УДК 681.3.07

ПРИМЕНЕНИЕ ОТНОШЕНИЙ НА МНОЖЕСТВЕ ХАРАКТЕРНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ ПРИ РАСПОЗНАВАНИИ НА ОСНОВЕ ГОЛОСОВАНИЯ

В. А. Гороховатский

ХНУРЭ, г. Харьков, Украина

Приведены новые результаты в теории построения систем распознавания изображений на основе голосования множеств характерных признаков. Основная идея состоит в применении отношений на множестве признаков, что в целом улучшает качественные характеристики распознавания. Обсуждаются вопросы теоретико-множественного представления отношений в процедурах голосования, обоснования вероятностной модели принятия решений на основе функции соответствия признаков, анализа способов использования отношений, оценки эффективности подхода.

КОМПЬЮТЕРНОЕ ЗРЕНИЕ, ИЗОБРАЖЕНИЕ, ХАРАКТЕРНЫЕ ПРИЗНАКИ, РАСПОЗНАВАНИЕ, ГОЛОСОВАНИЕ, ОТНОШЕНИЯ, ФУНКЦИЯ СООТВЕТСТВИЯ КЛАССУ, ВЕРОЯТНОСТНЫЕ МОДЕЛИ

Введение

Эффективным способом установления соответствия распознаваемых объектов с эталонами является метод голосования множеств характерных признаков (точек интереса), которыми представлены изображения [1-6]. Наряду с такими преимуществами как устойчивость и гибкость принятия решения при влиянии локальных помех и фона методы голосования по сравнению с другими подходами, например, с методами согласованной маркировки, обладают таким важным достоинством как простота построения, что при достаточно высокой надежности делает их конкурентоспособными среди известных подходов.

Использование различного типа отношений, построенных на множестве характерных признаков (ХП), дает возможность усовершенствования и улучшения качества работы систем голосования путем дополнения их информационных ресурсов новой высокоуровневой системой признаков.

Цель исследования — формализация и конкретизация аппарата отношений из теории множеств применительно к задаче распознавания визуальных объектов на основе голосования.

Задачи исследования — теоретико-множественное представление процедур голосования, построение отношений на множестве ХП, обоснование вероятностной модели принятия решений на основе функции соответствия признаков, анализ путей конкретного применения отношений, обоснование эффективности обсуждаемого подхода.

1. Формализация процедур голосования

Распознавание на основе множества ХП рассматривается как отображение заданных структурных компонентов модели на компоненты анализируемого изображения [2] и реализуется путем решения задачи оптимизации вида:

$$j^* = \arg\max_{j\in J} \psi(Z, Z(j)), \tag{1}$$

где J — множество классов ($j^* \in J$); $Z, Z(j)$ — соответственно множества характерных признаков распознаваемого изображения и эталона j-го класса; ψ — мера сходства для множеств. Множество классов J можно считать конечным множеством целых чисел.

В теоретико-множественном представлении решение (1) путем голосования сводится к построению на множестве Z разбиения $Z = \cup Z_j$ на непересекающиеся между собой подмножества $Z_i \bigcap\limits_{i\neq j} Z_j = \varnothing$, состоящие из элементов, каждый из которых получает метку класса j путем реализации однозначного отображения $\Lambda : Z \to J$. Отображение Λ реализует локальное решение на множестве ХП. Поиск оптимума в (1) представляется как выбор из сформированных подмножеств Z_j элемента максимальной мощности, то есть

$$j^* = \arg\max_{j\in J} \# Z_j , \tag{2}$$

где символ # означает мощность (количество элементов) множества. Разбиением $Z = \cup Z_j$ на множестве признаков Z фактически устанавливается отношение эквивалентности. Признаки, отнесенные к одному и тому же Z_j, считаются одинаковыми, так как принадлежат классу j. При этом в представлении $Z = \cup Z_j$ не обязательно присутствуют все номера эталонов из J, зато в реальных задачах состав разбиения часто дополняется множеством Z_0, что соответствует случаю неопределенности или отказу от локального решения в пользу одного из эталонов. Рисунок демонстрирует схему разбиения множества признаков Z при реализации однозначного отображения Λ.

