе ее расстояние Хаусдорфа [9].

Пусть X и Y – два непустых компактных подмножества метрического пространства M . Тогда расстояние Хаусдорфа $d_H(X, Y)$ между X и Y есть минимальное число r , такое, что замкнутая r -окрестность X содержит Y и также замкнутая r -окрестность Y содержит X . Другими словами, если $d(x, y)$ – базовое расстояние между точками x и y в M , то

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}.$$

Учитывая это, легко определить расстояние Хаусдорфа между проекциями форм

$$d_H(F^*, G^*) = \max \left\{ \sup_{f^* \in F^*} \inf_{g^* \in G^*} d(f^*, g^*), \sup_{g^* \in G^*} \inf_{f^* \in F^*} d(f^*, g^*) \right\}$$

и затем отобразить его на множество самих мозаичных форм:

$$d_H(F, G) = \max \left\{ \sup_{f \in F} \inf_{g \in G} \arccos K(f, g), \sup_{g \in G} \inf_{f \in F} \arccos K(f, g) \right\}. \tag{5.1}$$

Рассмотрим связь между расстоянием Хаусдорфа $d_H(F, G)$ (5.1) и введенным ранее морфологическим коэффициентом корреляции форм $K_{MM}^*(F, G)$ (2.7). Действительно, из (2.7) непосредственно следует, что

$$K_{MM}^*(F, G) = \min \{K_{MM}(G, F), K_{MM}(F, G)\} \Rightarrow \\ \Rightarrow K_{MM}^*(F, G) = \min \left\{ \min_{g \in G} \max_{f \in F} K(g, f), \min_{g \in G} \max_{f \in F} K(g, f) \right\}.$$

Отсюда

$$d_H(F, G) = \arccos K_{MM}^*(F, G) = \arccos \left(\min \left\{ \min_{i=1, n} \lambda_i(\mathbf{M}^T \mathbf{M}), \min_{j=1, m} \lambda_j(\mathbf{M} \mathbf{M}^T) \right\} \right)^{1/2}.$$

Другими словами, и эти показатели оказались связаны взаимно-однозначной зависимостью.

Таким образом, мы получили в точном смысле метрическое пространство мозаичных форм, причем указали явный алгоритм вычисления расстояния Хаусдорфа $d_H(F, G)$ между двумя мозаичными формами F и G , основанный на вычислении минимального собственного числа характеристической матрицы $\mathbf{M}^T \mathbf{M}$.

Следует отметить, что различные метрики сравнения мозаичных форм предлагались и ранее, например в [10, 11], однако они вводились эвристически и не были связаны с математическим аппаратом морфологии Пытьева ничем, кроме собственно предмета рассмотрения (мозаичных форм). Данный же подход целиком следует методологии сравнения изображений в морфологическом анализе Пытьева.

Заключение. В статье рассмотрена задача определения процедуры морфологического сравнения мозаичных форм, позволяющей осуществлять непосредственную оценку сходства/различия форм между собой и при этом согласованной с пытьевской морфологией в части сравнения изображений с формами. Исследованы четыре различных способа построения такой процедуры:

- на основе поиска критических точек морфологического коэффициента корреляции;
- на основе анализа свойств суперпозиции соответствующих им проекторов;
- на основе оценки раствора угла между многообразиями;
- на основе метрики Хаусдорфа между соответствующими многообразиями на единичной гиперсфере.

Суперпозиционный подход к сравнению форм ранее уже использовался Пытьевым, но лишь для определения понятия ранга корреляции форм [2], а не для оценки меры их сходства. Остальные подходы, насколько нам известно, применительно к данной задаче анализа изображений предложены и рассмотрены впервые.

Для всех четырех подходов независимо получены меры сходства или различия мозаичных форм. При этом меры сходства, полученные с использованием первых двух подходов (коэффициенты морфологической корреляции форм), совпали, а два других подхода привели к формированию мер отличия форм, связанных с коэффициентом морфологической корреляции форм взаимно-однозначными зависимостями. Такой результат является ожидаемым, поскольку с точки зрения геометрического смысла все эти подходы эксплуатируют одно и то же понятие “угла” между мозаичными формами F и G как многообразиями в гильбертовом пространстве изображений. Поэтому использование той или иной конкретной меры сходства (коэффициент корреляции как косинус минимального “угла” между изображениями — элементами многообразий) или различия (метрика Хаусдорфа как величина минимального “угла”, раствор как синус минимального “угла”) зависит от специфики решаемой задачи и удобства организации вычислений. В смысле же характеристик, достижимых в задачах распознавания, эти показатели полностью равноправны, поскольку никакой дополнительной информации ни один из них по отношению к другим не содержит.

