

## АНАЛИЗ И СЖАТИЕ СТРУКТУРНЫХ ОПИСАНИЙ НА ОСНОВЕ ГЕОМЕТРИЧЕСКИХ ИНВАРИАНТНЫХ ПРИЗНАКОВ

Т.В. ПОЛЯКОВА

Рассматриваются модели построения пространств геометрических инвариантных признаков в целях повышения эффективности сопоставления структурных описаний видеообъектов в системах компьютерного зрения. Предложенный подход позволяет управлять объемом описания для уменьшения вычислительной сложности, обеспечивая достаточный уровень достоверности распознавания. Обсуждаются результаты экспериментов на основе построенных моделей.

*Ключевые слова:* распознавание изображений, геометрические признаки, аффинные инварианты, структурное описание, фильтрация, базис, вычислительные затраты, достоверность.

### ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Структурные методы анализа видеоданных являются основными при решении современных прикладных задач компьютерного зрения [1]. Особую ценность для задач распознавания видеообъектов имеет геометрическая информация, содержащаяся в координатах характерных признаков (ХП), которые являются основой построения структурных описаний и отражают локально-пространственные свойства объектов в отдельных точках [2–7]. На основе координат ХП можно построить признаки более сложного плана, к которым относятся аффинные инварианты (АИ). Значения АИ непосредственно отражают геометрическую форму объекта и обладают инвариантностью к аффинным преобразованиям на плоскости [3, 4]. С другой стороны, АИ позволяют сохранять устойчивость при действии пространственных помех. Схема преобразования данных с использованием АИ представлена на рис. 1.



Рис. 1. Преобразование данных в процессе структурной классификации

Обозначим через  $n$  число ХП в структурном описании. Пусть  $A$  – множество АИ, сформированных для изображения на основе  $n$  пар координат ХП,  $A \subset R^3$ . 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  $\alpha_i$  формируются путем перебора всевозможных троек ХП, координаты которых не лежат на одной прямой. Для каждого базиса (тройки) вычисляются  $(n-3)$  значения АИ. Суть  $\alpha_i = (\zeta_i, \eta_i, \gamma_i)$  – аффинные проекции в базисе.

Максимально возможное число базисов из трех неколлинеарных точек равно числу комбинаций  $C_n^3$ . В результате мощность  $m$  множества  $A$  достигает величины

$$m = nC_{n-1}^3 = (n-3)C_n^3. \quad (1)$$

На рис. 2 приведен график зависимости числа аффинных инвариантов  $m$  от величины  $n$ . Анализ графика показывает, что количество АИ резко возрастает с ростом количества точек в структурном описании. Так, при числе ХП равном 35 количество АИ достигает 6 тысяч. Это обуславливает необходимость проведения исследований по сокращению объема описания в виде множества АИ.

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

Цель работы – анализ и сравнительная оценка эффективности применения структурных моделей пространств геометрических признаков, связанных с группировкой элементов описания, с точки зрения вычислительной сложности и достоверности для сопоставления описаний визуальных объектов при решении прикладных задач распознавания в системах компьютерного зрения.

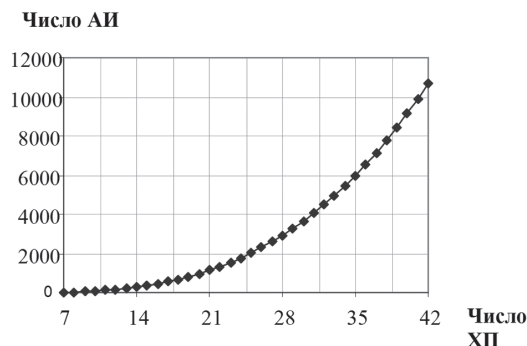


Рис. 2. Зависимость числа АИ от количества ХП

### 1. ПРЕДСТАВЛЕНИЕ СТРУКТУРНЫХ МОДЕЛЕЙ ПРОСТРАНСТВА АИ

Исходя из процедуры построения структурного описания, множество  $A$  можно интерпретировать как  $A = \bigcup_k \bigcup_q \alpha_{kq}$ , где  $\alpha_{kq}$  – конкретный

АИ, относящийся к  $k$ -му базису,  $q$  – номер АИ для точек описания.

Обсудим варианты представлений, анализа и групповой обработки множества  $A$  в целях эффективного сопоставления описаний.

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

$$\{\alpha\}_k = \langle e_k, \{\alpha_q^k\}_{q=1}^{n-3} \rangle, \quad (2)$$

где  $e_k$  – базис, вокруг которого осуществляется группировка.

