УДК 004.932.2:004.93’1

МЕТРИЧЕСКАЯ КЛАССИФИКАЦИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДА БЛИЖАЙШИХ СОСЕДЕЙ НА ОСНОВЕ ОПИСАНИЙ В ВИДЕ МНОЖЕСТВ ПРИЗНАКОВ

Н.В. ГОРОХОВАТСКАЯ

Исследуются возможности и свойства метода ближайших соседей в задаче метрической классификации видео-объектов, представленных описаниями в виде множества признаков. Приводятся результаты вычислительных экспериментов для различных баз данных, анализируются варианты и особенности применения подхода.

Ключевые слова: метрическая классификация, метод ближайшего соседа, описание в виде множества признаков, вероятность правильной классификации.

ВВЕДЕНИЕ

Решение прикладных задач распознавания образов часто сводится к отнесению анализируемого описания к некоторому классу [1–4]. Одним из эффективных практических подходов к распознаванию видео-объектов в компьютерном зрении есть сопоставление анализируемого описания с конечным набором эталонных описаний. Эти подходы называют метрическими [5]. Основу метрического подхода составляет построение описаний и синтез меры подобия для оценки их близости. Прикладные задачи такого плана встречаются в системах распознавания движущихся объектов [6] и метрического поиска в базах видеоинформации [7].

В ряде применений описание видео-объекта представляют как множество векторов с вещественными компонентами [2,3]. Это могут быть структурные описания в виде множеств дескрипторов характерных признаков (ХП), сформированных в результате применения технологий типа SIFT, SURF [2]. В системах распознавания на основе проекционных признаков объекты описывают в виде множества числовых функций или векторов. В этих пространствах метрическая классификация сводится к вычислению и оптимизации значения меры для сравнения конечных множеств на основе анализа подобия их элементов.

1. ПОСТАНОВКА ЗАДАЧИ

Традиционная задача классификации состоит в следующем [5]. Имеется пространство объектов и конечное множество имен классов J. На множестве Z задана функция расстояния $\rho_Z : Z \times Z \to [0,\infty)$. Существует зависимость $j : Z \to J$, значения которой известны только на объектах заданного подмножества $Z^h \in Z$ в виде набора $Z^h = (z_i, j_i)_{i=1}^{h}$, $z_i \in Z^h$, $j_i = j(z_i), j_i \in J$. Требуется построить отображение (метод классификации) $c : Z \to J$, аппроксимирующее зависимость $j(z)$ на все множество Z.

В соответствии с этим рассмотрим описание V анализируемого визуального объекта в виде конечного множества $V = \{v_i\}_{i=1}^{m}$ элементов, где m – число элементов в описании, $V = S \times U$, $v_i = (s_i, u_i)$, S – множество пространственных характеристик, $s_i \in S$, $u_i \in U$, $U \subseteq R^n$, R^n – n-мерное вещественное векторное пространство. Например, для признаков SURF значение n=64, $u_i \in [-1,1]$. Элементы $s_i \in S$ – это также числовые векторы конечной размерности, как правило, значительно меньшей n. В частности, для структурных признаков – это координаты $s_i = (x_i, y_i)$ или более сложные признаки типа аффинных инвариантов $s_i = f(\{x_i, y_i\})$ (f – некоторая функция построения признаков) [8]. Для проекционных признаков s_i равно непосредственно углу проецирования φ_i, $s_i = \varphi_i$, или некоторой функции $s_i = f(u_i, \varphi_i)$ от угла и значения признака [2].

Метрическая классификация объекта на основе описания V есть отображение $\Re : V \to \{V(j)\}$, где $\{V(j)\}_{j=1}^{J}$ – конечное множество эталонных описаний (база видеоданных), и реализуется через решение задачи оптимизации

$$c(V) = \arg \max_{j \in [1,\ldots,J]} \vartheta[V, V(j)], \qquad (1)$$

где $\vartheta \in R^1$ – мера подобия, $c(V)$ – класс объекта. Известные методы сопоставления описаний сосредотачиваются на вычислении и оптимизации по параметру j значения доли элементов описания V, нашедших соответствие (эквивалентность) в j-м эталоне [6-9].

