

Распознавание типовых портретных изображений в задаче автоматической классификации

М.В. Смирнов

И.Н.Сивяков

ЗАО ИВА “TheWebProduction”, Санкт-Петербург,
m.Smirnov@thewebproduction.com

Аннотация

Решение задачи распознавания узкого класса объектов имеет тенденцию к образованию кластеров [8]. Успех кластеризации изображений в значительной степени основан на эвристических алгоритмах. Эвристики могут быть сведены к формальным алгебрам [3] и в частности к нечетким множествам. Суть предлагаемого решения основывается на совместном применении независимых признаков изображений, таких как, контурные, яркостные, коэффициенты Фурье, в контексте модели визуального восприятия. Эксперименты показали, что система работоспособна и уверенно классифицирует типовые портреты. Правильное решение, для произвольного набора изображений, получено в 74% случаях.

1. Введение

Постановка задачи: построение и исследование кластеров типовых портретных изображений (поясных фото, фото на документы) на основе совокупности независимых признаков изображений: контурных, яркостных и коэффициентов разложения в ряд Фурье. Основное внимание уделяется выбору признаков, их формализации и формированию на их основе кластеров. Известно, что выбор признаков является наиболее важной задачей при построении систем распознавания изображений. Если результаты измерений позволяют получить исчерпывающий набор

различительных признаков, то процедура распознавания не вызывает особых затруднений [8]. Примером может служить распознавание символов. Однако в большинстве случаев оценка различительных признаков оказывается исключительно трудным делом. Для решения данной проблемы применяется предварительная обработка входных данных, позволяющая не только повысить различительные свойства признаков, но и снизить размерность вектора признаков. С этой точки зрения, большого внимания заслуживает применение методов нечетких (под)множеств, позволяющих представить нечеткие эвристические признаки и процедуры на формальном уровне [3] и свести некорректные процедуры к формальным алгебрам [4]. Большое внимание в работах по распознаванию, с помощью нечетких множеств, уделяется классификации линейно-контурных элементов изображения, что не удивительно, учитывая значимость контурной информации. В работах [10,11] с помощью нечетких алгебр были исследованы возможности распознавания линейных объектов и объектов, описываемых кривыми линиями.

Другое интересное направление в классификации образов связано с понятием нейронных сетей [2]. Основное преимущество нейронных сетей заключается в наличии универсального механизма обучения, который избавляет от необходимости выбирать ключевые признаки, оценивать их значимость [9]. Судя по литературе, методы нейронных сетей получили наибольшее распространение в задаче распознавания личности по изображению лица. Входными характеристиками для распознавания в нейронных сетях, чаще всего, являются расстояния между специфическими участками лица (нос, рот, глаза) [12]. При этом, во многих публикациях отмечается, что выбор исходного представления входных данных (признаков) существенно влияет на принятие решения в нейронных сетях.

Предлагаемая нами система распознавания основывается на разделении входных изображений на два класса. Первый класс составляют искомые портретные изображения, а второй класс - все остальные изображения (фоновые). Признаковое пространство портретных изображений формируется на основе описания изображений с помощью контурных элементов изображения, коэффициентов усеченного разложения функции изображения в ряд Фурье, яркостных признаков. Среди всех контурных элементов изображения рассматриваются только те контуры, которые удовлетворяют структурному описанию края полуплоскости [7].

2. Идея исследования

Основная идея исследования базируется на совместном использовании независимых и разнохарактерных признаков, для которых объединяющим фактором являются принципы зрительного восприятия. Гипотеза о подчеркивании контуров зрительным анализатором широко используется во многих приложениях по обработке и распознаванию изображений. При этом, в основу метода выделения контурных элементов нами положено структурно-вероятностное описание изображения края полуплоскости [7]. Применение данного подхода обеспечило успешное распознавание краев полуплоскостей и оценку разрешающей способности по аэрокосмическим изображениям вне зависимости от условий съемки [13].

