

УДК 004.932.2:004.93'1



## СТРУКТУРНО-ИЕРАРХИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ПРИ ВЛИЯНИИ ПРОСТРАНСТВЕННЫХ ПОМЕХ

В.А. Гороховатский

ХНУРЭ, г. Харьков, Украина, gorohovatsky-v@rambler.ru

Рассмотрена проблема распознавания изображений объектов при действии помех пространственного типа. Приведены результаты разработки структурно-иерархических методов, которые обеспечивают высокую вероятность правильного распознавания по сравнению с традиционными подходами. Обсуждаются пути оптимизации и оценки помехозащищенности, приводятся результаты экспериментов по распознаванию изображений реальных объектов.

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, СТРУКТУРНО-ИЕРАРХИЧЕСКИЕ МЕТОДЫ, ПРОСТРАНСТВЕННЫЕ ПОМЕХИ, СТРУКТУРНОЕ ОПИСАНИЕ ВИДЕО-ОБЪЕКТОВ

### Введение

Существующие методы анализа объектов на изображениях реальных визуальных сцен не позволяют в достаточной степени эффективно осуществить распознавание в условиях искажения отдельных частей объектов, т.к. модель их построения принципиально не учитывает влияния внешних воздействий, приводящих к формированию частичных и ложных описаний [1]. На рис. 1 демонстрируется действие фона и локальных помех как разновидностей пространственных помех. Искажение уже 30% описания видео-объекта существенно снижает вероятность его идентификации до уровня 0,6–0,7 [2]. Необходима разработка более универсальных структурных подходов, в основу которых должен быть положен иерархический анализ описаний пространственных признаков объектов и возможность принятия решения по представлению, включающему ложные структуры [3].

Для успешной реализации обсуждаемого подхода необходимо решение ряда самостоятельных задач, от которых в конечном итоге зависит эффективность распознавания. Для получения структурного представления нужно разработать надежные методы выделения структурных элементов (СЭ) в виде характерных признаков (ХП, англ. Keypoints), которые отражают локально-пространственные свойства объектов в виде структурного описания. Вторая задача связана с достоверным установлением сходства СЭ объекта и эталонов. Третья состоит в выборе способа функционального интегрирования элементарных сходств для формирования иерархического подобия как основы глобального решения о классе объекта. Функциональное интегрирование состоит в агрегировании и логической обработке множества локальных сходств СЭ и построении на этой основе функционала структурного сходства описаний [4]. Обсуждаемые методы опираются на структурное стратифицированное представление, характеризуются многоэтапной обработкой с включением иерархического анализа и названы структурно-иерархическими методами

(СИМ).

Цель работы – развитие теоретических основ построения и экспериментальное исследование СИМ как средства для эффективного решения проблемы распознавания видео-объектов при влиянии геометрических преобразований и искажений, вызванных воздействием пространственных помех.

Основной научный результат состоит в теоретическом обосновании принципа построения СИМ для обеспечения высокой достоверности распознавания объектов в реальной обстановке внешней среды.

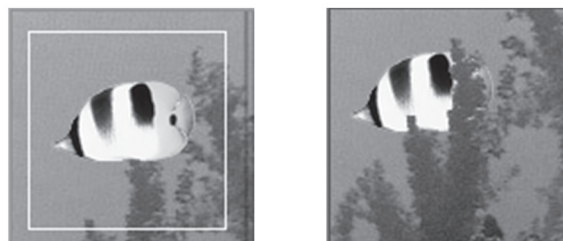


Рис. 1. Действие фона (а) и фона совместно с локальными помехами (б)

### 1. Принцип построения СИМ

Каждая точка  $(x, y)$  изображения  $B(x, y)$  – это объект  $B_0[T(x, y)]$ , полученный из эталона  $B_0(x, y)$  путем геометрических преобразований  $T$ , либо фон (локальная помеха), на которые дополнительно накладывается аддитивный шум. Функцию изображения  $B(x, y)$  в целях структурного анализа представим в виде конечного множества фрагментов  $\{b_k(x, y)\}$  или множества  $\Lambda = \{\lambda^k\}$  ХП. При знаку  $\lambda^k$  соответствует элемент  $c^k = (x^k, y^k)$  множества координат  $C = \{c^k\}$ , который определяет пространственные свойства описания. Визуальный объект представляется в виде кортежа  $Z = \langle \Lambda, C \rangle$ . Получение множества ХП как описания объекта направлено на решение проблемы распознавания и является ее важным этапом. От эффективности формирования ХП напрямую зависит результат распознавания.