Для определенности рассмотрим элементы пространства U. В качестве критерия при сопоставлении отдельных элементов $u_i \in U$ используем некоторую метрику $\rho(u_i, u_k)$, оценивающую различие u_i, u_k, $u_i \in U_1$, $u_k \in U_2$. Основным источником информации для анализа подобия и классификации есть матрица $\wp = \{\rho(u_i, u_k)\}$ соответствующих размеров.

Двумя главными факторами, от которых в решающей степени зависит качество классификации, как известно, являются подобие элементов внутри классов и различия между элементами разных классов [3, 4]. В статистической теории классификации эти факторы отражаются значениями критериев внутриклассового разброса и межклассовых расстояний. В связи с этим будем различать две принципиально разных модели метрической классификации видео-объектов:

а) элементы внутри одного эталонного класса однородны (примерно одинаковы), и их различие связано лишь со случайными отклонениями в виде флуктуационных помех аддитивного типа;

б) эталонные классы состоят из конечных наборов данных, причем отдельные элементы могут быть близкими для разных эталонных классов.

При решении задач на основе второй модели учитывается состав эталонных множеств, что существенно сложнее и требует дополнительной обработки.

Задачи компьютерного зрения в большей степени описываются второй моделью и специфичны в том плане, что распределения признаков в пространстве U обычно неизвестны и бывают достаточно близкими для разных классов, а описания классов заданы в виде непосредственного представителя (эталона) или конечного множества как описания эталона. Именно для таких ситуаций представляется эффективным применение метода ближайших соседей [1,5]. Использование решений, опирающихся на представителей из эталонных классов, более эффективно, чем решение, основанное на одном представителе (хотя и самом близком в некотором аспекте). Появляется новая возможность интегрированного анализа решений соседей, что может быть использовано для улучшения качества классификации. Метод ближайших соседей может быть применен как к множеству элементов $u_i \in U$, так и непосредственно к множеству описаний $U(j)$ базы распознавания.

Цель работы – применение метода ближайших соседей при решении задач метрической классификации видео-объектов по их описаниям в виде множеств и исследование свойств метода в сравнительном аспекте с другими подходами.

Задачи исследования состоят в формализации процесса применения метода ближайших соседей в задачах классификации, проведении экспериментальных исследований для определения эффективности работы метода, а также анализе вариантов метода с точки зрения качества классификации.

2. ФОРМАЛИЗАЦИЯ И ПАРАМЕТРИЗАЦИЯ МЕТОДА КЛАССИФИКАЦИИ

Метод k ближайших соседей (КБС) относят к непараметрическим решающим процедурам, где без явного применения вероятностных оценок строятся решающие функции об отнесении объекта к классу [3, 5]. Суть метода КБС применительно к нашей задаче заключается в выделении среди всех имеющихся наблюдений базы некоторого числа k априорно известных объектов (k ближайших соседей), похожих на новый неизвестный ранее объект с точки зрения расстояния $\rho(u_i,u)$ между элементами. На основе значений классов ближайших соседей выносится решение относительно нового объекта.

Если в пространстве, где осуществляется классификация, рассмотреть ячейку (шар) некоторого объема вокруг анализируемого объекта, в которой находится d объектов известных классов, причем d_i из которых принадлежит классу i, $d=\sum_i d_i$, то значение d_i/d можно считать оценкой апостериорной вероятности того, что состояние природы для анализируемого объекта есть класс i. Размер ячейки есть либо функцией от общего числа имеющихся данных, либо функцией от числа элементов, попавших в ячейку с заданным размером. Решение по максимуму величины d_i/d минимизирует уровень ошибки [3].

