

Т. В. ПОЛЯКОВА, аспирант каф. информатики ХНУРЭ, Харьков;
Ю. А. КУЛИКОВ, аспирант каф. информатики ХНУРЭ, Харьков;
В. А. ГОРОХОВАТСКИЙ, д-р. техн. наук, проф. ХИБД НБУ, Харьков

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРИЗНАКОВЫХ ПРОСТРАНСТВЕННЫХ СТРУКТУР

Досліджуються нові ознакові простори на множині координат структурних ознак зображень. Ознакові просторові структури відображають взаємне розташування та зв'язок окремих частин відео-об'єкта. Наведені результати розробки методів, які за рахунок скорочення надмірності описів знижують об'єм обчислювальних витрат. Обговорюються результати експериментів по класифікації зображень на основі запропонованих методів.

Ключові слова: класифікація зображень, характерні ознаки, структурний опис, афінні інваріанти, ознакові просторові структури, голосування, обчислювальні витрати, ймовірність правильної класифікації.

Исследуются новые признаковые пространства на множестве координат структурных признаков изображений. Признаковые пространственные структуры отражают взаимное расположение и связь отдельных частей видео-объекта. Приведены результаты разработки методов, которые за счет сокращения избыточности описания снижают объем вычислительных затрат. Обсуждаются результаты экспериментов по классификации изображений на основе предложенных методов.

Ключевые слова: классификация изображений, характерные признаки, структурное описание, аффинные инварианты, признаковые пространственные структуры, голосование, вычислительные затраты, вероятность правильной классификации.

There have been researched new feature spaces on the set of coordinates of structural features of images. The feature spatial structures reflect relative position and relationship of individual parts of the video object. The development results of the methods, that reduce the amount of computing costs by decrease descriptions redundancy, have been put. The results of experiments on image classification based on the proposed methods are discussed.

Keywords: image classification, characteristic features, structural description, affine invariants, feature spatial structures, voting, computing costs, probability of correct classification.

Введение. Структурные методы анализа видеоданных в настоящее время являются основными при решении прикладных задач компьютерного зрения [1]. Особую ценность для задач распознавания видео-объектов имеет геометрическая информация, которая содержится в координатах характерных признаков (ХП). ХП являются основой построения структурных описаний и отражают локально-пространственные свойства объектов в отдельных точках [2–7]. Человеческое зрение редко прибегает к анализу функции яркости визуальных объектов, в основном опираясь при решении практических задач исключительно на геометрические характеристики объектов, которые можно вычислить непосредственно на основе множества координат ХП.

Исключительно на основе координат ХП можно построить признаки более сложного плана, к которым относятся аффинные инварианты (АИ),

которые можно считать частным случаем более общих признаков пространственных структур (ППС). Значения АИ непосредственно отражают геометрическую форму объекта и обладают свойствами инвариантности к аффинным преобразованиям на плоскости. С другой стороны, АИ позволяют сохранять устойчивость при действии пространственных помех, т.к. наряду с ХП отражают локальные особенности взаимного расположения подмножеств точек объекта. Схема преобразования данных с использованием ППС представлена на рис.1.

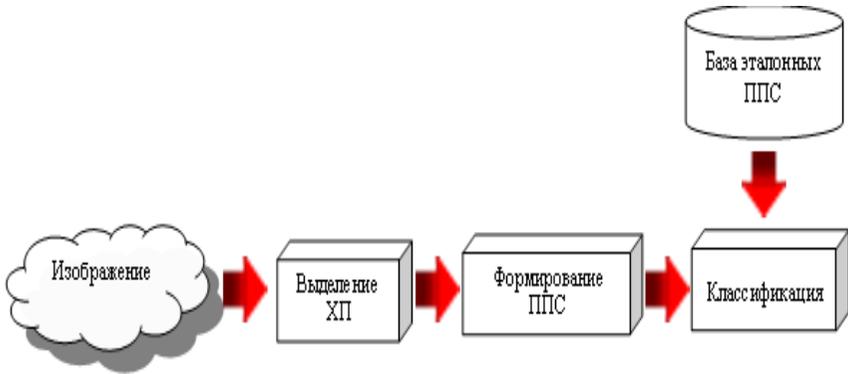


Рис. 1 – Преобразование данных в процессе структурной классификации

Описание визуального объекта формально представим в виде конечного множества $U = \{u_i\}_{i=1}^n$ элементов u_i , которые опишем кортежем:

$$u_i = \langle d_i, x_i, y_i, \{\alpha_q\}_{q=1}^{m_i} \rangle, \quad (1)$$

где $i = \overline{1, n}$, n – количество ХП в описании U , $\alpha_q = (\zeta_q, \eta_q)$ – значения АИ (или аффинных проекций) [5,6], $m_i = 6C_{n-1}^3$ – количество величин α_q для i -го ХП, $d_i \in R^k$ – k -мерный числовой вектор-дескриптор атрибутивных признаков, инвариантный к геометрическим преобразованиям из некоторой группы, который отражает свойства изображения в локальной окрестности точки (x_i, y_i) и полученный путем применения современных методов SIFT, SURF [4,7]. В общем случае каждое из множеств $\{x_i, y_i\}, \{d_i\}, \{\alpha_q\}$ есть мультимножество [9].