Путем перехода к множеству ХП или множеству информативных фрагментов удается достичь существенного снижения объема информации, в то время как характеристики распознавания (достоверность, помехозащищенность) при этом практически не снижаются. В пространстве ХП данные рассматриваем как мультимножества [5], т.к. значения ХП бывают близкими между собой даже внутри одного описания. В большинстве применений ХП представляет собой числовой вектор  $\lambda^k = (\lambda_1, \dots, \lambda_n)^k$ ,  $\lambda_i \in R^1, i = \overline{1, n}$ ,  $\lambda^k \in R^n$ . В некоторых предметных областях типы ХП классифицируют путем присвоения меток [1, 5].

Далее нужно оценить эквивалентность объектов по их структурным описаниям  $Z_i, Z_m$ , что сводится к установлению соответствий ХП и получением на их основе значения глобального соответствия [3]. Наличие геометрических преобразований значительно усложняет анализ соответствий ХП. Проблема инвариантности актуальна как на локальном уровне формирования ХП, так и на этапе сопоставления описаний. Существующие подходы решают проблему распознавания путем «согласования» описаний в некотором признаковом пространстве, что соответствует значению критерия – функционала близости  $\vartheta(Z_i, Z_m)$ .

Наиболее распространена идея получения меры  $\vartheta(Z_i, Z_m)$  в виде оценки части общих элементов. Получение точного значения меры с учетом значений и структурных взаимоотношений ХП представляет собой сложную в вычислительном плане проблему. В качестве компромисса целесообразно рассмотреть меры, основанные на принципах голосования СЭ, в которое могут быть включены как ХП, так и их отношения, формирующие многоуровневые признаки. Голосование основано на оптимальных статистических решениях относительно ХП и позволяет отделить множество ХП объекта от ложных элементов. Подход представляется универсальным в силу учета многообразия возможных ситуаций, связанных с видом изображений, эталонов, сегментов, фоновых образований, помех.

Множество отношений  $R_Z \subseteq Z \times Z \times \dots \times Z$  для ХП опишем в виде предиката  $R_Z : Z \times Z \times \dots \times Z \rightarrow \{0, 1\}$ , где 1 означает истинность  $R_Z$  для набора аргументов. В силу простоты описания и анализа распространены бинарные отношения  $R_Z \subseteq Z \times Z$ . Имеем три типа отношений ХП:  $R_\Lambda$  – на множестве  $\Lambda$ ,  $R_C$  – на множестве  $C$ , а также комбинированные отношения  $R_Z = R_{\Lambda C}$  вида  $([\lambda^i, c^i], [\lambda^k, c^k])$ , заданные на множестве пар признаков – координаты. Отношения  $R_\Lambda$  реализуют групповые свойства ХП, а  $R_C$  и  $R_Z$  отражают пространственные связи ХП [3, 6]. Использование отношений развивает систему распознавания путем дополнения ее ресурсов многоуровневыми призна-

ками. Комбинации признаков в виде отношений и построение решений на использовании групп ХП заложены в природе структурного представления как способа грануляции данных и получения общих знаний о распознаваемом объекте [6, 7].

Концепция применения СИМ при распознавании основана на оптимизации близости между описаниями объекта и эталонов. Иерархическое сходство описаний вычисляется через сходство ХП. Модель применяемой в СИМ меры локально-пространственного подобия представима как функция  $\vartheta(Z^1, Z^2) = \Upsilon[\{\rho(\lambda_i^1, \lambda_k^2)\}]$  от симметрической матрицы  $\{\rho(\lambda_i^1, \lambda_k^2)\}$  значений метрики  $\rho$  для пар  $\lambda_i^1 \in \Lambda^1, \lambda_k^2 \in \Lambda^2$  [3]. Инвариантность построения  $\Lambda^1, \Lambda^2$  абстрагирует от пространственной информации и упрощает решение. На локальном уровне предпочтение отдано применению метрики  $\rho(\lambda_i^1, \lambda_k^2)$ , за счет чего оценка соответствия ХП определяется однозначно. Для устранения пространственных помех введем функцию  $\Psi_1[\rho(\lambda_i^1, \lambda_k^2)]$  логического анализа значений  $\rho(\cdot)$ . Одним из вариантов есть предикат вида

$$\Psi_1[\rho(\lambda_i^1, \lambda_k^2), \delta] = \begin{cases} 0, & \rho(\lambda_i^1, \lambda_k^2) \leq \delta, \\ 1, & \text{otherwise,} \end{cases} \quad (1)$$