Структурное описание [7] края полуплоскости обеспечивает представление выделенных контуров в виде бинарных и целочисленных множеств. В свою очередь, использование свойств элементов множеств позволяет разбить их на классы и получить, в конечном счете, словесное описание изображений.

Меньшее распространение получила гипотеза о способности зрительной системы выполнять разложение в ряд Фурье [1,5]. Предполагается, что зрительная система разбивает изображение на фрагменты, соответствующие предполагаемым объектам и затем выполняет разложение в ряд Фурье участков изображения внутри фрагментов. Данная процедура получила название разложения Габора. Полученный набор коэффициентов сравнивается с набором эталонов, хранящихся в зрительной памяти. Таким образом, происходит распознавание объектов в зрительной системе. Кроме того, учет горизонтальной симметрии портретов позволяет сократить признаковое пространство коэффициентов Фурье, отказавшись от вычисления синусных составляющих в этом направлении.

Свойство горизонтальной симметрии портретных изображений, представляет собой одну из наиболее важных характеристик и используется для получения как контурных, так и яркостных признаков портретных изображений.

Контурным элементам изображения отвечают более высокочастотные компоненты пространственного спектра. А усеченному ряду Фурье отвечают низкие и средние частоты пространственного спектра изображения. В процессе принятия решения, если контурные признаки портретных изображений не

получают доминирующего значения, большой вклад дают коэффициенты разложения в ряд Фурье и наоборот.

3. Описание методов

В силу поставленной задачи разделения изображений на два класса и в силу тенденции портретных изображений к тесной группировке, целесообразно классифицировать образы с помощью функции расстояния. Формула линейной разделяющей функции для двух классов может быть записана в виде:

$$d(x) = \sum_{i=1}^N [x_i(M_{1i} - M_{2i}) - (M_{1i}^2 - M_{2i}^2)/2] / \lambda_i, \quad (1)$$

где $d(x)$ - разделяющая гиперплоскость, x - вектор входных признаков, M_{1i} и M_{2i} - математические ожидания (координаты центров кластеров) i -го признака для портретных и фоновых изображений, соответственно, N - число признаков, λ_i - нормирующий коэффициент. В частности, использовалось среднее средних:

$$\lambda_i = (M_{1i} + M_{2i})/2. \quad (2)$$

Процесс кластеризации начинается с выбора типичных представителей портретных (рис.1) и фоновых (рис.2) изображений. Затем кластеры пополняются исходя из условия:

$$M_{Ki} - \sqrt{D_{Ki}} \leq x_i \leq M_{Ki} + \sqrt{D_{Ki}}, \quad (3)$$

где K принимает значение 1 или 2, по числу классов, D_{Ki} - дисперсия i -го признака.



Рис. 1.



Рис. 2.

3.1. Описание контурных признаков

Для получения контурных признаков используется методика распознавания изображения края полуплоскости [7]. Методика основывается на вероятностном характере изменения яркости края полуплоскости в зависимости от изменений освещенности объектов съемки. Вероятность выделения наиболее резких перепадов яркости, превосходящих порог Π , определяется по формуле:

$$p = 2 \int_{\Pi}^{+\infty} P(\beta) d\beta,$$

где $P(\beta)$ - плотность вероятности перепадов яркости изображения. Если плотность $P(\beta)$ известна, то пороговый параметр Π является адаптивным и автоматически перестраивается в зависимости от изменения внешней освещенности.

Геометрическим представлением края полуплоскости является индикаторная линия резкого перепада яркости. Каждой индикаторной линии $L(\rho, \varphi)$ присваивается номер (идентификатор), определяется положение на плоскости изображения в полярных координатах (ρ, φ) , подсчитывается длина линии и геометрический центр тяжести. Совокупность идентификаторов индикаторных линий (подмножество W) образует множество целых чисел Z , а совокупность всех точек

индикаторных линий на изображении составляет бинарное множество B , свойством b которого является единичное значение яркости линий на нулевом фоне. Примеры выделения множества B для изображений на рис.1 и рис.2 представлены на рис.3 и рис.4.