Определим имеющуюся базу $U^M=\{U(j)\}_{j=1}^J$ в виде конечного множества $Z^M=\{u_i,c_i\}_{i=1}^M$ пар (u_i,c_i), $u_i \in U^M$, $c_i \in J$, где компонента u_i определяет значение признака, а c_i – значение его класса. Общее число M элементов базы определим как $M=\sum_{j=1}^J m(j)$, где $m(j)$ – число элементов базы, принадлежащих классу с номером j. Классифицируемый объект u в соответствии с методом КБС будет отнесен к классу, которому принадлежит большинство из его соседей – k ближайших к нему объектов из множества U^M, которое можно трактовать как обучающую выборку.

Для произвольного объекта u расположим объекты выборки $u_i \in U^M$ в порядке возрастания их расстояний до u:

$$\rho(u,u_{1u}) \le \rho(u,u_{2u}) \le \ldots \le \rho(u,u_{Mu}), \tag{2}$$

где через u_{qu} обозначается объект, который является q-м соседом элемента u в соответствии со значением ρ. Таким образом, произвольный объект порождает свою нумерацию $u_{1u},u_{2u},\ldots,u_{Mu}$ выборки U^M и соответствующую нумерацию $c_{1u},c_{2u},\ldots,c_{Mu}$ значений их классов.

Метод ближайшего соседа относит классифицируемый объект к классу, которому принадлежит ближайший объект из U^M, в соответствии с правилом: $c(u,Z^M)=c_{1u}$, где $c(u,Z^M)$ – функция определения класса объекта с использованием обучающего множества Z^M. Более общий метод КБС есть

$$c(u,Z^M)=\arg\max_{j\in J}\sum_{i=1}^M [c_{iu}=j]\,\omega(i,u), \tag{3}$$

где $[c_{iu}=j]=\begin{cases}1, \text{ класс } u_{iu} \text{ равен } j,\\ 0, \text{ иначе},\end{cases}$ $\omega(i,u)$ – весовая функция, которая отражает степень важности i-го соседа для классификации. В литературе функцию $\omega(i,u)$ называют функцией окна, в частности, она используется для интерполяции плотности распределения [3]. Полагают, что функция $\omega(i,u)$ неотрицательна и не возрастает по i. По-разному задавая $\omega(i,u)$, получают различные варианты метрической классификации [1, 4, 5]. Этой же схемой описывают методы парзеновского окна и потенциальных функций. Функция $\omega(i,u)$, кроме степени важности соседа $u_{iu} \in U^M$, часто отражающейся величиной параметра i, может выражать также и значимость $\alpha(u_{iu})$ этого элемента внутри эталонного описания $U(j)$. Особенно это свойственно описаниям видео-объектов, где отдельные признаки для визуального анализа могут быть намного важнее других, и их участие в решении более значимо. Схема применения метода КБС для элементов описаний приведена на рис. 1.

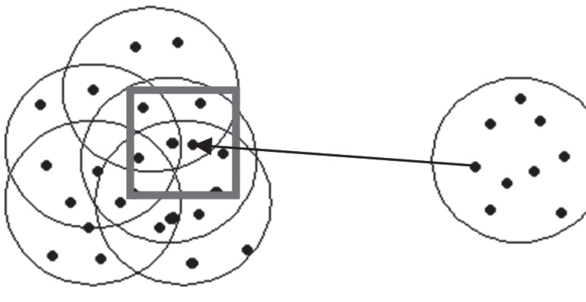

Рис. 1. Схема применения метода ближайших соседей

После классификации всех $u_i \in U$ решение $c(U, Z^M)$ о классе объекта примем по большинству среди классов $u_i \in U$

$$c(U,Z^M) = \arg\max_{j \in J} \sum_{i=1}^{m} [c(u_i, Z^M) = j], \qquad (4)$$

где m – число элементов в описании U. Формулировка (4) соответствует задаче (1).

