

УДК 621.372:519.72

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ МЕТОДОМ НАПРАВЛЕННОГО ПЕРЕБОРА НА ОСНОВЕ ПРИНЦИПА МИНИМУМА ИНФОРМАЦИОННОГО РАССОГЛАСОВАНИЯ*

© 2010 г.

А.В. Савченко

Государственный университет – Высшая школа экономики, Нижегородский филиал

savchenko@tecom.nnov.ru

Поступила в редакцию 15.09.2009

Ставится и решается задача автоматического распознавания изображений из большой базы данных на основе принципа минимума информационного рассогласования. Предложен метод направленного перебора альтернатив в информационной метрике Кульбака–Лейблера в противовес их полному перебору. Представлены программа и результаты экспериментального исследования. Показано, что предложенный метод характеризуется повышенной степенью точности и надежности распознавания изображений.

Ключевые слова: автоматическое распознавание изображений, распознавание образов, классификация с обучением, критерий минимума информационного рассогласования.

Введение

Принцип минимального информационного рассогласования (МИР) является, как известно [1], оптимальным в байесовском смысле средством решения широкого круга задач в области распознавания образов. Между тем до настоящего времени практически не исследовались преимущества принципа МИР в задачах автоматического распознавания изображений (АРИ). Исследованиям в этом актуальнейшем направлении и посвящена предлагаемая работа. Акцент в ней сделан на учете характерной особенности задачи АРИ – количество эталонов в реальных базах данных составляет сотни и тысячи единиц. Большинство известных алгоритмов [2], работающих на основе сплошного перебора эталонов, не могут быть реализованы в этой задаче в режиме реального времени. В противовес всем таким алгоритмам в данной статье предложен метод направленного перебора (МНП), основанный на теоретико-информационном подходе и отталкивающийся от метрических свойств решающей статистики МИР. Полученные результаты и сделанные по ним выводы рассчитаны на широкий круг специалистов в области современной теории и практики распознавания образов.

Критерий МИР

Пусть задано множество из $R > 1$ сложных, полутоновых изображений $X_r = \|x_{uv}^r\|, u = \overline{1, H}, v = \overline{1, W}$. Здесь H – высота изображения, W – его ширина; $x_{uv}^r \in \{1, 2, \dots, x_{\max}\}$ – интенсивность точки изображения с координатами (u, v) ; x_{\max} – максимальное значение интенсивности. Предполагается, что эталоны X_r задают некие классы изображений, например, как способ защиты от влияния помех. При этом каждый класс характеризуется тем, что принадлежащие ему объекты обладают некоторой общностью или сходством в характеристиках. То общее, что объединяет объекты в класс, и называют *образом*. Задача состоит в том, чтобы отнести вновь поступающее (на вход) изображение $X = \|x_{uv}\|$ к одному из R таких классов. Это типичный пример задачи диагностики (распознавания образов с обучением) для объектов нечисловой природы (ОНП) [3].

Напомним, что термин ОНП относится к элементам математического пространства, не являющегося линейным (векторным). Процедуры построения решающих правил для решения поставленной задачи в общем случае делятся на детерминированные и стохастические. В настоящее время наиболее часто используется первый, детерминистский, подход. В рамках такого подхода в пространстве ОНП определяется некое расстояние (мера близости) между

* Статья рекомендована к печати программным комитетом Международной научной конференции «Параллельные вычислительные технологии 2009» (<http://agora.guru.ru/pavt>).

любыми парами объектов. Зачастую [2, 4] для АРИ применяется критерий, основанный на стандартной l_1 -метрике

$$\rho_1(X/X_r) = \frac{1}{W \cdot H} \sum_{u=1}^H \sum_{v=1}^W |x_{uv} - x_{uv}^r| \rightarrow \min. \quad (1)$$

К сожалению, подобный подход не всегда позволяет получить удовлетворительные результаты. Это обстоятельство связано, во-первых, с известной [2] вариативностью зрительных образов, а во-вторых, с наличием во входном изображении X помех, таких как не определенная заранее интенсивность источников освещения или просто случайное искажение некоторых точек изображения.