Процедура выделения любого подмножества W , т.е. любой линии $L(\rho, \varphi)$, состоит в задании свойства значения идентификатора $W = \{w / w = z\}$.

3.1.1. Признак выпуклости индикаторных линий

Признак выпуклости характерен абрису головы человека. Из множества B выделяется подмножество Θ линий $L(\rho, \varphi)$,



Рис. 3.



Рис. 4.

обладающих свойством выпуклости в ограниченном диапазоне значений ρ :

$$\Theta = \{\rho / \rho_c < \rho_s, 0.3 \leq \rho < 1\},$$

где ρ_c – радиус вектор центра тяжести индикаторной линии, ρ_s – радиус вектор точки линии, ближайшей к центру тяжести. Признак выпуклости индикаторной линии задается следующим образом:

$$E = \ell(\rho_s - \rho_c) / \ell_0, \quad ,$$

где ℓ - длина линии подмножества Θ , ℓ_0 - длина всех индикаторных линий. Признак E принимает наибольшее значение

для круга. На рис. 5 представлен пример выделения выпуклых линий изображения на рис. 1.



Рис.5.

3.1.2. Признак случайности

Предпосылкой использования признака случайности V_ρ , как различительного, служит факт более упорядоченного (детерминированного) расположения индикаторных линий класса портретных изображений, относительно класса фоновых изображений. Мерой случайности может служить стандартное отклонение радиус-вектора. Признак V_ρ принимает наименьшее значение для круга. Таким образом, признак случайности оценивается как вариации радиус вектора:

$$V_\rho = \int [\rho(\varphi) - \bar{\rho}]^2 d\varphi,$$

где $\bar{\rho} = \int \rho(\varphi) d\varphi$ - среднее значение.

3.1.3. Признак симметрии

Горизонтальная симметрия портретных изображений, является, пожалуй, одним из наиболее очевидных и наиболее значимых признаков. Вместе с тем, условие горизонтальной симметрии, должно допускать наличие наклонов и сдвигов головы на портретах.

Признак симметрии определяется как корреляция точек индикаторных линий относительно центра горизонтальной симметрии $X_h = \rho_h$ всего изображения. Вычисление скользящей корреляции обеспечивает инвариантность к сдвигу, а применение полярных координат - инвариантность к повороту. Учитывая, что $B = \{b / b = 1\}$, коэффициент корреляции вычисляется как сумма совпадений симметричных точек:

$$C_h = \iint L(\rho - \rho_h, \varphi - \varphi_h) L(\rho + \rho_h, \varphi) d\rho_h d\varphi_h,$$

при этом линия симметрии для диапазона углов $\Delta\varphi$ определяется, как $X_h = \rho_h \cos \varphi$. На рис.6 представлен пример вычисления коэффициентов C_h для портретного и фонового изображения. Значения C_h нормируются на длину индикаторных линий. Структура отчетов корреляции C_h на плоскости изображения, представлена на рис. 7. На рис.8 показано зеркальное отображение совокупности линий $L(\rho - \rho_h, \varphi_1)$ относительно линии симметрии, а на рис.9 показано совмещение линий $L(\rho - \rho_h, \varphi_1)$ для угла φ_1 с линиями $L(\rho + \rho_h, 0)$ для угла $\varphi = 0$.

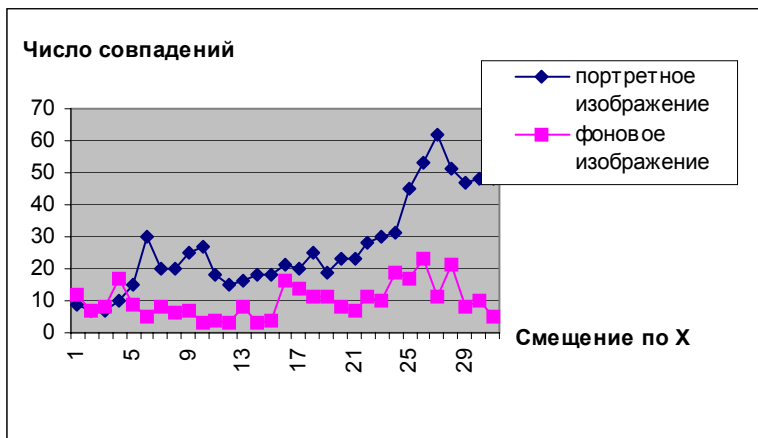


Рис.6.