Наряду с этим, при построении голосования возможно также применение множественного (неоднозначного) отображения Λ, когда каждый из признаков $z \in Z$ формирует не один, а несколько элементов $j_1,...,j_q$; $j_i \in J$ из множества классов J. В реальности это происходит по причине близости значений признаков разных классов. Представление $Z = \cup Z_j$ в такой постановке не подходит под классическое определение разбиения [7], так как элементы могут повторяться в разных Z_j. Множества Z_j в рамках описанной модели трансформируются так, что могут иметь общие элементы, то есть $Z_i \bigcap_{i \neq j} Z_j \neq \varnothing$.

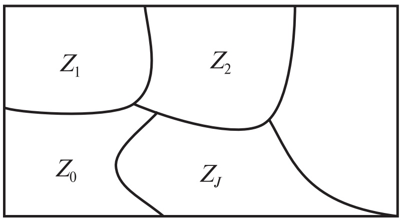

Рис. Схема разбиения множества ХП

Для того чтобы при реализации такого решения не усиливать влияния множественных признаков, их общий голос нужно представить в виде суммы $\sum_{i=1}^{q} \alpha_i = 1$, где α_i — вес отдельного ХП, а q — количество повторных значений классов для конкретного признака. Для упрощения можно считать $\alpha_i = 1/q$. При такой нормировке число голосов остается равным мощности элементов множества Z. Правило (2) для этой модификации остается в силе, а мощность множества вычисляется по формуле

$$\# Z_j = \sum_i \beta_i,\ \beta_i = \begin{cases} 1, \text{ для одиночного голоса,} \\ \alpha_i, \text{ для множественного} \end{cases}, \quad (3)$$

и при необходимости округляется до ближайшего целого значения.

В основу функционирования систем голосования ХП могут быть положены также мажоритарные структуры [8], которые часто связаны с достаточно сложными схемами голосования. В зависимости от используемой схемы возможны парадоксальные ситуации, когда даже в условиях полного равноправия элементов сколь угодно маленькая их доля оказывается решающей коалицией [8]. Не углубляясь в подробности построения сложных схем голосования, будем использовать для наших целей лишь наиболее простые из них. При прямом голосовании решение принимается большинством, превышающим половину голосующих элементов. Прямое, но не равное, голосование состоит в том, что элементы имеют разные веса, сумма которых равна единице: $\sum_{z \in Z} \gamma(z) = 1$. Такая нормировка требуется в целях принятия решения путем управления значением величины, характеризующей решающее число голосов и называемой коалицией. Решение принимается коалицией $\widetilde{Z} \subset Z$ в виде $\sum_{z \in \widetilde{Z}} \gamma(z) \geq \varepsilon$, где обычно считают $0{,}5 < \varepsilon \leq 1$ — порог голосования [8]. Коалиция для нашей задачи — это максимальное по мощности из множеств Z_j.

При снятии ограничения на правило «больше половины» значение ε может быть выбрано из условия $\varepsilon < 0{,}5$. При этом заранее предполагается возможное решение на основе относительно небольшого числа элементов [3]. Такое решение не вписывается в рамки традиционных статистических решений. Его надежность должна быть заранее оговорена и обоснована, так как речь идет о том, что в процессе голосования объект может быть представлен незначительной частью своих элементов.

Важным моментом, усиливающим эффективность распознающих систем на базе голосования, является использование комбинаций локальных признаков, основанных на построении отношений между ними [1,4-6]. Построение и использование отношений в целом дает возможность на основе имеющегося множества ХП получить устойчивые признаки более высокого уровня. Применение отношений между признаками естественным образом заложено в самой природе структурного представления как путь получения целостной информации о распознаваемом объекте. В конечном итоге любой визуальный объект можно формально описать путем построения отношений на множестве его частей [1].

С другой стороны, очевидно, что для визуального объекта в зависимости от условий его восприятия системой распознавания существует оптимальное в смысле степени глобальности представление, позволяющее осуществить его распознавание с достаточной надежностью. Это следует из опыта работы зрительного анализатора человека. Вся проблема состоит в том, чтобы успешно адаптироваться к этим условиям. Если помехи и фон вообще отсутствуют, то лучше применить методы интегрального типа, не предполагающие разбиения на компоненты. Структурный анализ в этой ситуации есть смысл применять только в случаях сильной схожести распознаваемых объектов. Основное предназначение структурного представления в виде множества характерных признаков — надежное распознавание в условиях частичных искажений объектов [3].