Координаты $c_i = (x_i, y_i)$ для элемента $u_i \in U$ могут быть представлены в некотором базисе $\{e_0, e_1, e_2\}$ в виде пары (ζ, η) АИ [1]:

$$c_i = e_0 + \zeta(e_1 - e_0) + \eta(e_2 - e_0). \quad (2)$$

В общем случае при перестановке порядка элементов базиса для одного ХП значения (ζ, η) изменяются. Обозначим $\gamma = 1 - \zeta - \eta$ и запишем (2) в виде [6]:

$$c_i = \gamma e_0 + \zeta e_1 + \eta e_2. \quad (3)$$

Путем представления (3) после упорядочивания (γ, ζ, η) значение числа АИ m_i сокращается в 6 раз, т.е. в описании (1) можно принять $m_i = C_{n-1}^3$, $\alpha_q = \{\zeta_q, \eta_q, \gamma_q\}$. Это значительно снижает общие вычислительные затраты при анализе и сопоставлении совокупностей АИ.

Введем определение пространства A признаков пространственных структур как подпространства $A \subset R^3$, рассматривая его как обобщение множества АИ, и исследуем свойства этого пространства применительно к задаче инвариантной классификации в базах видеоданных. Для определенности будем рассматривать задачу классификации видео-объектов на конечном множестве классов как вычисление и оптимизацию меры подобия их описаний [1,2].

В частности, используем расстояние между значениями ППС в евклидовой метрике:

$$\rho_a(\alpha_q, \alpha_p) = \sqrt{(\gamma_q - \gamma_p)^2 + (\zeta_q - \zeta_p)^2 + (\eta_q - \eta_p)^2}. \quad (4)$$

Эффективным средством устранения повреждений отдельных элементов структурного описания есть логическая обработка значений используемой меры подобия, например, в виде предиката $L_V(\rho_V, \varepsilon_V, a, b) = \begin{cases} 1, & \rho_V(a, b) \leq \varepsilon_V \\ 0, & \rho_V(a, b) > \varepsilon_V \end{cases}$, где $a, b \in V$ – признаки или множества признаков из некоторого пространства V , ε_V – порог для значения метрики, который определяет точность при установлении эквивалентности компонентов описания a и b .

Обозначим элемент $\alpha_i \in A$, где $i = \overline{1, m}$, $\alpha_i = (\zeta_i, \eta_i, \gamma_i)$, m – мощность (число точек) множества A . Элементы α_i формируются путем перебора всевозможных базисов (троек ХП), координаты которых не лежат на одной прямой. Для каждого базиса вычисляются $(n-3)$ значений АИ.

На рис. 2 приведен график зависимости числа элементов ППС m от величины n . Анализ графика показывает, что количество элементов ППС резко возрастает с ростом количества точек в описании. Так, при числе ХП равном 20, количество ППС достигает 19 тысяч. Нелинейный характер роста числа ППС в зависимости от значения n в описании требует применения

специальных методов структурирования и фильтрации множества ХП в целях сокращения описания.

Применяя обсуждаемые в данной статье варианты анализа множества ППС с точки зрения группировки его элементов вокруг отдельных ХП или вокруг отдельных базисов, можно достичь улучшения характеристик распознавания за счет использования разработанных процедур фильтрации множества ППС.