Указанные проблемы при детерминистском подходе обычно решаются путем добавления новых изображений к множеству эталонных изображений (МЭИ), что приводит, в свою очередь, к резкому увеличению его объема. Во всех перечисленных случаях на помощь приходит второй, статистический, подход [2]. В рамках данного подхода выполним нормировку интенсивности изображений из МЭИ, то есть перейдем к рассмотрению нормированных изображений

$$\bar{X}_r = \|\bar{x}_{uv}^r\|, \text{ где } \bar{x}_{uv}^r = \frac{x_{uv}^r}{S_r}, u = \overline{1, H}, v = \overline{1, W}.$$

Здесь введено обозначение $S_r = \sum_{u=1}^H \sum_{v=1}^W x_{uv}^r$.

Аналогичная процедура нормировки выполняется и для входного изображения:

$$\bar{X} = \|\bar{x}_{uv}\|, \bar{x}_{uv} = \frac{x}{S}, S = \sum_{u=1}^H \sum_{v=1}^W x_{uv}.$$

Подобная нормировка [5] применяется в задаче АРИ во избежание указанной выше проблемы случайной интенсивности источника освещения.

Можно предположить, что нормированное изображение-эталон \bar{X}_r определяет собой распределение некой (гипотетической) двумерной дискретной случайной величины. Подобная интерпретация выглядит вполне обоснованной, если учесть, что для \bar{X}_r выполняются все свойства дискретного распределения:

$$\sum_{u=1}^H \sum_{v=1}^W \bar{x}_{uv}^r = 1$$

(условие нормировки) и $\bar{x}_{uv}^r > 0, u = \overline{1, H}, v = \overline{1, W}$ (условие регулярности [1]). Задача состоит, в таком случае, в проверке R гипотез о распределении $X_r, r = \overline{1, R}$, сигнала изображения на входе X . Как показано в работе [1], оптимальное в байесовском смысле решение задачи проверки гипотез о распределении дискретной случайной

величины дает принцип МИР с решающим правилом следующего вида

$$\rho_{KL}(X/X_r) = \sum_{u=1}^H \sum_{v=1}^W (\bar{x}_{uv} \ln(\bar{x}_{uv} / \bar{x}_{uv}^r)) \rightarrow \min. \quad (2)$$

Статистика $\rho_{KL}(X/X_r)$ здесь определяет информационное рассогласование (ИР) по Кульбаку–Лейблеру [6] между наблюдаемым сигналом изображения X и его r -м эталоном из МЭИ $\{X_r\}$.

Таким образом, процедура АРИ в данном случае реализуется по схеме многоканальной обработки, в которой число каналов R определяется количеством изображений-эталонных. Решение принимается по критерию минимума решающей статистики из выражения (1) – для традиционных методов решения задачи АРИ или из выражения (2) – при использовании принципа МИР.

Метрические свойства решающей статистики МИР

Рассмотрим наиболее актуальный и для теории, и для практики случай $R \gg 1$, когда решается задача АРИ с объемом МЭИ в сотни и даже тысячи изображений. В указанных условиях практическая реализация оптимального решающего правила (2) по схеме R -канальной обработки наталкивается на очевидную проблему его вычислительной сложности и даже практической реализуемости, особенно если учитывать трудоемкую процедуру выравнивания изображений по множеству их параметров: размерам, цвету, ракурсу и т.п. В поиске путей решения указанной проблемы за счет отказа от сплошного перебора МЭИ и состоит центральная идея настоящей работы.

Прежде всего, отметим метрические свойства решающей статистики МИР $\rho_{KL}(X/X_r) \geq 0$ с равенством ее нулю лишь в идеальном случае совпадения входного и эталонного сигналов. Поэтому вначале преобразуем критерий МИР (2) к упрощенному (в его практической реализации) виду [6]

$$W_v(X): \rho_{KL}(X/X_v) < \rho_0 = \text{const}. \quad (3)$$

Здесь ρ_0 – порог для допустимой величины ИР на множестве одноименных изображений за счет известной их вариативности. Значение такого порога нетрудно установить опытным путем. По своей сути выражение (3) определяет условие «останова» при переборе альтернатив в рамках проверочной процедуры по критерию МИР (2).

Таким образом, при принятии решения на основе принципа МИР (2) требуется просматривать не все эталоны, а вычислять величину ИР лишь до тех пор, пока оно не будет меньше некоторого порогового уровня. Нетрудно понять, что само по себе указанное обстоятельство позволит сократить объем перебора в среднем на 50%. Иными словами, благодаря использованию правила останова (3) удастся в два раза сократить объем выполняемых вычислений и этим существенно ослабить проблему практической реализуемости АРИ в режиме реального времени. В этом состоит принципиальное преимущество МИР по сравнению со всеми его наиболее известными статистическими аналогами, в которых применяются классические (байесовские) критерии: минимума среднего риска, максимума апостериорной вероятности и др. [2]. Между тем, как это выясняется ниже, рассмотренный выигрыш в вычислительной сложности и производительности далеко не исчерпывает всех преимуществ принципа МИР в задаче распознавания изображений.