В теории системного анализа рассматривают два основных способа изучения структуры некоторого объекта [8]. Один из них состоит в том, чтобы «препарировать» его внутреннее содержание, установить состав и структуру частей, составляющих этот объект. Другой способ состоит в «проектировании» объекта на некоторую совокупность близких объектов. По свойствам проекций выносится суждение о внутреннем содержании изучаемого объекта. Считается, что для сложно организованного объекта второй способ более приемлем. При анализе и распознавании объектов в системах компьютерного зрения нашли применение оба способа [1-6].

2. Отношения на множестве ХП

Отношение (relation) — это абстрактное математическое понятие. Наибольшую известность в анализе изображений получили топологические отношения, которые учитывают взаимное расположение, а также метрические, оценивающие расстояние между элементами. Такие топологические отношения, как связность, смежность, инцидентность обладают устойчивостью к распространенным дефектам изображений, приводящим к нарушению метрических отношений. Примером метрических отношений есть расстояние, которое определено для любых двух элементов изображения. Расстояние в координатном понимании инвариантно относительно поворотов и переносов, но не относительно масштабирования.

Сопоставление с применением топологических и метрических отношений используют в целях расширения множества элементов, для которых установлено соответствие [1].

Пусть $A=\{a^i\}$ — некоторое множество элементов, которыми могут быть, например, ХП. Совокупность упорядоченных наборов (кортежей) из k элементов множества A вида $<a^1,a^2...,a^k>$, $a^i \in A$ называется k-местным отношением R_A на множестве A. Иначе говоря, отношение R_A — это подмножество декартова произведения $R_A \subseteq A \times A \times ... \times A$, в котором множество A участвует k раз. Величина k — это число «мест» или «арность» отношения. Более формально, отношение — это упорядоченная пара $<R_A, A>$. Множество с заданным набором отношений получило в математике название «модель» [8]. Таким образом, кортеж $\mathcal{A}=<A,\mathcal{R}>$, где $\mathcal{R}=\{R_A^1,...,R_A^n\}$, представляет модель с n отношениями, в терминах которой описывается распознаваемый визуальный объект.

Различают унарные, бинарные или k-арные ($k>2$) отношения в зависимости от количества участвующих в них элементов [7]. Унарные отношения характеризуют индивидуальные признаки элемента (ХП). Бинарные отношения наиболее популярны из-за простоты их анализа и построения. Кроме того, имеются простые и атрибутивные типы отношений. В атрибутивных отношениях кортежу из k элементов сопоставляется некоторый атрибут (число), представляющий, например, степень выполнимости отношения.

В задачах анализа изображений часто используют бинарные отношения, связанные с пространственной близостью элементов (в смысле координат), а также отношения близости в признаковом пространстве, фиксирующие эквивалентность значений признаков [4]. Примерами выражений, задающих указанные отношения, являются

$$|x_1 - x_2| + |y_1 - y_2| \le \delta,$$

$$\sum_{i=1}^{n} |z_i^1 - z_i^2| \le \varepsilon,$$

где x, y — координаты; z_i^k — компоненты вектора признаков; δ, ε — пороги.

Именно на отношении пространственной близости точек построено описание локальной окрестности, пиксели которой используются при формировании ХП. На отношениях близости значений в пространстве признаков так или иначе реализуются все подходы, связанные с принятием решений о типе признака или классе объекта. Другой популярный тип — отношения порядка, когда для пары элементов установлено отношение вида «<» или «$\le$».

Будем для определенности рассматривать бинарные отношения. Тот факт, что элементы множества $a^i, a^k \in A$ находятся в отношении R_A, записывается в виде $(a^i, a^k) \in R_A$ или $a^i R_A a^k$. Множество бинарных отношений имеет вид $R_A \subseteq A \times A$. Для унарных отношений $a^i \in R_A$, при этом $R_A \subseteq A$. Атрибутивное отношение записывается как $d(a^i, a^k) \in R_A$, где d — число или вектор, отображающий степень выполнимости отображения R_A на этой паре. Для простых отношений принято $d=1$. Отношения называются симметричными, если из условия $(a^i, a^k) \in R_A$ вытекает $(a^k, a^i) \in R_A$. Бинарные отношения могут быть заданы матрицей или графом.