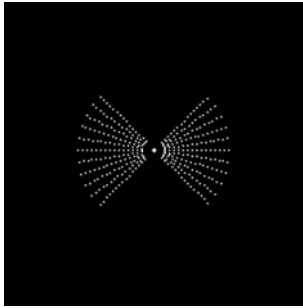


Рис. 7.



Рис. 8.

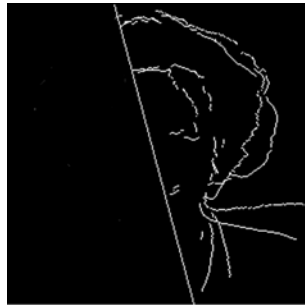


Рис. 9.

В качестве результирующего значения признака контурной симметрии принимается максимальное значение коэффициента корреляции C_h , нормированное на длину линий ℓ_0 .

3.1.4. Признак формы

Чаще всего признаком формы описываются плоские односвязные фигуры. В нашем случае “фигуры” состоят из разрозненных индикаторных линий, не имеющих четких границ. Если граница определена, с помощью некоторой меры, то изотропность фигуры оценивается разбросом значений радиус-вектора и длиной дуги, пропорциональной $\Delta\varphi$. В качестве меры разбиения будем использовать свойство близости индикаторных линий.

Таким образом, признак формы определяем для множеств индикаторных линий, обладающих свойством близости

(группировка по принципу “ближайший сосед”). Первоначально, из множества B группируются подмножества индикаторных линий R , удовлетворяющие условию близости

$$R = \{r/r \leq 3\sigma_\rho\} \quad , \quad (4)$$

где r - парные расстояния между всеми точками, составляющими множество индикаторных линий, σ_ρ - стандартное отклонение радиус-векторов ρ . Затем для каждого подмножества R вычисляется коэффициент формы $K_f = (\rho_{\max} - \rho_{\min})/\Delta\varphi$, где ρ_{\max} и ρ_{\min} - максимальное и минимальное значение радиус-вектора, $\Delta\varphi$ - диапазон полярного угла множества R . Коэффициент K_f принимает наименьшее значение для круга - самой изотропной фигуры. Условие близости (4) позволяет разбить индикаторные линии на подмножества (классифицировать их). На рис. 10 и на рис. 11 представлены два наиболее “мощных” класса из подмножеств R , в состав которых входит более 97% индикаторных линий.

Парные расстояния r образуют симметричную квадратную матрицу $r_{i,j} = r_{j,i}$ размерностью $\ell_0 \times \ell_0$, диагональные элементы которой $r_{k,k} = 0$.



Рис. 10.



Рис. 11.

Положение (столбец) i минимального расстояния r_{\min} в каждой строке j матрицы $r_{i,j}$, позволяет построить цепочку идентификаторов:

$$W_i \times W_j = \{(w_i; w_j)\} \quad (5),$$

которые связаны условием минимума расстояния (4), где w_i - значение рассматриваемого идентификатора, а w_j - значение ближайшего к нему идентификатора. Пары целых чисел w_i и w_j образуют замкнутые пространства значений, и поэтому могут быть разбиты на классы на формальном уровне.

3.2. Фурье-анализ изображений

Исходными данными для вычисления коэффициентов ряда Фурье являются значения яркости матрицы изображения $g_{k,j}$. Использование коэффициентов Фурье является одним из наиболее разработанных методов в обработке и распознавании изображений. Имеются данные [1,5,6], свидетельствующие, о том, что при анализе изображений зрительной системой, в некоторой форме, используются эти же принципы.