параметрически определяемый порогом  $\delta$ . Управление изменением  $\delta$  в (1) в допустимом диапазоне  $[\delta_{\min}, \delta_{\max}]$  задает окрестность в пространстве ХП, в пределах которой  $\lambda_i^1, \lambda_k^2$  считаются эквивалентными. На практике порог  $\delta$  выбирают экспериментально для анализируемой базы данных. Предикат в виде отрицания  $\Psi_1[.]$  в теории интеллекта назван предикатом «узнавания».

Вторым важным моментом есть способ  $\Psi_2$  обобщения (свертывания) в общую глобальную меру локальных подобий, прошедших обработку  $\Psi_1$ . Здесь могут быть применены корреляционные, ассоциативные, дисперсионные меры подобия, а также метрики. Одним из вариантов есть метрика Танимото

$$\Psi_2(\Lambda^1, \Lambda^2) = \rho(\Lambda^1, \Lambda^2) = \mu(\Lambda^1 \Delta \Lambda^2) / \mu(\Lambda^1 \cup \Lambda^2), \quad (2)$$

где  $\mu(\cdot)$  – мощность множества;  $\Delta, \cup$  – операции симметрической разности и объединения множеств; эквивалентность элементов понимается в смысле предиката (1). Конкретно СИМ в соответствии с (1)-(2) задается кортежем  $\langle \Xi, \Psi_1, \Psi_2 \rangle$ , включающим способ  $\Xi$  формирования ХП и процедуры  $\Psi_1, \Psi_2$  анализа множеств ХП на локальном и глобальном уровнях обработки.

Распознавание объекта  $Z$  на основе СИМ рассмотрим как отображение  $\mathfrak{R} : Z \rightarrow \{Z(j)\}$ , где  $\{Z(j)\}$  – множество эталонных описаний, и реализуем его путем решения задачи оптимизации

$$j^* = \arg \max_{j \in J} \vartheta(Z, Z(j)), \quad j^* \in J. \quad (3)$$

В процессе решения (3) вычисляется мощность  $\mu(Z_j)$  подмножества  $Z_j \subseteq Z(j)$  описания объекта  $Z$ , для которых установлена эквивалентность элементам  $Z(j)$  и определяется  $Z_j$  наибольшей мощности с номером  $j^*$ . Выражения  $\mu(Z_{j^*})/\mu(Z)$  или  $\mu(Z_{j^*})/\mu(Z(j))$ , отражающие значение доли отданных за класс  $Z(j^*)$  голосов, трактуем как оценку апостериорной вероятности отнесения  $Z$  к классу  $j$ . Вычисление  $\mu(Z_j)$  в теоретико-множественном представлении соответствует голосованию и сводится к построению на множестве  $\Lambda$  (или  $Z$ ) представления  $\Lambda = \cup \Lambda_j$  в виде разбиения ( $\Lambda_i \cap \Lambda_j = \emptyset$ ) или покрытия ( $\Lambda_i \cap \Lambda_j \neq \emptyset$ ) из элементов, которые получают метку  $i^j$  класса  $j$  путем отображения  $\Lambda \rightarrow J$  [3]. Распознавание с применением СИМ включает применение основных способов изучения структуры объектов в системном анализе, связанных как с «препарированием», так и с «проектированием» объекта на совокупность близких объектов. Схема применения СИМ при распознавании представлена на рис. 2.

Важным фактором при принятии решения (3) есть мощность  $\mu(Z_{j^*})$  результирующей коалиции голосов. При выборе  $\mu(Z_{j^*}) < 0,5$  требуется дополнительное обоснование [3], т.к. при этом предполагается возможность решения на основе относительно небольшого числа ХП, что не вписывается в рамки традиционных статистических решений, где принято  $\mu(Z_{j^*}) \geq 0,5$ . Заметим, что отождествление описаний по части одинаковых ХП или фрагментов в СИМ приводит к отождествлению целых классов изображений, где эти элементы совпадают, а другие в то же время могут различаться. Это требует контроля решений.