На базе множества $\mathcal{R}=\{R_A^1,...,R_A^n\}$ отношений элементов из A можно получить отношения более высокого уровня. Примером такой иерархии отношений может быть пирамидальное или вейвлетное представления. Теория отношений лежит в основе теории выбора и принятия решений [7, 8].

Очевидно, что максимальная мощность множества пар элементов, удовлетворяющих симметричному бинарному отношению R_A, определяется числом комбинаций $m_R = C_m^2$, где m — мощность множества A, C_m^2 — число комбинаций. В результате мощность множества бинарных отношений R_A можно оценить величиной $m_R = m(m-1)/2$, что в $(m-1)/2$ раза больше числа элементов исходного множества A.

По причине роста количества признаков, построенных на множестве отношений, перед их применением необходимо каким-либо образом структурировать информацию, чтобы добиться сокращения числа анализируемых элементов. Кроме того, часто так выбирают содержимое отношений, чтобы уменьшить количество признаков в сформированной системе [1,7].

Пусть $V=\{v^i\}$ — множество значений ХП изображения, $C=\{c^i\}$ — множество их координат, а $Z=V \cup C$ — их объединение, $z^i=\{(v^i,c^i)\}$. Между элементами множеств V и C имеется однозначное соответствие вида $C \to V$. Будем рассматривать три типа отношений: R_V как отношения на множестве V, R_C — отношения на множестве C, а также комбинированные отношения $R_Z = R_{VC}$, заданные на множестве пар элементов признак — координаты, то есть $([v^i,c^i],[v^k,c^k]) \in R_{VC}$.

Заметим, что отношения, как и признаки, могут обладать свойством инвариантности к геометрическим преобразованиям. Этот факт можно использовать в процессе применения отношений.

Пусть имеется некоторая база данных $E=\{E_1,...,E_J\}$, содержащая J моделей (эталонов) объектов. Модель объекта E_j, в свою очередь, включает множество значений $Z(j)$ и множество $R_{Z(j)}$ отношений признаков, то есть $E_j=\{Z(j),R_{Z(j)}\}$. При этом вид отношений R_Z будем считать известным и фиксированным для базы данных и для распознаваемого изображения. Унарные отношения находят свое отражение непосредственно в значениях признаков $Z(j)$. Если рассматривать $R_{Z(j)}$ как множество бинарных отношений, то E_j будет представлять собой совокупность пар z_t,z_ν, характеризуемых значениями конечного множества атрибутов $d_1,...,d_\theta$, где θ — количество отношений, то есть $E_j=\{z_t,z_\nu,d_1,...,d_\theta\}$, $z_t,z_\nu \in Z(j)$. Эталонная информация, на основе которой строится голосование, представима в виде следующей таблицы.

z_t	z_ν	d_1	...	d_θ

Аналогичная таблица строится и для всех пар ХП из множества Z.

При внесении описанных изменений схема распознавания (1), (2) не изменится, меняются лишь вид и число элементов, участвующих в голосовании, а также мера для сопоставления этих элементов. Член голосования с учетом отношений представляет собой вектор $r_{t\nu}[d]=(z_t,z_\nu,d_1,...,d_\theta)$, $d_i \in R_Z$.

Одним из вариантов голосования является представление, когда отношения между признаками задаются неявным образом, то есть без вычисления значений $(d_1,...,d_\theta)$. Примером такого задания отношений есть алгоритмический (процедурный) путь, когда признаки из множества Z при формировании новой системы признаков R_Z выбираются посредством некоторой процедуры.

Теоретико-множественная формулировка задачи распознавания с учетом отношений сводится к получению представления $R_Z=\cup R_{Z(j)}$ на множестве отношений R_Z множества Z и решения задачи оптимизации (2) в виде

$$j^* = \arg\max_{j\in J} \#R_{Z(j)}. \quad (4)$$

В связи с появлением дополнительной информации в виде значений атрибутов или процедурных отношений расширяется разнообразие возможных схем принятия как локального Λ, так и глобального решений. При этом перспективно совместное применение правил (2) и (4), когда одновременно используются отношения и значения признаков. Еще один вариант комбинированного построения состоит в поэтапном использовании признаков и их отношений. Например отношения строятся и проверяются лишь в случае, если принято одинаковое локальное решение относительно входящих в них признаков.