Дальнейшая разработка морфологических инструментов сравнения форм может, по-видимому, проводиться в следующих основных направлениях:

- теоретическое исследование топологических и геометрических свойств введенного метрического пространства и их влияния на качество метрических алгоритмов распознавания образов;
- экспериментальное исследование свойств метрик сравнения форм в задачах привязки изображений различной физической природы;

расширение описанных подходов на другие классы форм, более содержательные или более инвариантные, чем простые разбиения кадра.

Здесь подход, основанный на введении метрики Хаусдорфа, представляется, возможно, наиболее перспективным, поскольку он позволяет естественным образом получать метрики для классов форм, не являющихся, в отличие от мозаичных, линейными подпространствами пространства изображений. Другой интересный путь обобщения данного подхода заключается в построении метрики Громова–Хаусдорфа [9], которая могла бы обеспечить сравнение мозаичных форм, инвариантное не только к яркостным, но и к определенным типам геометрических преобразований.

СПИСОК ЛИТЕРАТУРЫ

1. *Serra J.* Image Analysis and Mathematical Morphology, London: Acad. Press, 1982.
2. *Пытьев Ю.П., Чуличков А.И.* Методы морфологического анализа изображений. М.: Физматлит, 2010. 336 с.
3. *Визильтер Ю.В., Желтов С.Ю.* Проективные морфологии и их применение в структурном анализе цифровых изображений // Изв. РАН. ТиСУ. 2008. № 6. С. 113–128; *Vizilter Yu.V., Zheltov S.Yu.* Projective Morphologies and Their Application in Structural Analysis of Digital Images // J. Computer and Systems Sciences International. 2008. V. 47. № 6. P. 944–958.
4. *Визильтер Ю.В.* Обобщенная проективная морфология // Компьютерная оптика. 2008. Т. 32. № 4. С. 384–399.
5. *Визильтер Ю.В., Желтов С.Ю.* Использование проективных морфологий в задачах обнаружения и идентификации объектов на изображениях // Изв. РАН. ТиСУ. 2009. № 2. С. 125–138; *Vizilter Yu.V., Zheltov S.Yu.* The Use of Projective Morphologies for Object Detection and Identification in Images // J. Computer and Systems Sciences International. 2009. V. 48. № 2. P. 282–294.
6. *Визильтер Ю.В.* Конструирование операторов сегментации и сжатия данных на основе проективных морфологических разложений // Изв. РАН. ТиСУ. 2009. № 3. С. 89–104; *Vizilter Yu.V.* Construction of Data Segmentation and Data Compression Operators Based on Projective Morphological Decompositions // J. Computer and Systems Sciences International. V. 48. № 3. 2009. P. 415–429.
7. *Vizilter Y.V., Zheltov S.Y.* Geometrical Correlation and Matching of 2D Image Shapes // ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-3. 191–196. doi:10.5194/isprsannals-I-3-191-2012. 2012.
8. *Ахиезер Н.И., Глазман И.М.* Теория линейных операторов в гильбертовом пространстве. М.: Наука, Физматлит, 1966.
9. *Бураго Д.Ю., Бураго Ю.Д., Иванов С.В.* Курс метрической геометрии. Москва–Ижевск: Ин-т компьютерных исследований, 2004. 512 с. ISBN 5-93972-300-4.
10. *Визильтер Ю.В., Рубис А.Ю.* Метрическое пространство форм изображений // 9-я Междунар. конф. “Интеллектуализация обработки информации”. Черногория. г. Будва. М.: Торус Пресс, 2012. С. 406–410.
11. *Визильтер Ю.В., Рубис А.Ю., Горбачев В.С.* Реляционные модели формы изображений и метрики их сравнения // 9-я Междун. конф. “Интеллектуализация обработки информации”. Черногория. г. Будва. М.: Торус Пресс, 2012. С. 410–414.