Этапы СИМ: 1) формирование описания в виде множества ХП или фрагментов; 2) сопоставление ХП для вычисления локальных характеристик подобия; 3) установление меры локально-пространственного подобия между описаниями объекта и эталона; 4) оптимизация на множестве описаний эталонов. Каждый из этапов имеет внутреннюю структуру.

## 2. Оптимизация и помехозащищенность

СИМ есть обобщение традиционных интегральных подходов на случай влияния пространственных помех, где уровень локальности при анализе задается параметрически. Принцип оптимальности выбора параметров СИМ представляется как оптимизация в соответствии с аналитически заданными статистическими критериями с учетом меры подобия ХП, разновидности  $\Xi, \Psi_1, \Psi_2$ , а также распределений помех [2].

Сложность и многообразие представления аналитического вида критерия на примере среднего риска определяется распределениями эталона, объекта, аддитивной и пространственной помех, а также функцией сходства ХП. Одно из аналитических выражений для случая равномерного распределения эталона, функции сходства в виде суммы модуля разности и фрагментов без общих точек имеет вид

$$R(\delta_1, \delta_2) = \sum_{k=\delta_2+1}^s C_s^k p_1^k q_1^{s-k} + \sum_{k=1}^{\delta_2} C_s^k p_0^k q_0^{s-k}, \quad (4)$$

где  $\delta_1, \delta_2$  – пороги (параметры оптимизации);  $p_0, p_1$  – вероятности отклонения значения сходства в пределах порога  $\delta_1$ , которые вычисляются путем интегрирования  $n$ -мерной плотности ( $n$  – размерность ХП). Из-за сложного нелинейного вида критерия (4) оптимальные величины  $\delta_1, \delta_2$  получены численно.

Еще более сложный аналитический вид имеют математические выражения для критерия среднего риска или вероятности ошибки с учетом модели пространственной помехи [2].

Анализ показывает, что при существенных отличиях характеристик помехи и эталона СИМ имеет ощутимые преимущества: численно минимум вероятности ошибки в эксперименте составил 0,16 против 0,35 для традиционного метода. В то же время при незначительных отличиях, задаваемых отклонением по яркости, преимущества практически нет: минимумы примерно одинаковы. Это объясняется более существенными интегральными свойствами традиционного метода, потому

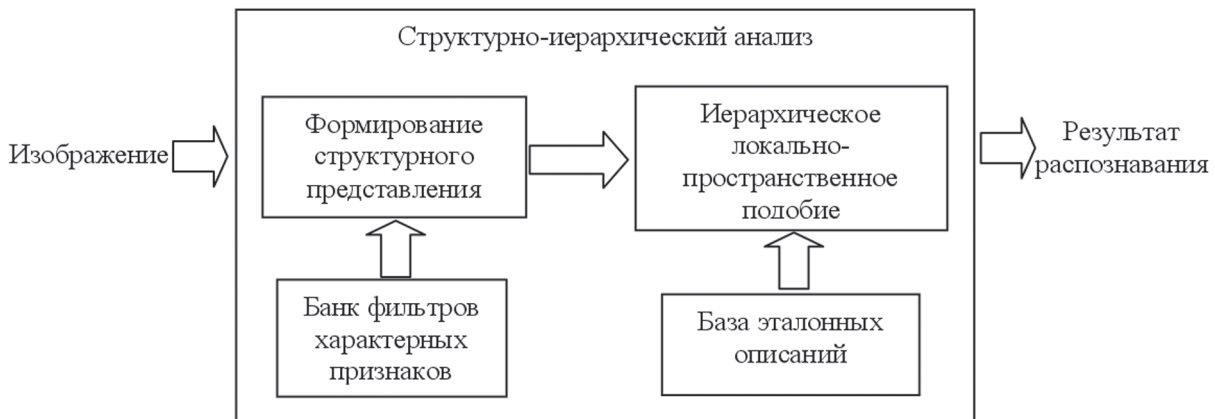


Рис. 2. Применение СИМ при распознавании

что при незначительных отличиях действие пространственных помех можно считать эквивалентным аддитивным флуктуациям. Этот факт подчеркивает тесную связь более универсального СИМ и традиционного метода, который можно считать его частным случаем. Преимущества применения СИМ усиливаются с увеличением отличий между значениями помехи и распознаваемого объекта.