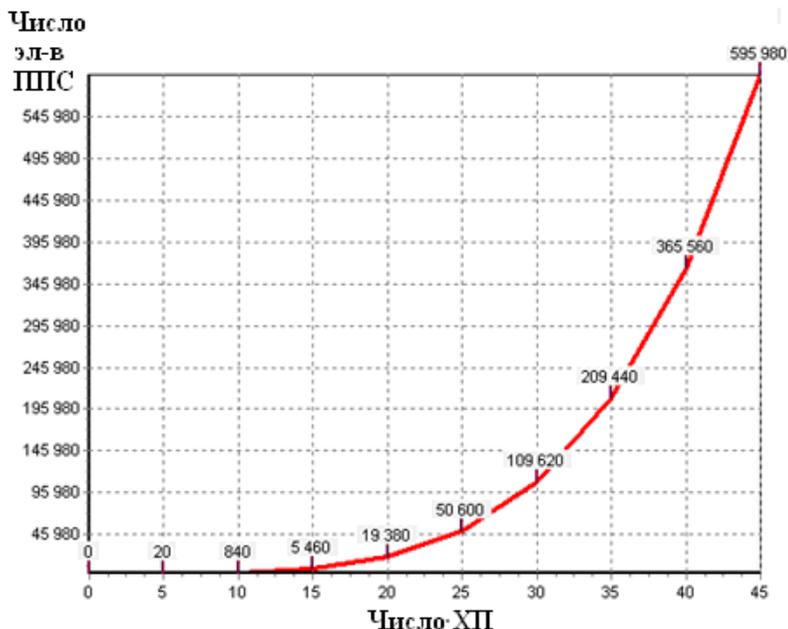


Рис. 2 – Зависимость числа ППС от количества ХП

Цель работы – исследование моделей для представления признаков пространственных структур, основанных на группировке элементов структурного описания для повышения эффективности решения задач классификации с точки зрения вычислительной сложности и достоверности.

Задачи исследования: формализация процедур сопоставления описаний в виде множеств признаков пространственных структур, оценка эффективности разработанных моделей на основе анализа аффинных базисов или характерных признаков в рамках одного базиса.

Структурные модели пространства признаков структур. Исходя из процедуры построения структурного описания, множество A можно интерпретировать как $A = \bigcup_k \bigcup_q \alpha_{kq}$, где α_{kq} – конкретный АИ, относящийся к

k -му базису, q – номер ППС для точек описания. Обсудим варианты представлений, анализа и групповой обработки множества A в целях эффективного сопоставления описаний при классификации.

1. Рассмотрим представление в виде объединения (покрытия): $A = \bigcup_k \{\alpha\}_k$, где $k = \overline{1, C_n^3}$, $\{\alpha\}_k \in A$ – подмножество значений ППС для k -го базиса. Подмножество $\{\alpha\}_k$ представляет собой набор векторов – троек значений ППС для всех ХП в системе координат одного базиса (рис. 3). Структурно множество $\{\alpha\}_k$ имеет следующий вид:

$$\{\alpha\}_k = \langle e_k, \{\alpha_q^k\}_{q=1}^{n-3} \rangle, \quad (5)$$

где e_k – базис, вокруг которого осуществляется группировка.

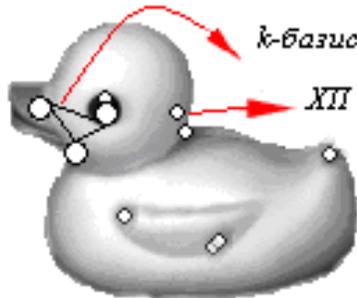


Рис. 3 – Иллюстрация формирования $\{\alpha\}_k$

Сопоставление двух разных описаний A и B сводится к вычислению меры подобия, отражающей число эквивалентных элементов множеств [2]. Традиционная схема реализуется как перебор элементов по правилу «каждый с каждым». Пусть A – эталон с числом ППС m_A , а B – анализируемый объект с числом ППС m_B , тогда их подобие можно определить через установление эквивалентности отдельных базисов, представленных подмножествами ППС. Например, будем считать два базиса эквивалентными, если для конкретного базиса из описания A найдено более половины эквивалентных элементов из текущего базиса описания B . Если такая эквивалентность установлена, проверка остальных базисов из состава B не

осуществляется. В результате мера подобия определяется как доля эквивалентных базисов к общему числу базисов эталона A .

Другой вариант сводится к установлению соответствия между отфильтрованными множествами базисов эталона и объекта [3]. В целях сокращения структурного описания введем процедуру $F_1: A \rightarrow A_1$ фильтрации на множестве базисов. В результате получим сжатое описание $A_1 = \bigcup_k \{\alpha\}_k$, где $k = \overline{1, s}$, s – число элементов в сжатом описании. Коэффициент сжатия равен C_n^3/s , что соответственно сокращает время классификации. Одним из примеров F_1 может быть процедурное априорно определенное формирование A_1 на принципах участия каждого эталонного ХП хотя бы в одном из базисов, используемых при сопоставлении в целях классификации.