Слабым местом традиционного Фурье-анализа изображений является его чувствительность к поворотам и изменению масштаба. В нашем случае, масштабные различия и вариации поворотов, в значительной степени ограничены постановкой задачи и преодолены за счёт усреднения при создании эталонов.

Согласно [1], зрительная система для опознания элементарных объектов использует порядка 5×5 гармоник пространственного спектра изображения. С увеличением номеров гармоник их величина быстро падает. Значения гармоник вычислялись по следующим формулам:

$$\begin{aligned}
S_{im}^c &= \sum_k^{N_k} \text{Cos}\left(\frac{2\pi m}{N_k}\left(k - N_k/2\right)\right) \sum_j^{N_j} g_{kj} \cdot \text{Cos}\left(\frac{2\pi i}{N_j}\left(j - N_j/2\right)\right) \\
S_{im}^s &= \sum_k^{N_k} \text{Sin}\left(\frac{2\pi m}{N_k}\left(k - N_k/2\right)\right) \sum_j^{N_j} g_{kj} \cdot \text{Cos}\left(\frac{2\pi i}{N_j}\left(j - N_j/2\right)\right)
\end{aligned} \quad (6),$$

где S_{im}^c и S_{im}^s - косинусные и синусные гармоники, соответственно, $N_j \times N_k$ - размер изображения.

Синусные гармоники S_{im}^s вычислялись только по вертикали, поскольку асимметрия по горизонтали, как признак портрета, не является значимой. Косинусные гармоники S_{im}^c вычислялись для обоих направлений. Постоянная составляющая не рассматривалась. Как при создании эталонов, так и в процессе распознавания использовалось 64 гармоники по вертикали и 64 гармоники по горизонтали. Каждая гармоника рассматривалась как независимый признак. После вычисления коэффициентов Фурье, последние нормировались на экстремальные значения, в пределах данного пространственного спектра. Это позволило привести все спектры к одному контрасту. В противном случае, возможна инверсия значимой части спектра.

Эталонные спектры портретных изображений были получены путём усреднения по 27 изображениям, а эталонные спектры фоновых изображений по 22 изображениям. На рис.12 приведено изображение, полученное обратным преобразованием Фурье усреднённых гармоник эталонных портретных изображений.

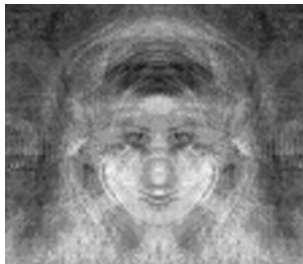


Рис. 12.

Для оценки положения классифицируемого изображения в пространстве признаков коэффициентов Фурье, относительно математических ожиданий классов портретных и фоновых изображений, использовалось выражение (7):

$$F = \sum_i \left(s_i (M_{in} - M_{i\phi}) - (M_{in}^2 + M_{i\phi}^2) \cdot \sigma_{i\phi}^2 / \sigma_{in}^2 \right) / \sigma_{i\phi}^2 / \sigma_{in}^2, \quad (7)$$

где M_{in} и $M_{i\phi}$ – математические ожидания i -ой гармоники s_i признака, соответственно для портретных и фоновых изображений, а σ_{in}^2 и $\sigma_{i\phi}^2$ – соответствующие дисперсии. Это выражение, в ходе проводившихся экспериментов, показало наилучшие результаты среди ряда исследованных нами статистик.

3.3. Оценка симметрии изображения по значениям яркости

В отличие от коэффициента контурной симметрии C_h (п.3.1.3), здесь оценивается горизонтальная симметрия значений яркости изображения $g_{k,j}$. Оценкой симметрии служит взаимная корреляция значений яркости $g_{k,j}$ относительно линии симметрии $X_h = \rho_h \cos \varphi$ в диапазоне углов $\Delta\varphi$:

$$K_h = \iint [g(\rho - \rho_h, \varphi - \varphi_h) - \bar{g}_{-h}] \times \\ \times [g(\rho + \rho_h, \varphi - \varphi_h) - \bar{g}_{+h}] d\rho_h d\varphi_h / Q_h.$$

Коэффициенты корреляции K_h являются центрированными и нормированными коэффициентами, где $\bar{g}_{\pm h}$ – средние яркости, а нормировка Q_h равна:

$$Q_h = \iint g_{-h}^2 g_{+h}^2 d\rho_h d\varphi_h,$$

где $g_{\mp h}^2$ - дисперсии яркости. Структура отсчетов K_h , на плоскости изображения, аналогична структуре на рис.7. На рис.13 представлен пример вычисления коэффициентов K_h для портретного и фонового изображения, соответственно.