### 3. Экспериментальные результаты

Проведено компьютерное моделирование с целью сравнительной оценки предложенных методов распознавания (СИМ) для тематической базы данных из 30 полутоновых изображений аквариумных рыб [2]. ХП получены методом SIFT, для сопоставления дескрипторов ХП использовано евклидово расстояние, а порог  $\delta$  выбран как 1% от максимума метрики. По результатам экспериментов наилучшие показатели отличия в пространстве признаков и адекватной реакции на локальные искажения имеют модификации СИМ на базе метрик Хемминга, Танимото, а также мер с применением подсчета бинарных соответствий ХП.

Допустимый уровень локальных искажений для СИМ составил 75-80% от объема описания при вероятности распознавания больше 0,98. Это значительно эффективнее традиционных методов (до 30%). Время обработки с использованием СИМ меньше, чем для известных методов. Анализ бинарных отношений в мерах, основанных на многозначных соответствиях, обеспечивает лучшую чувствительность к помехам исключения и большую достоверность решений, определяемую относительным локальным оптимумом значений мер.

### Выводы

Применение структурно-иерархических методов обеспечивает высокую вероятность распознавания и дает возможность осуществить оценку параметров объектов в условиях действия искажений пространственного типа. Проведенные эксперименты для систематизированных баз видеоданных в целях анализа вероятности распознавания, достоверности, помехозащищенности и быстродействия дают возможность получить обобщенные оценки для произвольных классов изображений. Решение практических задач идентификации протяженных объектов, космического мониторинга внешней среды, при контроле качества оборудования подтверждает прикладную ценность разработанных методов. Полученные результаты в плане эффективности распознавания открывают перспективу научно обоснованного решения ряда новых прикладных задач, где анализ осуществляется в присутствии ложных объектов.

Перспективы дальнейших исследований связаны с развитием адаптационных подходов в целях оптимального выбора параметров СИМ в зависимости от уровня пространственных помех.

**Список литературы:** 1. Шапиро, Л. Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман.; [пер. с англ. А.А. Богуславского, под ред. С.М. Соколова]. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 2. Путятин, Е.П. Методы та алгоритми комп'ютерного зору [Текст]: навч. посіб. / Е.П. Путятин, В.О. Гороховатський, О.О. Матат. – Х.: ТОВ «Компанія СМІТ», 2006. – 236 с. 3. Гороховатский, В.А. Структурное распознавание изображений на основе моделей голосования признаков характерных точек [Текст] / В.А. Гороховатский, Е.П. Путятин // Реєстрація, зберігання і обробка даних. – 2008. – Т.10. – №4. – С.75-85. 4. Gorokhovatsky V.A. Image Likelihood Measures of the Basis of the Set of Conformities / V.A. Gorokhovatsky, Ye. P. Putyatın // Telecommunications and Radio Engineering. – 2009, 68 (9). – P. 763–778. 5. Шлезингер, М. Десять лекций по статистическому и структурному распознаванию [Текст] / М. Шлезингер, В. Главач. – К.: Наукова думка, 2004. – 535с. 6. Гороховатский, В.А. Применение отношений на множестве характерных признаков изображений при распознавании на основе голосования [Текст] / В.А. Гороховатский // Бионика интеллекта. – 2008. – №1(68). – С. 87–93. 7. Kinoshenko D. A Partition Metric for Clustering Features Analysis / D. Kinoshenko, V. Mashtalir, V. Shlyakhov // International Journal «Information Theories and Applications». – 2007. – V. 14. – №3. – P. 230–236.

Поступила в редакцию 3.06.2010.

УДК 004.932.2:004.93'1

**Структурно-ієрархічні методи розпізнавання зображень при впливі просторових завад** / В.О. Гороховатський // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 58–61.

У роботі запропоновано та досліджено структурно-ієрархічні методи розпізнавання зображень об'єктів при дії завад просторового типу. Наведено принципи оптимізації та оцінки завадостійкості розроблених методів, приводяться результати комп'ютерного моделювання.

Л. 2. Бібліогр.: 7 найм.

UDC 004.932.2:004.93'1

**Structural-hierarchical methods of image recognition under the conditions of the influence of spatial distortions** / V.O. Gorokhovatskyi // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 58–61.

The structurally-hierarchical methods of image objects recognition under the conditions of spatial distortions influence are suggested and investigated. Principles of optimization and noise-stability estimation of the suggested methods and the results of computer modeling are shown.

Fig. 2. Ref.: 7 items.