Сопоставление описаний  $A$  и  $B$  сводится к вычислению меры подобия, отражающей число эквивалентных элементов множеств [7,8]. Традиционная схема реализуется как перебор элементов по правилу «каждый с каждым». Считая  $A$  эталоном с числом АИ  $m_A$ , а  $B$  – анализируемым объектом с числом АИ  $m_B$ , их подобие можно определить через установление эквивалентности отдельных базисов, представленных подмножествами АИ. 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  $A$ .

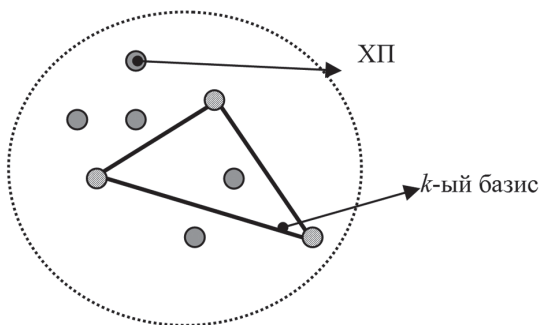


Рис. 3. Иллюстрация формирования  $\{\alpha\}_k$

Другой вариант сводится к установлению соответствия между сокращенными в результате фильтрации или просеивания множества базисов эталона и объекта. В целях сокращения структурного описания введем процедуру  $F_1: A \rightarrow A_1$  фильтрации на множестве базисов каждого эталона. В результате получим сжатое описание  $A_1 = \cup_k \{\alpha\}_k$ , где  $k = \overline{1, s}$ ,  $s$  – число элементов в сжатом описании. Коэффициент сжатия равен  $C_n^3 / s$ , что приводит к соответствующему сокращению времени распознавания.

Одним из примеров построения  $F_1$  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  $A = \cup_q \{\alpha\}_q$ ,  $q = \overline{1, n}$ ,  $\{\alpha\}_q$  – подмножество АИ для ХП с номером  $q$ . Подмножество  $\{\alpha\}_q$  содержит совокупность тех  $\alpha$ , которые представляют один ХП во всех базисах (рис. 4). Структура элемента описания в этом случае примет вид кортежа:

$$\{\alpha\}_q = \langle x_q, y_q, \{\alpha_k^q\}_{k=1}^{m/n} \rangle. \quad (3)$$

Так как множество АИ содержит информацию о форме объекта, а множество АИ 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 (3) 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 сокращает описание [5].

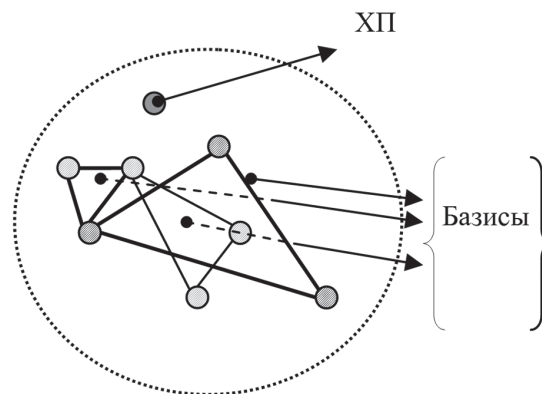


Рис. 4. Иллюстрация формирования  $\{\alpha\}_q$

Обсуждаемые методы сжатия описаний зависят от выбранной базы эталонов, а также от ряда факторов: метода выделения 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 исключающими, 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 АИ, которое построено на множестве имеющихся ХП.

## 2. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

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

Результаты экспериментов при данных условиях показали, что для сокращенного пространства структурных признаков эталона среднее время сопоставления и классификации изображения с применением модели представления вида 2 примерно в 5,5 раза меньше, чем для полного пространства. При этом объем описания (количество используемых ХП), влияющий на затраты памяти, сократился приблизительно в 1, 2 раза.

Цель применения процедуры  $F_1$  состоит в сокращении числа базисов на эталонном описании с сохранением 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. Экспериментальные исследования показали, что в результате применения  $F_1$  время сопоставления сократилось в 2,5 раза.

Критерием, характеризующим 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 критерия  $\theta$  служит показателем уверенности в принятии решения на основе максимума голосов, отданных за конкретный эталон. Глобальное решение в случае меньшего значения  $\theta$  в большей степени значимо по отношению к локальному, соответствующему

наиболее близкому из остальных (возможно, ложных) классов.

Достоверность классификации по критерию  $\theta$  с применением структурных моделей пространства геометрических признаков значительно улучшилась в сравнении с подходом независимого голосования АИ. Для традиционного подхода (с голосованием полного множества АИ) 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  $\theta = 0,01$ . Достоверность повышается за счет процедуры группового сопоставления элементов подмножеств объекта и эталона, а быстродействие – за счет процедур фильтрации и принятия решения по части элементов эталона. Полученные данные говорят о существенных преимуществах обработки, предложенной в данном исследовании.