Ясно, что в общем случае достоверность решения (4) определяется числом γ элементов, которые формируют максимум. Значение $\gamma \in [1, m]$. Чем больше γ, тем в большей степени можно доверять принятому решению о классе объекта. В то же время, осуществление классификации в условиях искажений, связанных с потерей элементов, приводит к необходимости решения при $\gamma << m$. Обычно используют некоторый порог δ_γ для величины γ. Можно также ввести ограничения на число элементов γ_j, представляющих отдельный эталон при решении (4).

Для метода КБС важно, чтобы среди объектов из U^M отсутствовали выбросы, иначе классификация может потерять интерпретируемость. Поэтому приходится применять специальные методы отбора эталонов, т.е. тщательно анализировать множество U^M. Предварительная обработка U^M осуществляется в целях отсеивания неинформативных, периферийных, шумовых выбросов, что одновременно существенно снижает и временные затраты, т.к. метод КБС опирается на использование всех элементов выборки U^M, и размерность U^M есть одним из проблемных вопросов [5,9]. Другим немаловажным способом сокращения объема вычислений есть использование эффективных структур данных для организации быстрого поиска ближайших соседей, например, kd-деревьев [4].

Принципиально важным для метода КБС есть подбор k. Этот этап является одним из основных при реализации метода. В ряде прикладных задач добиваются, чтобы число k было как можно меньше. С другой стороны, значение k определяет интегральные способности метода с точки зрения борьбы с помехами, и с увеличением k растет помехозащищенность, но и одновременно теряется свойство выявления локальных особенностей групп элементов из U^M. На практике k определяют по критерию скользящего контроля методом исключения объектов по одному [5].

В значительной мере на эффективность метода КБС влияет также и выбор метрики ρ. Традиционно, особенно в пространстве числовых векторов, применяют евклидово либо манхэттенское расстояния, хотя оптимизация выбора метрики с точки зрения согласования с типом используемых данных может значительно улучшить результаты классификации.

Проблемным вопросом метода КБС являются пространства большой размерности, где все объекты примерно одинаково далеки друг от друга, и выбор соседей становится практически произвольным [9]. Проблема решается путём отбора относительно незначительного числа информативных признаков путем селекции имеющегося набора. Например, в алгоритмах вычисления оценок строится множество наборов признаков (опорных множеств), для каждого формируется своя функция близости, затем по всем функциям производится голосование.

3. ВАРИАНТЫ И ОСОБЕННОСТИ ПРИМЕНЕНИЯ МЕТОДА

Рассмотрим некоторые подходы и особенности использования метода при вычислении подобия описаний.

Учет весов соседей. Модификации метода КБС связаны с обогащением моделей (3), (4) путем введения дополнительных параметров. Это могут быть априорно заданные веса элементов из U^M (объекты базы видеоданных имеют разную полезность с точки зрения классификации новых объектов), а также веса, вычисленные в процессе классификации и зависящие от распознаваемого объекта.

Так как в основе КБС лежит предположение о подобии близко расположенных объектов, можно считать, что ближайшие соседи оказывают различное влияние на результат. Например, допускается, что чем ближе сосед к классифицируемому элементу, тем существенней его влияние. Такое предположение реализуется путем введения весов ближайших соседей, которые учитывают значение близости [10]. Пример весовой функции:

$$\beta(i,u) = \frac{\exp[-\rho(u_i,u)]}{\sum_i \exp[-\rho(u_i,u)]}, \qquad (5)$$

где $\exp[-\rho(u_i,u)]$ – функция экспоненты от расстояния, $u_i \in U^M$. Веса, определенные в соответствии с (5), удовлетворяют условию $\sum_i \beta(i,u) = 1$.