2. Представим описание в формате покрытия $A = \bigcup_q \{\alpha\}_q$, $q = \overline{1, n}$, $\{\alpha\}_q$ – подмножество ППС для ХП с номером q . Подмножество $\{\alpha\}_q$ содержит совокупность тех α , которые представляют один ХП во всех базисах (рис. 4). Структура элемента описания в этом случае примет вид кортежа:

$$\{\alpha\}_q = \langle x_q, y_q, \{\alpha_k\}_{k=1}^{C_{n-1}^3} \rangle. \quad (6)$$

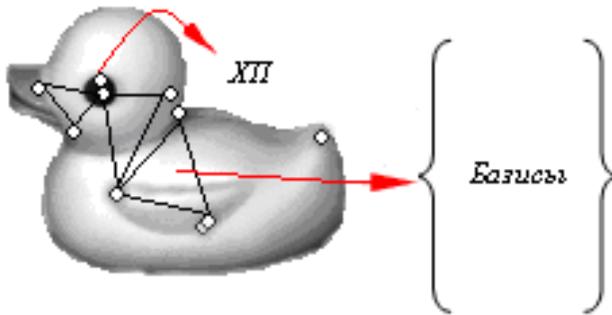


Рис. 4 – Иллюстрация формирования $\{\alpha_k\}_q$

Так как множество ППС содержит информацию о форме объекта, а множество ППС формируется на основе всевозможных комбинаций базисов для $n-1$ ХП, то всегда можно выбрать некоторое подмножество $n_v < n$ элементов, от которых в определяющей мере зависят геометрические

характеристики распознаваемого объекта. В этом случае подобие описаний можно установить на основе эквивалентности n_v элементов эталона. За счет сокращения количества точек в представлении эталона снижается и число базисов, которые формируются на основе этих точек, что в результате снижает вычислительную сложность распознавания.

Учитывая в представлении (6) дублирование структурных данных для разных ХП, целесообразно осуществить фильтрацию множества ХП, что одновременно с сокращением данных часто обеспечивает достаточно высокий уровень достоверности. Конкретным примером F_2 является проверка точек описания на коллинеарность, что фильтрует и в значительной мере сокращает объем описания [5].

Обсуждаемые методы сжатия описаний путем фильтрации зависят от выбранной базы эталонов, а также от ряда факторов: метода для вычисления ХП, порога на коллинеарность, требуемой достоверности распознавания, процедур выбора базисов и ХП в сжатом описании. Заметим, что введенные процедуры построения эффективных описаний обладают важным достоинством в плане управляемости необходимым объемом анализируемых при распознавании данных для достижения нужного уровня достоверности. Понятно, что факторы достоверности и вычислительных затрат являются взаимно исключаящими, что дает возможность выбора решения на основе оптимального соотношения их взаимосвязи или предпочтения одного из этих факторов в зависимости от решаемой проблемы. Оба варианта обработки реализуют схему рис. 1, относятся к покрытиям множества, близки друг другу в плане использования одного и того же исходного множества ППС, которое построено инвариантных характеристиках координат множества ХП.

Компьютерное моделирование. Экспериментальное исследование рассмотренных в работе моделей анализа структурных описаний в виде множеств ППС проведено для базы видео-данных Coil-20 [2,5]. Суть процедуры F_2 сводилась к проверке коллинеарности координат базиса, а сравнение осуществлялось путем установления эквивалентности подмножеств $\{\alpha\}_q$ объекта и эталона. Порог для проверки коллинеарности для множества дискретных координат выбран равным значению 3. Два подмножества $\{\alpha\}_q$ считались эквивалентными, если для объекта найдено более половины эквивалентных элементов, соответствующих эталонному подмножеству. Использовались параметры процедуры фильтрации, для которых вероятность правильной классификации оставалась равной 1.

Результаты экспериментов при данных условиях показали, что для отфильтрованного пространства ППС эталона среднее время классификации с применением модели 2 примерно в 5,5 раза меньше, чем для полного пространства. При этом объем описания, непосредственно влияющий на затраты памяти, сократился приблизительно в 1,2 раза.

Цель применения процедуры F_1 состоит в сокращении числа базисов на эталонном описании при сохранении достаточного уровня достоверности. Исходя из этих требований, суть F_1 заключалась в отборе не менее 30% базисов, в частности, выбирался каждый третий базис из полного множества. Эксперименты показали, что в результате применения F_1 время сопоставления сократилось в 2,5 раза.

Критерием, отражающим достоверность классификации для конечной базы данных, выбрана величина $\theta = h_1 / h_2$, где h_2 – максимальное число голосов, h_1 – ближайший к нему максимум. Значение θ служит показателем уверенности в принятии решения на основе максимума голосов, отданных за эталон. Глобальное решение в случае меньшего значения θ в большей степени значимо по отношению к локальному максимуму, который соответствует наиболее близкому из остальных (возможно, ложных) классов.