Теория отношений не ограничивается применением лишь к структурным элементам распознаваемых объектов. Учитывая, что отношения можно построить также для элементов разных множеств, саму процедуру голосования и подсчета голосов можно трактовать как реализацию некоторого бинарного отношения.

Формально бинарное отношение $Q(j)$ для двух разных множеств Z и $Z(j)$ опишем как тройку $\langle Q(j),Z,Z(j)\rangle$, где $Q(j)\subseteq Z\times Z(j)$. Во множество $Q(j)$ могут быть включены, например, пары равных между собой в некотором смысле элементов множеств $Z,Z(j)$. Заметим, что во множества признаков $Z,Z(j)$, в свою очередь, могут включаться внутренние отношения элементов этих множеств. При такой формулировке решение (2) или (4) выглядит как поиск среди $\{Q(j)\}$ множества максимальной мощности и определение в соответствии с этим значения класса. При этом трактовка (2) фактически основана на построении пересечения множеств $Z,Z(j)$, а (4) — на пересечении множеств отношений $R_Z,R_{Z(j)}$. При этом считается, что множества совпадают, если любой объект первого множества содержится во втором и наоборот, каждый элемент второго множества входит в первое.

3. Примеры построения отношений ХП

Рассмотрим вначале применение отношений на примере метода частных корреляций [2, 3]. В его классическом представлении локальное решение принимается, если сходство двух соответствующих фрагментов объекта и эталона будет выше некоторого порога δ_1. Глобальное же решение принимается, если количество положительных локальных решений превышает порог $n=\delta_2 m$, где m — общее количество фрагментов, соответствующее максимально возможному числу локальных решений. Здесь имеем n-арное отношение на множестве значений локальных сходств, на выполнении которого основано построение глобального решения. Функционально это отношение фиксируется местоположением фрагментов (множеством номеров) в построенной системе. Словесно данное отношение можно сформулировать так: «величины сходства n любых фрагментов из общего количества m превышают порог δ_1».

Схема независимого голосования (без применения отношений) состоит в том, что каждый фрагмент принимает самостоятельное решение, а далее на их основе реализуется глобальное решение путем поиска максимума среди подмножеств локальных решений для отдельных классов. Сравнивая схемы, можем рассчитывать на большую надежность схемы, основанной на отношениях по сравнению с независимым голосованием. Это предположение основано на использовании несколько большего объема информации об объектах, связанной с учетом структуры системы фрагментов.

В рассмотренной схеме метода частных корреляций на глобальном и локальном уровнях используется отношение типа R_V, хотя на локальном уровне неявно при вычислении сходства применено также отношение типа R_C, строго устанавливающее соответствие координат сравниваемых фрагментов в структуре объекта.

Другой пример использования отношений на множестве ХП описан в [5]. Здесь отношения строятся путем формирования новой системы высокоуровневых признаков, которые включают: значение (или голос) центрального для некоторой окрестности ХП; условие, что более половины из имеющегося множества ХП в данной окрестности имеют то же значение; условие совпадения углов между тройками согласованных ХП с соответствующими углами для объекта. Эти свойства названы «полулокальными условиями». Такие условия, по мнению авторов, существенно увеличивают селективную силу исходной системы голосования, что явно усиливает максимум в гистограмме голосов, подавляя локальные максимумы [6]. Здесь имеем комбинированный вариант R_{VC} отношений для ХП, где параллельно используются их координаты и значения. Описанный тип отношений можно характеризовать как «поддержку» локальных решений, когда окрестность ХП подтверждает или опровергает фиксированное локальное решение.