В качестве результирующего значения признака симметрии изображения $g_{k,j}$ принимается максимальное значение функции взаимной корреляции K_h . Положение максимума взаимной корреляции так же позволяет скорректировать геометрический центр тяжести.

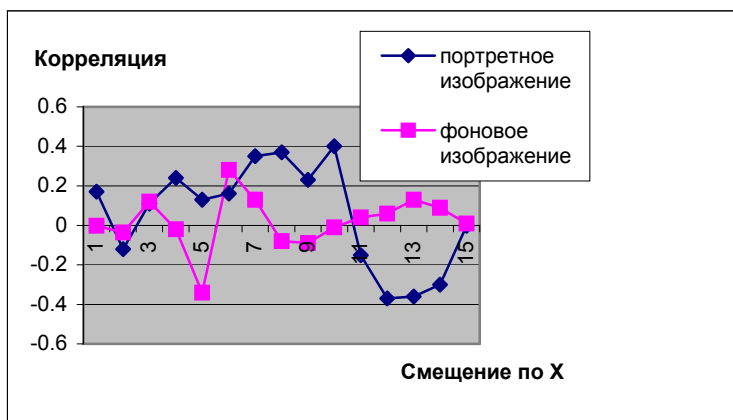


Рис.13.

3.3. Процесс кластеризации

Несмотря на то, что процедура принятия решения (1) определена, проблема разбиения анализируемых данных на кластеры остается открытой. Одним из наиболее популярных подходов для решения этой задачи является оценка качества кластеризации по критерию суммы квадратов ошибок (расстояний) всех m образов, входящих в кластер, до центра кластера. Частным случаем данного подхода является мера (3), где ошибка кластеризации определена стандартным отклонением значений i -го признака. Если условие (3) не выполняется, то из кластера исключается i -ый признак m -го

образа. Если условие (3) выполняется для числа признаков n большего половины $n > N/2$ всех признаков, то образ идентифицируется как принадлежащий данному кластеру. Таким образом, построение кластеров сводится к задаче обучения без учителя.

4. Выводы и обсуждение результатов

Эксперименты показали, что система работоспособна и уверенно классифицирует типовые портретные изображения. Правильное решение, для произвольного набора изображений, получено в 74% случаях, ошибка I-го рода составила 33%, а ошибка II-го рода - 14%. Кластер портретных изображений был сформирован по 8 изображениям, а кластер фоновых изображений по 12. Произвольный набор изображений составил 200 изображений. Как кластерные наборы, так и произвольный набор изображений формировались по запросам в поисковых системах Интернета. Примером может служить запрос “каталог мужских анкет и фотографий” в поисковой системе Яндекса:

<http://www.yandex.ru/yandsearch?text=%CA%E0%F2%E0%EB%EE%E3+%EC%F3%E6%F1%EA%E8%F5+%E0%ED%EA%E5%F2+%E8+%F4%EE%F2%EE%E3%F0%E0%F4%E8%E9&stype=www>

Одной из причин превалирования ошибок I-го рода является естественное присутствие изображений, схожих с портретными изображениями. Типичным примером может служить изображение на рис.14, которое классифицируется как портретное.

Пути повышения вероятности распознавания, мы связываем с потенциальными возможностями методов разбиения множеств индикаторных линий на классы при помощи свойств элементов множеств (п.3.1.4). Успешным примером такого разбиения служит применение меры близости (4), которая обеспечила разбиение групп индикаторных линий $L(\rho, \varphi)$ на классы на формальном уровне (5) (рис.10 и рис.11).