Действительно, общая формулировка задачи (2) позволяет рассматривать ее как задачу оптимизации и применять алгоритмы поиска оптимального решения с заданным условием останова (3). В такой задаче на множестве эталонных изображений $\{X_r\}$ требуется найти такое изображение X_v , которое будет минимизировать статистику МИР. В этом случае метод, сводящийся к полному перебору МЭИ, является одним из множества известных методов оптимизации систем. Главным препятствием для применения в нашей задаче более эффективного оптимизационного метода является то, что, во-первых, задача относится к области дискретной математики и, во-вторых, в ней требуется найти глобальный минимум решающей статистики (2). По-видимому, наиболее обоснованным способом поиска глобального экстремума в указанных условиях можно считать метод случайного поиска. К сожалению, известные классические методы оптимизации (такие как генетический алгоритм) не учитывают информации о самих изображениях, рассогласованиях между ними. Более того, в большинстве дискретных алгоритмов оптимизации не существует четкого критерия останова поиска. На помощь снова приходит принцип МИР. Действительно, на основе того же выражения (3) мы можем сформулировать искомый критерий останова алгоритма оптимизации. В этом случае появляется гарантия того, что решение задачи, если оно существует (то есть если входное изображение

относится к одному из классов, заданному $\{X_r\}$), будет найдено. Естественным развитием этой идеи может служить предложенный ниже метод направленного перебора (МНП) МЭИ, в котором метрические свойства решающей статистики МИР (3) используются в наиболее полной степени.

Идея МНП

Следуя общей схеме вычислений (2), (3), сведем задачу АРИ X к проверке сначала N первых вариантов X_1, \dots, X_N из заданного R -множества альтернатив $\{X_r\}$ при условии $N \ll R$. Если по крайней мере одна из них, а именно $X_v, v \leq N$, отвечает требованию останова (3), процесс поиска оптимального решения по критерию МИР (2) на ней и завершается. Однако в общем случае можно предположить, что ни одна из первых N альтернатив проверку (3) на первом шаге не проходит. В таком случае можно проверить вторую группу из N эталонных изображений в пределах множества $\{X_r\}$, потом третью группу и т.д. – до момента выполнения условия (3). Но есть и иной, более рациональный, вариант решения поставленной задачи.

Расставим изображения из нашей первой контрольной выборки X_1, \dots, X_N в порядке убывания их ИР $\rho_{KL}(X/X_n), n = \overline{1, N}$. В результате будем иметь упорядоченную (ранжированную) последовательность эталонных изображений вида

$$\{X_{i_j}\} = \{X_{i_1}, X_{i_2} \dots X_{i_N}\}, i_j \leq N.$$

Соответствующая последовательность $\{\rho_j\}$ их ИР $\rho_j = \rho_{KL}(X/X_{i_j}), i_j \leq N$, будет иметь характер монотонно убывающей зависимости.

Заметим, что ИР вида (3) представляет собой непрерывную функцию своих аргументов – относительных интенсивностей двух изображений. В этом случае можно предположить, что если некоторый эталон из МЭИ X^* удовлетворяет условию (3) $\rho_{KL}(X/X^*) < \rho_0 \ll 1$, то для произвольного образа X_i из МЭИ выполняется условие $|\rho_{KL}(X/X_i) - \rho_{KL}(X^*/X_i)| \ll 1$.