Следующим примером эффективного использования отношений может быть способ, основанный на сравнении координат групп признаковых точек, представленных в двумерных аффинных базисах [1]. Этот метод получил название «геометрическое хеширование». Важной особенностью подхода есть инвариантность полученных характеристик локальных признаков (аффинных координат) при действии группы аффинных преобразований. Достоинства метода состоят в достаточно высоком быстродействии на этапе распознавания путем построения объемной хеш-таблицы признаков на этапе предварительной обработки. Здесь каждая признаковая точка на основе ее координат представляется в базисе трех других неколлинеарных между собой точек посредством аффинных координат — пары чисел $(\vartheta_1, \vartheta_2)$, по значениям которых создается хеш-таблица. Этим фактически устанавливается и затем сравнивается на множестве эталонов некоторое пространственное отношение каждой четверки особенных точек объекта. Итоговое голосование осуществляется на основе значений величин $(\vartheta_1, \vartheta_2)$, вычисленных для точек объекта. Метод построен на использовании только координатной информации (отношение типа R_C), однако его можно усилить и обобщить за счет применения значений признаков из R_V.

4. Меры отношений и вероятностные модели

Рассмотрим теперь вероятностную модель системы распознавания, построенной на голосовании отношений для ХП. Пусть содержательный смысл отношения состоит в том, что два или более признака голосуют за один и тот же эталон. Это отношение типа R_V.

Введем в рассмотрение полную группу событий $\Omega = \{\omega_0, \omega_1, ..., \omega_J\}$, где J — количество классов распознаваемого объекта, ω_0 соответствует классу неопределенности («не знаю»). Событие $\omega_j(z)$ состоит в том, что признак z относится к классу j: $\omega_j(z) = \{z \in j\}$. Считаем события ω_j несовместными для разных j, тогда полная группа событий $\Omega = \sum_{j=0}^{J} \omega_j$.

В качестве меры выполнимости события ω_j будем использовать значение *функции* $\xi(z, j)$ *соответствия* признака z классу j. Один из способов получения значения $\xi(z, j)$ состоит в вычислении расстояния от точки z до множества $Z(j)$ эталонных признаков класса j. Его можно определить через величину расстояния ρ в пространстве значений признака. Одним из вариантов есть вычисление расстояния ρ между признаком z и его ближайшим соседом z_j^{NN} в классе j [4]. Другой вариант основан на расстоянии до среднего значения признака класса, если величины средних для классов имеют значимые отличия [6].

Пусть по правилу ближайшего соседа $\xi(z, j) = \rho(z, z_j^{NN})$, где $z_j^{NN} = \arg \min\limits_{k=1,K(j)} \rho(z, z_j^k)$; $K(j)$ — количество элементов во множестве $Z(j)$; $z_j^k \in Z(j)$. Вид расстояния ρ определяется множеством значений и типом признака z. Для этого может быть использовано расстояние в векторном пространстве.

Вычисление функции соответствия, в свою очередь, устанавливает некоторое отношение между элементами множеств признаков Z и $Z(j)$. Атрибут этого отношения задается значением функции $\xi(z, j)$.

Учитывая, что значение меры соответствия естественным образом должно возрастать с уменьшением расстояния до эталона, в качестве $\xi(z, j)$ целесообразно использовать функции вида:

$$\xi(z, j) = \exp\{-\lambda \rho(z, z_j^{NN})\}, \tag{5}$$

где экспоненциальное преобразование классического расстояния ρ кроме получения необходимых свойств позволяет нормировать величину $\xi(z, j)$

в пределах отрезка $[0,1]$, что упрощает анализ значений $\xi(z,j)$. Величина λ отражает степень чувствительности построенной меры к изменению расстояния и выбирается в пределах фиксированного отрезка [4]. Появление события ω_0 будем характеризовать условием $\xi(z,j) \le \varepsilon_0$, где ε_0 — порог для величины незначительного сходства, которое не может быть засчитано ни одному из классов.

Вероятность $p(z,j)$ того, что признак z относится к классу j (метка признака z равна j), можно вычислить через соотношение значений мер, выражающих число случаев, благоприятных этому событию, и общего числа случаев в виде

$$p(z,j) = \frac{\xi(z,j)}{\sum_{i=0}^{J} \xi(z,i)} . \qquad (6)$$

Заметим, что здесь выполняется условие

$$\sum_{j=0}^{J} p(z,j) = 1 ,$$

так как в результате признак будет обязательно отнесен к одному из классов, то есть выполнится одно из событий ω_j.