Основываясь на матрице расстояний $r_{i,j}$ можно разбить бинарные множества $B = \{b / b = 1\}$ и/или целочисленные

множества $W = \{w / w = z\}$ на подмножества, обладающие свойствами минимального остового дерева, пирамидального дерева, свойством двухсвязности и получить, в конечном счете, формальные словесные описания изображений.



Рис.14.

Дальнейшее совершенствование предусматривается при формировании признакового пространства коэффициентов ряда Фурье (6). В первую очередь потребуется существенное сжатие признакового пространства гармоник. Если сжатие связать с предварительным выделением головы (лица) на портрете, то предполагается уменьшение ошибок II-го рода.

Повышение эффективности классификации мы так же связываем с определением оптимальной нормировки (2) решающего правила (1). В качестве нормировочных множителей мы использовали варианты средних и дисперсий признаков двух искомых классов. При этом наблюдались 3-5% вариации вероятности правильного распознавания.

5. Литература

1. Глезер В.Д. Зрение и мышление, Л-д, Наука, 1993 – 284с.
2. Головок В.А. Нейроинтеллект: Теория и применения. Книга 2. Самоорганизация, отказоустойчивость и применение нейронных сетей – Брест: БПИ, 1999, - 228с.

3. Журавлев Ю.И., Корректные алгебры над множествами некорректных (эвристических) алгоритмов, Кибернетика, № 4, 6 1977, № 2, 1978.
4. Заде Л.А. Размытые множества и их применение в распознавании образов и кластер-анализе. В сб.: Классификация и кластер. М: Мир, 1980, с.208-247.
5. Красильников Н.Н., Шелепин Ю.Е., Красильникова О.И. Пространственно-временная функциональная модель первичных звеньев зрительной системы, Оптический журнал, 2004, т.71, №7, с.24-31.
6. Логвиненко А.Д. Чувственные основы восприятия пространства, М., Изд. МГУ, 1985 – 224с.
7. Смирнов М.В. Распознавание изображения края полуплоскости в задаче количественного анализа изображений //Сб. статей: Применение методов оптической обработки изображений, Л., РИТ ЛИЯФ, 1986.
8. Ту Дж., Гонсалес Р., Принципы распознавания образов. - Мир, М., 1978.
9. Jacobsen C., Zscherpel U., Perner P.: A Comparison between Neural Networks and Decision Trees - Machine Learning and Data Mining in Pattern Recognition, Springer Verlag, LNAI 1715, 1999, pp. 144-158.
10. Nishida, H., A structural approach to representation of curved objects, Pattern Recognition, Vol. 30, No.1, 1997, pp. 45-56.
11. Nishida Hirobumi, Mori Shunji: Algebraic Description of Curve Structure. IEEE Trans. Pattern Anal. Mach. Intell. 14(5): 516-533 (1992).
12. Yoon K. S., Ham Y. K. and Park R.-H. Hybrid approaches to frontal view face recognition using the Hidden Markov Model and Neural Network. Pattern Recognition 1998 Vol. 31, pp. 283-293.
13. Smirnov M.V., Korolev A.N., Sivjakov I.N., Zvezdin D.I.. Method and software of automatic resolution evaluating of optic-electronic telescopic land remote sensing system // AeroSense'96 Symposium. Visual Information Processing V, Proc. SPIE Vol. 2753, Orlando, USA, April 1996.

Typical portrait images recognition in the automatic classification problem

Smirnov M.V., Sivyakov I.N.

Abstract

The decision of a problem of recognition of a narrow class of objects tends to formation of clusters [8]. The success images clusterization is substantially based on heuristic algorithms. Heuristics can be shown to formal algebras [3] and in particular to fuzzy ensemble. The essence of the offered decision is based on joint application of independent attributes of images, such as, image edge, brightness and Fourier coefficients, in a context of visual perception model. Experiments have shown, that the system is efficient and confidently classifies typical portraits. The correct decision, for arbitrary set of images, is received in 74 % cases.