

УДК 004.93'1

В.А. ГОРОХОВАЦКИЙ, д-р техн. наук, проф., зав. каф.,
Харьковский институт банковского дела университета банковского
дела Национального банка Украины, Харьков,

Н.В. ВЛАСЕНКО, канд. техн. наук, преп., Харьковский
национальный экономический университет им. С. Кузнеця, Харьков,

Т.В. ПОЛЯКОВА, канд. техн. наук, асс., Харьковский
национальный автомобильно-дорожный университет, Харьков

ПОСТРОЕНИЕ МОДИФИКАЦИЙ И АНАЛИЗ СВОЙСТВ СТРУКТУРНЫХ ОПИСАНИЙ ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Обсуждаются принципы построения, трансформации и свойства модификаций для структурных описаний изображений в виде множества характерных признаков. Приведены результаты экспериментов по оцениванию достоверности и помехозащищенности предложенных методов, которые в плане достоверности и быстродействия лучше известных методов. Ил.: 2. Библиогр.: 8 назв.

Ключевые слова: характерный признак, структурное описание, трансформация, достоверность, помехозащищенность.

Постановк проблемы и анализ литературы. В современных системах компьютерного зрения получили практическое применение структурные методы распознавания визуальных объектов, описание которых представлено в виде множества характерных признаков (ХП) [1 – 4]. ХП – это числовой вектор $z \in R^n$ фиксированной размерности n , рассматриваемый совместно с его координатами (x, y) . Каждый вектор z содержит информацию о функции яркости изображения в локальной окрестности точки (x, y) и формируется специальными методами, например, SURF [2]. Конечное множество векторов $\{z\}$ образует описание изображения, его элементы обладают рядом важных для распознавания свойств, таких как инвариантность к геометрическим преобразованиям анализируемого объекта в поле зрения, а также равенством 1 модуля вектора z . Эти свойства позволяют рассматривать все множество векторов $\{z\}$ как подмножество векторного пространства $\{z\} \subset R^n$.

В результате описание Z анализируемого изображения приобретает вид конечного множества $Z = \{z_i, x_i, y_i\}_{i=1}^s$, где s – его мощность.

© В.А. Гороховацкий, Н.В. Власенко, Т.В. Полякова, 2014

Учитывая разнородный состав множества Z , его можно рассматривать как объединение двух множеств $Z = \{z\} \cup \{(x, y)\}$, где $\{z\}$ отражает информацию о функции яркости объекта и в общем случае есть мультимножество, множество координат $\{(x, y)\}$ содержит пространственную информацию о взаимосвязи элементов $\{z\}$. Процедура распознавания объекта сводится к сопоставлению полученного описания Z с некоторым эталонным описанием $Z(j)$ из класса j на основе вычисления некоторой меры подобия описаний $\mathfrak{A}[Z, Z(j)]$ и оптимизации значения этой меры на конечном множестве классов $j = 1, \dots, J$ [4]. В самом общем виде классификация есть отображение $\mathfrak{R}: Z \rightarrow \{Z(j)\}$, где $\{Z(j)\}_{j=1}^J$ – конечное множество эталонных описаний (база изображений). Отображение \mathfrak{R} , как правило, реализуется путем решения задачи оптимизации

$$c(Z) = \arg \operatorname{opt}_{j \in \{1, \dots, J\}} \mathfrak{A}[Z, Z(j)], \quad (1)$$

где $\mathfrak{A} \in R^1$ – мера подобия, $c(Z)$ – функция получения класса объекта.

Следуя теории метрических пространств [5] и соотношению $\{z\} \in R^n$, для элементов $z_1, z_2 \in \{z\}$ можно задать произвольную метрику $\rho(z_1, z_2)$, которая определяет расстояние для пары произвольных конечных вещественных векторов в пространстве R^n , в результате чего появляется возможность оценить в теоретическом аспекте эквивалентность, а в практическом плане – близость двух векторов на основе условия $\rho(z_1, z_2) \leq \delta_z$, где δ_z – некоторый порог (параметр) точности. Исходя из этого, на основе сравнения составляемых элементов можно рассмотреть также метрическое пространство для множеств Z^1, Z^2 путем введения $\mathfrak{A}[Z^1, Z^2]$ как метрики для конечных множеств. Однако более практичным для объектов-изображений в условиях действия помех есть применение мер подобия $\mathfrak{A}[Z^1, Z^2]$, которые подсчитывают нормированное число эталонных элементов в анализируемом описании (голосование), на основе чего по максимуму функции подобия выбирается наиболее подходящий класс изображения.

Цель исследования – анализ свойств и эффективности применения модификаций систем структурных признаков при инвариантном распознавании изображений в компьютерном зрении.

Задачи работы состоят в модернизации систем признаков на основе исходного структурного описания, сравнительной оценке их эффективности с традиционными подходами, экспериментальном исследовании предложенных методов для баз изображений.