Вес $\beta_j(u)$ классифицируемого элемента за класс j в общем виде представляется как функция $\beta_j(u) = f(k, \delta_\rho, \beta(\{u_i\}_j), \{\alpha_i\}_j)$ от числа k используемых соседей, порога δ_ρ для эквивалентности элементов, веса $\beta(\{u_i\}_j$ в зависимости от номера i в списке соседей, вектора априорных весов $\{\alpha_i\}_j$, который учитывает значимость эталонного элемента-соседа внутри класса j.

Ограничение области решений. Обозначим для конкретных k, u множество ближайших соседей как $U(k,u) \subseteq U^M$. Традиционная схема основана на определении значения класса по наиболее мощному подмножеству $U(k,u)$. Модификация «с отказами» предполагает, что элемент u будет классифицирован только тогда, если все элементы $u_i \in U(k,u)$ принадлежат одному классу, в противном случае принимается решение об отказе от классификации [1]. Другими словами, решение выносится только в том случае, если все k элементов принадлежат одному и тому же классу. Понятно, что такой подход существенным образом повышает достоверность классификации, однако на практике предположение о доминировании представителей одного класса в выбранной окрестности классифицируемого объекта является не совсем приемлемым и обоснованным. Как вариант, решение может быть принято, если не менее чем $k' = \lambda k$, $\lambda \in [0,1]$ элементов принадлежат одному классу. При этом проблема выбора λ трудно формализуема и сходна с определением числа k. Как вариант, λ определяет долю от k, например, $\lambda = 0{,}5$.

Модели «элемент-множество» и «множество-множество». Используем правило КБС как основу для построения мер подобия, если эталон $U(j)$ рассматривать как целостный элемент базы (множество). Обозначим $U(k,u,j)$ множество КБС из класса j при фиксированных k, $u \in U$. Классификацию можно осуществить либо для элемента $u \in U$:

$$c(u) = \arg\max_{j \in J} \theta[U(k,u,j)], \qquad (6)$$

либо для множества U в целом

$$c(U) = \arg\max_{j \in J} \chi[U, U(k,u,j)], \qquad (7)$$

где функционал $\theta[U(k,u,j)]$ отражает подобие элемента классу, а функционал $\chi[U, U(k,u,j)]$ – подобие множеств U и $U(j)$. Особенностью соотношений (6), (7) есть то, что класс элементов обучающей выборки априорно задан. Модель (7) в силу использования принципа групповых решений может оказаться надежнее в условиях помех, чем (6). Рассмотренные схемы сопоставления проиллюстрированы на рис. 2.

Соотношение (7), в отличие от классического применения метода КБС типа (6), связанного построением аппроксимации для плотности распределения, не имеет непосредственной интерпретации, хотя и использует принцип анализа ближних соседей.