Отсюда и вытекает главная идея МНП: использовать последний элемент X_{i_N} из упорядоченной контрольной выборки $\{X_{i_j}\}$ как наилучшее приближение к искомому изображению

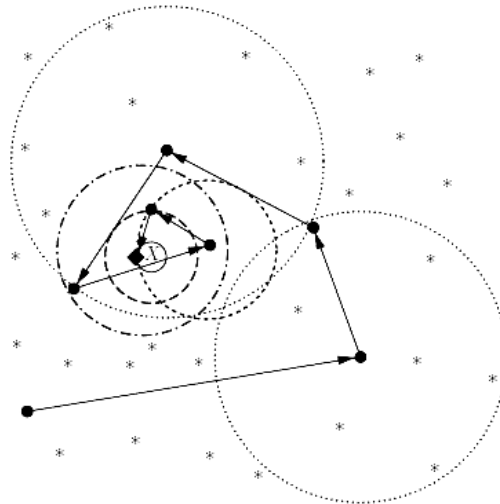


Рис. 1. Поисковая процедура МП

X в роли точки отсчета для поиска наиболее подходящих «кандидатов» в очередную контрольную выборку. При этом ориентиром для определения максимально допустимых различий (в теоретико-информационном смысле) изображений-эталонов из будущей контрольной выборки по отношению к «точке отсчета» X_{i_N} будет служить рассогласование $\rho_{KL}(X/X_{i_N})$.

Проиллюстрируем сказанное с помощью диаграммы поисковой процедуры МП на рис. 1.

Здесь звездочками обозначены все имеющиеся изображения-эталоны, буквой X – входное изображение, а ромбиком – наиболее близкий к X эталон. Он и определяет искомое оптимальное решение задачи. Траектория поиска отображается на рис. 1 ломаной направленной линией. Жирными точками на ней обозначена последовательность наиболее близких к оптимуму изображений X_{i_N} после нескольких подряд этапов вычислений. Окружностями здесь отмечены границы соответствующих контрольных точек-выборок X_1, \dots, X_N . Их радиусы определяются согласно ИР до входного изображения X . Хорошо видно, что траектория поиска имеет вид скручивающейся спирали.

Синтез алгоритма

Следуя определению ИР (2) составим $(R \times R)$ -матрицу $P = \|\rho_{ij}\|$ значений ИР

$\rho_{ij} = \rho_{KL}(X_i/X_j)$, $i, j \leq R$. Эту весьма сложную в вычислительном отношении операцию требуется выполнить лишь раз: на предвари-

тельном этапе вычислений и для каждого конкретного МЭИ. После этого, как это было описано выше, зададимся в пределах имеющегося R -множества эталонов $\{X_r\}$ произвольной первой контрольной выборкой X_1, \dots, X_N некоторого фиксированного объема N , по ней получим ранжированный по критерию МИР (2) ряд данных $\{X_{i_j}\}$ и, наконец, найдем из него первый

локальный оптимум X_{i_N} . На этом завершается первый этап вычислений. На втором этапе для выделенного изображения-эталона X_{i_N} по матрице P найдем множество из $M < R$ изображений $X^{(M)} = \{X_{i_{N+1}}, \dots, X_{i_{N+M}}\}$, $i_j \leq R$, находящихся от изображения X_{i_N} на «расстоянии» (2), не превышающем порогового значения $\rho_N = \rho_{KL}(X/X_{i_N})$, или

$$\begin{aligned} & (\forall X_i \notin X^{(M)}) (\forall X_j \in X^{(M)}) \\ & \Delta\rho(X_i) \geq \Delta\rho(X_j). \end{aligned} \quad (4)$$

Здесь $\Delta\rho(X_i) = |\rho_{KL}(X_i/X_{i_N}) - \rho_N|$ – отклонение ИР между входным изображением X и локальным оптимумом X_{i_N} относительно ИР между парой изображений X_i и X_{i_N} . На рис. 1 каждое такое множество ограничивается соответствующей окружностью с центром в точке X_{i_N} . Добавим к этому множеству еще один, $(M+1)$ -й элемент $X_{i_{N+M+1}}$ из числа не попавших в состав контрольной выборки по результатам предыдущего этапа вычислений. Этим мы внесем в поисковую процедуру определенный элемент случайности как способ достижения гло-

бального оптимума за конечное число шагов оптимизации (этапов вычислений). В результате получаем вторую контрольную выборку изображений-эталонов $\{X_{i_N}, \dots, X_{i_{N+M+1}}\}$, $i_j \leq R$, для анализа. Далее все вычисления первого этапа циклически повторяются до тех пор, пока на некотором K -м этапе для элемента X_{i_N} не будет выполнено условие останова (3), т.е. при условии