## ВЫВОДЫ

Применение рассмотренных моделей построения структурного пространства признаков видеообъектов приводит, в первую очередь, к сокращению избыточности и обеспечивает уменьшение объема вычислительных затрат. В то же время и достоверность распознавания на сокращенном описании оказывается лучше. Управление объемом структурного описания дает возможность применить на практике другие известные интеллектуальные подходы к анализу данных, например, связанные с визуализацией информации, с применением кластеризации и нейронных сетей для анализа и др.

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 впервые показано, что применение разработанных моделей представления пространства признаков, основанных на перегруппировке структурных элементов, обеспечивает повышение показателей достоверности и быстродействия классификации. На примере геометрических инвариантных признаков изображений получено экспериментальное подтверждение эффективности предложенных методов для автоматизированного интеллектуального анализа видеоданных.

Практически важным является получение предпочтительных характеристик 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 конкретных баз видеоинформации.

## Литература

- [1] Шапиро Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
- [2] 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.
- [3] Форсайт Д. Компьютерное зрение. Современный подход: Пер. с англ. // Д. Форсайт, Ж. Понс // – М.: Издательский дом «Вильямс», 2004. – С. 928.
- [4] Rahtu E. Object classification with multi-scale autoconvolution / E. Rahtu, J. Heikkila // In Proceedings of ICPR04. - Cambridge, England 2004. – vol. 3. – P. 37–40.
- [5] Гороховатский В.А. Повышение быстродействия метода аффинных инвариантов при сопоставлении структурных описаний визуальных объектов / В.А. Гороховатский, Т.В. Полякова // Системи управління, навігації та зв'язку: зб. наук. пр. – К.: ЦНДІ навігації і управління, 2009. – Вип. 2(10). – С. 85–89.
- [6] Гороховатский В.А. Метод структурного распознавания изображений с использованием разнотипных признаков / В.А. Гороховатский, Т.В. Полякова // VII міжн. науково-практ. конф. «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2009)». Тези доповідей. – Дніпр-ськ: ДНУ, 25-27 листопада 2009. – С. 71–72.
- [7] Гороховатский В.А. Критерии и модели структурной классификации с применением принципа голосования / В.А. Гороховатский, Т.В. Полякова // Системи обробки інформації: зб. наук. пр. – Х.: Харк. ун-т повітряних сил ім. І. Кожедуба. – 2011. – Вип. 2 (92). – С. 12–16.
- [8] Гороховатский В.А. Оптимальные методы сопоставления структурных описаний видеообъектов / В.А. Гороховатский, Т.В. Полякова // Бионика интеллекта. – 2011. Вип. 3(77). – С. 85–89.
- [9] Рассел С. Искусственный интеллект: современный подход; 2-е изд. / С. Рассел, П. Норвиг; пер. с англ. – М.: Изд. дом «Вильямс», 2006. – 1408 с.

Поступила в редколлегию 06.06.2012



**Полякова Татьяна Викторовна**, ассистент кафедры информатики ХНАДУ, аспирантка ХНУРЭ. Область научных интересов: модели, методы, информационные технологии анализа и распознавания изображений.

УДК 004.932.2:004.93'1

**Аналіз та стиснення структурних описів на основі геометричних інваріантних ознак** / Т.В. Полякова // Прикладна радіоелектроніка: наук.-техн. журнал. – 2012. – Том 11. № 3. – С. 417–420.

Розглядаються моделі побудови просторів геометричних інваріантних ознак з метою підвищення ефективності порівняння структурних описів відеооб'єктів. Запропонований підхід дозволяє керувати обсягом опису для зменшення обчислювальної складності, забезпечуючи достатній рівень достовірності розпізнавання. Обговорюються результати експериментів на основі побудованих моделей.

*Ключові слова:* розпізнавання зображень, геометричні ознаки, афінні інваріанти, структурний опис, фільтрація, базис, обчислювальні витрати, достовірність

Л. 4. Бібліогр.: 9 найм.

UDC 004.932.2:004.93'1

**Analysis and compression of structural descriptions based on geometric invariant features** // T.V. Polyakova // Applied Radio Electronics: Sci. Journ. – 2012. Vol. 11. № 3. – P. 417–420.

The paper considers models of constructing spaces of geometric invariant features to increase the efficiency of comparison of structural descriptions of video objects. The suggested approach allows to control description size in order to reduce computational complexity, providing a sufficient level of recognition reliability. Experimental results based on the constructed models are discussed.

*Keywords:* pattern recognition, geometric features, affine invariants, structural description, filtering, basis, computational cost, reliability.

Fig. 4. Ref.: 9 items.