Рассмотрим теперь событие, состоящее в том, что два признака z_t, z_ν будут отнесены к классу j. Оно состоит в одновременном выполнении событий $\omega_j(z_t)$ и $\omega_j(z_\nu)$, то есть в реализации бинарного отношения для признаков z_t, z_ν. Вероятность того, что оба признака принадлежат классу j, определяется как произведение вероятностей двух независимых событий в виде:

$$p(z_t, j \mid z_\nu, j) = \\ = p(z_t, j) p(z_\nu, j) = \frac{\xi(z_t, j)\xi(z_\nu, j)}{\sum_{i=0}^{J} \xi(z_t, i) \sum_{i=0}^{J} \xi(z_\nu, i)} . \qquad (7)$$

Вероятность (7) можно получить аналогично (6) как соотношение значений мер, соответствующих положительному исходу и общему числу вариантов для пар признаков. По аналогии с (7) вычисляются вероятности для трех и более признаков, реализующих k-арные отношения.

Для вероятностей вида (7) из-за отсутствия пересечения событий ω_j выполняется условие $\sum_{j=0}^{J} \sum_{i=0}^{J} p(z_t, j \mid z_\nu, i) = 1$.

Анализируя соотношения (6) и (7), можно заметить, что с увеличением количества признаков, входящих в объединяющее их отношение R_Z, вероятность их отнесения к одному и тому же классу уменьшается, так как произведение двух вероятностей всегда меньше или равно максимуму среди них. Этот факт можно трактовать следующим образом. Имеем, что для произвольной пары признаков и их меры соответствия вероятность отнесения обоих признаков к одному и тому же классу не больше, чем для каждого из признаков в отдельности. Это означает, что надежность распознавания, основанного на паре признаков, будет выше, чем для одного признака.

Заметим, что значения (6), (7) являются функциями от нескольких параметров: вида функции соответствия ξ, способа формирования и значений признаков эталонов, величин признаков распознаваемого объекта. Основываясь на этом факте, величины (6), (7) можно использовать для оценки качества системы распознавания, построенной на сопоставлении множеств локальных признаков. Для конкретного набора эталонов путем вычисления, например, значения (7), можно оценить вероятность «совпадения» между собой пар признаков внутри множества эталонов. Если эта вероятность велика, значит, необходимо менять либо эталоны, либо способ вычисления признаков, либо вид функции соответствия, либо увеличить количество признаков, формирующих отношение. Фактически таким способом определяются предельные возможности системы распознавания. На этой основе для конкретного множества (базы) эталонов можно подобрать необходимое для голосования количество и вид признаков, используя фиксированную функцию ξ.

На основе выражения (6) можно подсчитать вероятность того, что ровно m признаков из общего количества s будут отнесены к классу j. Это событие, которое обозначим как H_m, соответствует m-арному отношению на множестве признаков. Учитывая, что вероятности $p_k = p(z,j)$, кроме номера класса, зависят от значения признака z (эта зависимость отражена индексом k), вычисление вероятности события H_m отличается от схемы Бернулли и определяется более сложным распределением. Если обозначить через $(i[1],...,i[m])$ выборку объема m из последовательности чисел от 1 до s, то вероятность события H_m равна

$$P(H_m) = \sum_{(i[1],...,i[m])} p_{i[1]} \dots p_{i[m]} q_{r[1]} \dots q_{r[s-m]} , \qquad (8)$$

где $p_{i[\nu]}$ — значения p_k с соответствующими номерами; $(r[1],...,r[s-m])$ — дополнение выборки $(i[1],...,i[m])$ до s элементов; $q_{i[\nu]} = 1 - p_{i[\nu]}$. Тогда вероятность отнесения m и более признаков к классу j исходя из (8) равна

$$Q(p_k, m) = \sum_{t=m}^{s} P(H_t) . \qquad (9)$$

Если же предположить, что значение вероятности p_k одинаково для всех признаков из Z, то есть выполнено условие $\forall k \; p_k = p$, то вероятность Q в соответствии со схемой Бернулли равна

$$Q(p,m) = \sum_{t=m}^{s} C_s^t p^t q^{s-t} , \qquad (10)$$

где $q = 1 - p$. Учитывая сложность вычисления $Q(p_k, m)$ в виде (9), связанную с перебором всевозможных кортежей из m элементов признакового пространства, в качестве верхней оценки искомой вероятности вместо (9) можно использовать значение $Q(p,m)$ из (10), положив, например, $p = \max_k p_k$. Другой вариант оценки можно построить на усреднении значений p_k.