$$\rho_{KL}(X / X_{i_N}) < \rho_0. \quad (5)$$

На рис. 1 в такой момент входное изображение оказывается в пределах границ множества контрольных точек последнего этапа вычислений. Решение здесь принимается в пользу наиболее близкого образа X^* . Или, в худшем случае, после перебора всех альтернатив из множества $\{X_r\}$, но в отсутствие решения из (5), делается вывод о том, что входное изображение X нельзя отнести ни к одному классу из МЭИ и необходимо задействовать режим переспроса. В общем же случае суммарное число $N + (M + 1) \cdot K \leq R$ выполняемых согласно (5) проверок может существенно выигрывать по сравнению с объемом используемого МЭИ. Этот выигрыш обусловлен, в частности, тем обстоятельством, что для рассогласования Кульбака–Лейблера (как, впрочем, и для многих других расстояний, в частности метрики l_1) вероятность p того, что искомый эталон X^* принадлежит множеству $X^{(M)}$, как правило, существенно превышает вероятность того, что X^* будет одним из M наудачу выбранных эталонов:

$$p = P\{X^* \in X^{(M)}\} \gg p_0 = \frac{M}{R}. \quad (6)$$

В этом и состоит эффект направленного перебора. А отличия в количестве этапов K алгоритма для разных экспериментов объясняются тем, что вероятность p зависит не только от свойств используемой метрики, но и от свойств входного изображения и МЭИ.

Таким образом, система выражений (2)...(5) и определяет, в конечном итоге, предлагаемый МНП в задаче АРИ.

Результаты экспериментальных исследований

Для проведения эксперимента воспользуемся реальной базой данных фотографий [7]. Из 5000 фотографий 300 различных людей были отобраны в качестве эталонов $R=1200$ наиболее различающихся изображений. В

этом случае $p_0 = \frac{20}{1200} \approx 0.017$. Кроме того, для тестирования были взяты еще 1000 фотографий тех же людей. Для начала рассмотрим случай, когда все изображения освещены одинаково.

Итак, решалась задача АРИ X из множества всех его допустимых альтернатив $\{X_r\}$. Сначала для этого применялась метрика l_1 из выражения (1). Параметры МНП при этом были выбраны следующим образом: $N=2$, $M=20$. Порог $\rho_0 = 15$ был подобран экспериментально. Здесь в 98.3% случаев было получено точное решение $X^* = X_r$. На каждое такое решение в среднем потребовалось проверить (перебрать) 145 изображений, или 12% от объема всего МЭИ. При этом для 9.5% входных изображений X не было выполнено условие останова (3), поэтому алгоритм проверил все R эталонов. Условие (6) в рамках данного эксперимента примет вид

$$p = P\{X^* \in X^{(M)}\} = 0.41 \gg p_0.$$

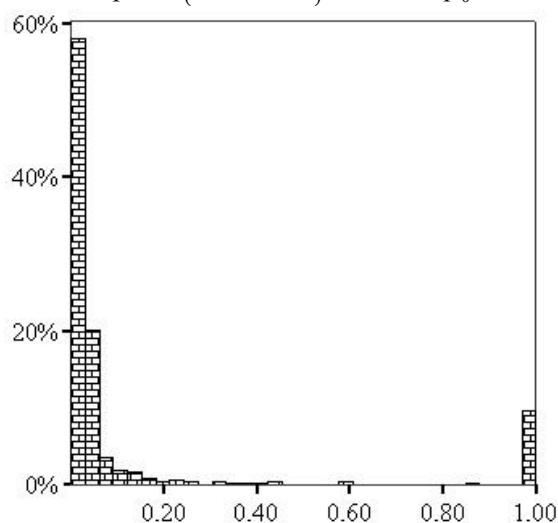


Рис. 2. Гистограммы количества проверок изображений. На оси абсцисс приведено отношение количества вычислений (согласно МНП) ИР к объему МЭИ R

Аналогичная задача с использованием того же МНП решалась позже в метрике Кульбака–Лейблера (2). Параметры N и M МНП были выбраны те же, что и для метрики l_1 . Порог (3) установлен $\rho_0 = 2.5$. В результате в 95% случаев было получено точное решение $X^* = X_r$. Для этого случая гистограмма количества проведенных проверок согласно алгоритму (2)...(5) показана на рис. 2. Здесь среднее количество проверок составило примерно 14% от объема МЭИ. С вероятностью 85% оно не превы-

шает 178 или 14.9% от общего числа изображений-эталонов для проверки. Для 10.2% исходных изображений условие (3) не было выполнено ни для одного эталона из МЭИ и были проверены все R альтернатив. В этом случае вероятность того, что искомым эталон X^* принадлежит множеству $X^{(M)}$, будет равна $p = 0.32$.