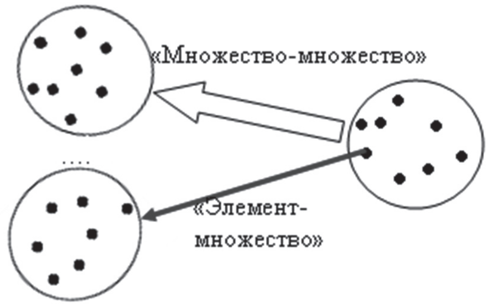

Рис. 2. Схемы сопоставления

4. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В первой части имитационных экспериментов выполнено моделирование метрического подхода на основе КБС применительно к описаниям в виде множеств случайных векторов. База эталонов сформирована следующим образом: сгенерировано 5 независимых классов, каждый из которых включает 10 векторов размерностью 64. Компоненты вектора с $k = 1,3,5$ считаем независимыми случайными величинами с нормальным распределением $N(\mu, \sigma^2)$, где μ – математическое ожидание, σ^2 – дисперсия. Таким образом, классы задавались параметрами μ_i, $i = \overline{1,5}$ при общем для базы фиксированном значении σ^2. В качестве классифицируемых описаний (10 векторов) сформированы векторы по аналогичному принципу. Действие помехи задавалось величиной дисперсии шума σ_s^2, действующего на классифицируемое описание, значение σ_s^2 увеличивалось, начиная с 0. Цель экспериментов – оценка вероятности правильной классификации в зависимости от параметра k, а также от уровня шума σ_s^2. При конкретно выбранных значениях $\mu_i = \{0{,}17; 0{,}33; 0{,}50; 0{,}67; 0{,}83\}$ до уровня $\sigma_s^2 = 0{,}15$ обсуждаемый метод (схема «элемент-множество») для $k = 1,3,5$ обеспечивает безошибочную классификацию. При $\sigma_s^2 > 0{,}15$ вероятность правильной классификации начинает снижаться. При $\sigma_s^2 = 0{,}17$ значения вероятностей составили: 0,978 ($k = 1$), 0,991 ($k = 3$), 0,994 ($k = 5$), а при $\sigma_s^2 = 0{,}2$ – 0,91 ($k = 1$), 0,935 ($k = 3$), 0,943 ($k = 5$). Эта картина наблюдалась для разных метрик (Евклидова, манхэттенская, Чебышева). Как видим, увеличение k повышает вероятность правильной классификации, хотя при $k = 3$ и $k = 5$ отличается незначительно. Преимущества метода при $k > 1$ по сравнению с вариантом $k = 1$ можно объяснить более общей и универсальной схемой анализа данных, для которой $k = 1$ является частным случаем.

Другая серия испытаний проведена на множестве инвариантных признаков SIFT (размер вектора – 128) для изображений базы рыб [8] в целях сравнения эффективности двух методов, основанных: 1) на принципе голосования элементов объекта и 2) на основе КБС (с весами вида (5) и без них), в плане оценки их качества функционирования при влиянии аддитивных помех. Помеха моделировалась путем добавления к эталонным компонентам вектора SIFT значения нормально распределенной случайной величины с нулевым матожиданием и фиксированной дисперсией σ_a^2. Значение σ_a^2 определяет соотношение сигнал–шум. В обоих методах использовалась модель сопоставления элемент-множество, параметр $k=3$. Решение о классе в первом методе принималось по максимуму числа набранных голосов, во втором – по максимуму значения суммы весов ХП или наиболее часто встречающегося класса элементов объекта. В результате моделирования установлено, что до уровня $\sigma_a^2=0{,}1$ все три варианта обсуждаемых методов обеспечивают безошибочную классификацию. С дальнейшим повышением σ_a^2 вероятность снижается и при $\sigma_a^2=0{,}12$ достигает значения 0,97, а для варианта со взвешиванием – 0,94.

Эксперименты показали, что правило формирования весов решающим образом влияет на результаты классификации и должно быть как можно лучше согласовано с моделью данных, отражающую структуру и содержание эталонных множеств. Это в большей степени относится к описаниям SIFT, которые более неоднородны по составу, чем случайные вектора в первой серии экспериментов. Применение принципа КБС более эффективно в ситуациях, когда описания содержат повторяющиеся элементы, потеря отдельных из которых из-за действия помех компенсируется другими.

ЗАКЛЮЧЕНИЕ

Метод ближайших соседей относится к непараметрическим подходам классификации и обладает рядом важных достоинств, в частности, отсутствием предположений о типе распределений признаков. Кроме того, метод считается устойчивым к аномальным выбросам, так как вероятность попадания аномальной записи в число ближайших соседей на практике весьма мала. Даже если это происходит, то влияние на голосование при $k>2$ незначительно, и, следовательно, результат классификации не искажается.

Метрическая классификация на основе принципа КБС более эффективна в ситуациях структурированных описаний и может быть универсально применена для разнообразия признаковых пространств.

В статье впервые показано, что применение принципа ближайших соседей в целях усовершенствования моделей метрической классификации для описаний, представленных множествами, обеспечивает высокую вероятность правильной классификации в условиях помех. Весовые коэффициенты соседей должны быть согласованы с объектом распознавания, эффективность модификации повышается для структурированных описаний.