Вероятности для отношений признаков, полученные в соответствии с (6)-(10), могут быть положены в основу работы системы распознавания путем голосования [4]. На базе (6)-(10) вычисляется апостериорная вероятность отнесения множества ХП к одному из классов. Решение принимается по максимуму апостериорной вероятности.

5. Некоторые практические аспекты

В плане особенностей практической реализации предложенного подхода отметим следующее.

Исходные данные для системы распознавания можно представить в виде совокупности матриц эталонов и распознаваемого объекта

$$Z(j) = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ . & . & . \\ z_{K[j]1} & \cdots & z_{K[j]n} \end{bmatrix}, \; j = \overline{1,J}, \quad Z = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ . & . & . \\ z_{s1} & \cdots & z_{sn} \end{bmatrix}.$$

Матрицы в общем случае имеют разное количество строк, а каждая строка состоит из фиксированного числа n элементов, соответствующих признаку z. Отношения устанавливаются между одинаковыми по числу подмножествами строк. Например, бинарные отношения предполагают установление соответствия для пар строк. Распознавание сводится к сопоставлению строк матриц или их отношений и установлению матрицы из $\{Z(j)\}$, наиболее близкой в смысле критерия распознавания к матрице Z. Степень соответствия отдельных строк выражается значением функции $\xi(z, j)$. Вероятностные характеристики распознавания могут быть получены из (6)-(10).

При реализации схемы распознавания для сходства комбинаций строк должны быть установлены пороги, равные половине количества голосующих строк или их комбинаций (максимальное правдоподобие), либо один общий порог для всех эталонов.

Учитывая, что сравнение осуществляется на основе полного перебора подмножеств строк, при сопоставлении необходимо использовать эффективные способы поиска, например, хеш-структуры.

Выводы

Отношения, построенные на множестве ХП распознаваемого объекта и эталонов, являются мощным источником информации, позволяющим сформировать признаки более высокого уровня. Применение этой новой высокоуровневой системы признаков в целом позволяет повысить надежность функционирования и расширить возможности распознающих систем, основанных на голосовании.

Научная новизна проведенных исследований состоит в формализации процесса применения отношений при голосовании признаков и построении вероятностных моделей, позволяющих оценить качество распознавания.

Практическая значимость работы заключается в построении конкретных схем распознавания, обладающих повышенными характеристиками надежности и селективных свойств по сравнению со схемами независимого голосования признаков.

Дальнейшее развитие подхода может быть связано с выбором оптимального для конкретных условий распознавания количества характерных признаков, составляющих отношение.

Список литературы: 1. *Шапиро Л., Стокман Дж.* Компьютерное зрение: Пер. с англ. — М.:Бином, 2006. – 752 с. 2. *Путятін Є.П., Гороховатський В.О., Матат О.О.* Методи та алгоритми комп'ютерного зору: Навч. посібник. — Харків: ТОВ «Компанія СМІТ», 2006. – 236 с. 3. *Путятин Е.П., Гороховатский В.А., Кузьмин С.В.* Распознавание изображений в пространстве инвариантных локальных признаков // Радиоэлектроника и информатика. — 2006, №1(32). — С. 69–73. 4. *Kim S., Kweon I.-S.* Biologically motivated perceptual feature: Generalized robust invariant feature. — In Asian Conference of Computer Vision (ACCV-06), 2006, P. 305-314. 5. *Schmid C., Mohr R.* Local Grayvalue Invariants for Image Retrieval. — IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997 — 19(5): 530-534 p. 6. *Форсайт Д., Понс Ж.* Компьютерное зрение. Современный подход: Пер. с англ. — М.: Изд. дом «Вильямс», 2004. — 928 с. 7. Нечеткие множества и теория возможностей. Последние достижения / Под ред. Р. Ягера. — М.: Радио и связь, 1986. – 408 с. 8. *Шрейдер Ю.А., Шаров А.А.* Системы и модели. — М.: Радио и связь, 1982. — 152 с.

Поступила в редколлегию 4.04.2008