Кажется, что результаты, полученные на основе метрики I_1 , существенно превосходят по качеству результаты использования рассогласования Кульбака–Лейблера. Однако это превосходство основано на том, что метрика I_1 в формулировке (1) использует условное предположение о том, что все входные изображения освещены так же, как и эталоны из МЭИ. Для подтверждения немного изменим наш предыдущий эксперимент. Выполним искусственное затемнение тестовых изображений и проведем процедуру АРИ еще раз. В этом случае для метрики Кульбака–Лейблера (2) и МНП (3)...(5) с параметрами N , M и ρ_0 из предыдущего эксперимента получаем среднее количество проверок 15.5% от объема МЭИ при вероятности ошибки 5%. Таким образом, результаты использования рассогласования Кульбака–Лейблера (2) ухудшились несущественно по сравнению с предыдущим экспериментом без искажений тестовых изображений. Однако при использовании метрики I_1 из выражения (1) и алгоритма МНП с порогом $\rho_0 = 15$ результаты получаются значительно хуже. Так, вероятность ошибки повысилась до 16%, при этом среднее число проверок составило 88% от объема МЭИ. В 84% случаев условие (3) не было выполнено ни для одного эталона.

Заключение

Вопрос о повышении скорости вычислений без потери качества распознавания образов вызывает повышенный интерес среди специалистов в области как теории, так и практики АРИ. Применение принципа МИР совместно с критерием останова (3) позволяет, с одной стороны, выполнить

отбраковку сомнительных с точки зрения надежности решений, с другой – сократить объем вычислений в среднем в 2 раза даже в самом невыгодном, переборном, варианте своей реализации. При использовании же предложенного метода в формулировке (2)...(5) количество вычислений и вовсе составляет 10–15% по сравнению с методом сплошного перебора. При этом не утрачивается (по сравнению со сплошным перебором МЭИ) и качество достигаемого решения. Более того, проведенные эксперименты наглядно показывают, что МНП может давать хорошие результаты на основе не только информационной метрики Кульбака–Лейблера (2), но и традиционной «евклидовой» метрики I_1 (1). А условие (6) позволяет в таком случае сделать обоснованный выбор в пользу одной из метрик для проверки гипотез по МНП при учете особенностей применяемого МЭИ.

Таким образом, благодаря проведенному исследованию предложен новый метод для решения задачи АРИ на основе принципа МИР, обладающий широкими функциональными возможностями и высокими эксплуатационными свойствами.

Список литературы

1. Савченко В.В., Савченко А.В. Принцип минимального информационного рассогласования в задаче распознавания дискретных объектов // Известия вузов России. Радиоэлектроника. 2005. Вып. 3. С. 10–18.
2. Горелик А.Л., Гуревич И.Б., Скрипкин В.А. Современное состояние проблемы распознавания: некоторые аспекты. М.: Радио и связь, 1985. 160 с.
3. Орлов А.И. Математические методы исследования и диагностика материалов // Заводская лаборатория. 2003. Т. 69. № 3. С. 53–64.
4. Алексеев К.В., Егорова С.Д. Метод идентификации рукописных данных на основе персонального кода // Известия вузов России. Радиоэлектроника. 2005. Вып. 3. С. 23–31.
5. Phill Kyu Rhee, InJa Jeon, and EunSung Jeong Adaptive normalization based highly efficient face recognition under uneven environments // Lecture Notes in Computer Science. 2005. Vol. 3611.
6. Кульбак С. Теория информации и статистика. М.: Наука, 1967. 408 с.
7. <http://cswww.essex.ac.uk/mv/allfaces/index.html> (дата обращения: 01.04.2009 г.).

IMAGE RECOGNITION BY THE DIRECTED SEARCH METHOD ON THE BASIS OF MINIMUM-INFORMATION-MISMATCH CRITERION

A.V. Savchenko

The problem has been set and solved on the automatic recognition of images from a large database on the basis of minimum-information-mismatch criterion. A directed search method using Kullback-Leibler information distance has been proposed as opposed to exhaustive search. The program and experimental results of the method's investigation have been presented. The proposed method has been shown to have extra accuracy and reliability in automatic image recognition.

Keywords: automatic image recognition, pattern recognition, classification with training, criterion of the minimum of information mismatch.