Практически важным результатом есть экспериментальное подтверждение предложенных модификаций как для множеств случайно сформированных векторов, так и для реальных видеоданных, полученных детекторами характерных признаков. Это говорит о целесообразности применения предложенных подходов для широкого класса прикладных задач.

Перспективой развития исследований может быть построение модификаций мер подобия для рассмотренного вида описаний, обладающих новыми свойствами в плане эффективности распознавания.

Литература.

[1] *Патрик Э.* Основы теории распознавания образов/ Э. Патрик; пер. с англ. под ред. Б. Р. Левина.– М.: Сов.радио, 1980. – 408 с.

[2] *Шапиро Л.* Компьютерное зрение/ Л. Шапиро, Дж. Стокман; пер. с англ.– М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.

[3] *Duda R.O.* Pattern classification/ Duda R.O., Hart P.E., Stork D.G. – 2ed., Wiley, 2000.–738p.

[4] Прикладная статистика: Классификация и снижение размерности: Справ. изд. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин; под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1989. — 607 с.

[5] *Воронцов К.В.* Лекции по метрическим алгоритмам классификации [Электронный ресурс]/ Режим доступа: http://www.ccas.ru/voron/download/MetricAlgs.pdf.

[6] *Путятин Е.П.* Нормализация изображений: достижения и нерешенные задачи/ Е.П. Путятин // Бионика интеллекта. – 2007. –Вып. 1(66). – С. 5-16.

[7] *Машталир В.П.* Точечно-множественные методы обработки информации / В.П. Машталир.– Х.: Бизнес Информ, 2001. – 199 с.

[8] *Гороховатский В.А.* Структурное распознавание изображений на основе моделей голосования признаков характерных точек / В.А. Гороховатский, Е.П. Путятин // Реєстрація, зберігання і обробка даних.– 2008.–Т.10. – №4.– С.75-85.

[9] *Рассел С.* Искусственный интеллект: современный подход; 2-е изд./ С. Рассел, П. Норвиг; пер. с англ.– М.: Изд. дом «Вильямс», 2006. – 1408 с.

[10] *Shepard, Donald.* A two-dimensional interpolation function for irregularly-spaced data/ D. Shepard// Proceedings of the 1968 ACM National Conference. – 1968. – p. 517–524.

Поступила в редколлегию 27.01.2012

Гороховатская Наталия Владимировна, аспирантка кафедры Информатики ХНУРЭ. Область научных интересов: модели, методы, информационные технологии анализа и распознавания многомерных данных.

УДК 004.932.2:004.93'1

Метрична класифікація з використанням метода найближчих сусідів на основі опису у вигляді множин / Н.В. Гороховатська // Прикладна радіоелектроніка: наук.-техн. журнал. – 2012. Том 11. № 1. – С. 42-47.

Досліджуються можливості і властивості методу найближчих сусідів в задачі метричної класифікації відео-обє'ктів, опис яких представлений у вигляді множини ознак. Наводяться результати обчислювальних експериментів для різних баз даних, аналізуються варіанти та особливості застосування підходу.

Ключові слова: метрична класифікація, метод найближчих сусідів, опис у вигляді множини ознак, ймовірність правильної класифікації

Іл. 2. Бібліогр.: 10 найм.

UDC 004.932.2:004.93'1

Metric classification using the method of the nearest neighbours on the basis of descriptions in the form of sets of attributes / N.V. Gorohovatskaya // Applied Radio Electronics: Sci. Journ. – 2012. Vol. 11. № 1. – P. 42-47.

Features and properties of the method of the nearest neighbours in the problem of metric classification of video objects, represented by descriptions as a set of attributes are investigated. The results of numerical experiments for different databases are presented, options and features of the approach are analyzed.

Keywords: metric classification, nearest neighbour, description as a set of attributes, probability of correct classification.

Fig. 2. Ref.: 10 items.