Достоверность классификации по критерию θ с применением структурных моделей пространства ППС значительно улучшилась в сравнении с подходом независимого голосования. Для традиционного подхода (с голосованием полного множества ППС) вычислена величина критерия $\theta = 0,78$, в то время как для подхода с использованием предложенных моделей на базе процедур F_1 и F_2 значение критерия составило $\theta = 0,01$, что существенно меньше. Достоверность повышается за счет процедуры группового сопоставления элементов подмножеств объекта и эталона, а быстроедействие – за счет процедур фильтрации и принятия решения по части элементов эталона. Полученные в эксперименте данные говорят о существенных преимуществах обработки, предложенной в данном исследовании.

Выводы. Представление описаний видео-объектов в виде множеств признаков пространственных структур предназначено для учета в процессе классификации геометрических свойств и соотношений между составными частями описания.

Применение рассмотренных моделей построения структурного пространства признаков видео-объектов ориентировано в первую очередь на сокращение избыточности, что обеспечивает снижение объема вычислительных затрат. В то же время и достоверность на сокращенном описании в пространстве ППС оказывается лучше. Управление объемом структурного описания дает возможность применить на практике другие интеллектуальные подходы к анализу данных, например, связанные с визуализацией информации, с применением кластеризации и нейронных сетей для анализа и др.

Рассмотренные процедуры фильтрации F_1 , F_2 позволяют управлять размером признакового пространства в зависимости от требований

прикладных задач. Такая гибкость может быть одним из путей решения проблемы хранения описаний видеоданных при создании визуальных баз знаний.

В статье показано, что применение разработанных моделей представления пространства геометрических инвариантных признаков, основанных на перегруппировке структурных элементов, обеспечивает повышение показателей достоверности и быстродействия классификации. На примере построенного пространства признаков получено экспериментальное подтверждение эффективности предложенных методов.

Практически важным является получение предпочтительных характеристик результатов сопоставления по сравнению с известными методами, что говорит о целесообразности развития и применения предложенных моделей обработки в задачах компьютерного зрения.

Перспективы исследования состоят в развитии предложенных моделей представления и фильтрации в плане их адаптации к свойствам и особенностям конкретных баз видеoinформации.

Список литературы: 1. *Шануро Л.* Компьютерное зрение / *Л. Шануро, Дж. Стокман.*; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 2. *Gorokhovatsky V. A.* Image Likelihood Measures of the Basis of the Set of Conformities / *V. A. Gorokhovatsky, Ye. P. Putyatin* // *Telecommunications and Radio Engineering.* – 2009. – Vol. 68 (9). – P. 763–778. 3. *Gorokhovatskiy V. A.* Compression of Descriptions in the Structural Image Recognition / *V. A. Gorokhovatskiy* // *Telecommunications and Radio Engineering.* – 2011. Vol. 70, No 15. – P. 1363 – 1371 4. *Lowe D. G.* Distinctive Image Features from Scale-Invariant Keypoints / *D. G. Lowe* // *International Journal of Computer Vision.* – 2004. – Vol. 60 (2). – P. 91–110. 5. *Гороховатский В. А.* Повышение быстродействия метода аффинных инвариантов при сопоставлении структурных описаний визуальных объектов / *В. А. Гороховатский, Т. В. Полякова* // *Системы управління, навігації та зв'язку: зб. наук. пр. – К.: ЦНДІ навігації і управління.* – 2009. – Вип. 2 (10). – С. 85–89. 6. *Rahtu E.* Object classification with multi-scale autoconvolution / *E. Rahtu, J. Heikkila* // *In Proceedings of ICPR04.* – Cambridge, England. – 2004. – Vol. 3. – P. 37–40. 7. *Bay H.* Surf: Speeded up robust features / *H. Bay, T. Tuytelaars, L. Van Gool* // *European Conference on Computer Vision.* – 2006. – P. 404–417. 8. *Гороховатский В. А.* Применение процедур голосования в структурных методах распознавания визуальных объектов / *В. А. Гороховатский* // *Вестник НТУ ХПИ. Системный анализ, управление и информационные технологии.* – Х.: НТУ ХПИ. – 2006. – № 39. – С. 132–140. 9. *Гороховатский В. А.* Формализм мультимножеств в задачах структурного распознавания и поиска в базах видеоданных / *В. А. Гороховатский, Ю. А. Куликов* // *Искусственный интеллект.* – 2012. – №1. – С. 76–85.

Надійшла до редколегії 08.12.2012