Системы признаков на основе множества ХП. Исходя из сути информационного разделения элемента описания (z_i, x_i, y_i) на вектор z_i инвариантного описания локальной окрестности и координаты (x_i, y_i) центра этой окрестности, целесообразно рассмотреть следующие пути построения, обобщения и сопоставления признаков описаний:

- на основе множества векторов-дескрипторов $\{z_i\}$;
- на основе множества координат $\{x_i, y_i\}$;
- на основе гибридного представления с одновременным включением $\{z_i\}$ и $\{x_i, y_i\}$.

В работе [6] применено преобразование множества $\{x_i, y_i\}$ координат описания к множеству $\{\gamma_k, \varsigma_k, \eta_k\}$ пространственных структур (ПС), $\gamma_k, \varsigma_k, \eta_k \in R^1$, отражающих геометрическое расположение точек $\{x_i, y_i\}$ между собой внутри описания. Значения $\gamma_k, \varsigma_k, \eta_k$ вычисляются несложным путем как решение системы линейных уравнений с учетом координат точки и базиса из трех неколлинеарных точек. Эти признаки получили название "аффинные инварианты" [1]. Их вычисление можно представить некоторой функцией F преобразования множеств $\{\alpha\} = \{\gamma_k, \varsigma_k, \eta_k\} = F\{x_i, y_i\}$. ПС $\alpha = (\gamma_k, \varsigma_k, \eta_k)$ геометрически закрепляет положение точки описания в пространстве других точек, превращая описание во множество "гроздей" из четверок ХП (один признак представлен в базисе трех других, рис. 1).

Применение аппарата ПС способствует повышению достоверности распознавания за счет включения в процесс распознавания информации о пространственных связях элементов описания. В основном это достигается за счет снижения вероятности "ложной тревоги", т.к. отдельные элементы структурных описаний разных изображений и даже группы элементов могут быть близкими между собой [6]. В то же время при применении ПС признаков возникает проблема сокращения объема полученного синтезированного пространства, т.к. мощность множества $\{\alpha\} = \{\gamma_k, \varsigma_k, \eta_k\}$ оказывается значительно большей, чем исходного множества $\{x_i, y_i\}$. Максимальное число признаков ПС непосредственно может быть оценено через исходное число ХП в описании через число

сочетаний. Другим недостатком рассмотренных методов можно считать некоторое снижение помехоустойчивости в случае действия пространственных помех, искажающих отдельные ХП описания. Из-за включения помеховых ХП в интегрированное описание структуры произойдет искажение и вновь созданного структурного элемента.

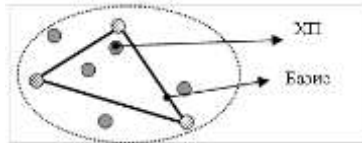


Рис. 1. Иллюстрация схемы формирования $\{\alpha\}_k$

Однако следует отметить, что рассмотренные недостатки нельзя считать критическими, т.к. объемом нового описания можно управлять параметрически (строить столько структур из возможного объема, сколько нужно). Кроме того, процедуру построения структур также можно адаптировать к прикладной задаче. Снижение же помехозащищенности к пространственным помехам, как показывают эксперименты [7], компенсируется увеличением устойчивости к аддитивному шуму за счет большей агрегации данных. Поэтому с применением ПС появляется возможность адаптивного выбора нужного числа признаков для обеспечения высокого уровня эффективности распознавания в условиях допустимого уровня помех разных типов.

Комбинирование путей 1 и 2 при построении признаков описаний приводит к гибридным представлениям в обработке [6, 7], причем пространственные признаки здесь могут выбираться как из исходного множества $\{(z_i, x_i, y_i)\}$, так и из преобразованного $\{\gamma_k, \zeta_k, \eta_k\}$ в ПС. Основная идея агрегирования состоит в вычислении значения некоторой комбинированной меры подобия между двумя "партнерскими" описаниями фактически из разных пространств. Исследованы следующие способы комбинирования:

- логическая обработка или анализ значений мер подобия в двух пространствах (геометрия и дескрипторы ХП), например, с использованием двух различных порогов для значений подобия, что устанавливает наличие логической связи между подобиями;

- последовательная модель сопоставления, когда один из типов признаков используется как основной при принятии решения, а второй играет роль контроля или подтверждения; например, находится минимум значения метрики по одному типу признаков, а по второму проверяется

условие удовлетворения порогу; из-за простоты построения эта модель получила наибольшее применение на практике;

– ассоциативная модель сопоставления, в которой фигурирует совместное использование пространственно-атрибутивных признаков в единой мере, например, в виде расстояния между элементами s_{iA}, s_{qB} описаний A, B

$$\rho(s_{iA}, s_{qB}) = \frac{1}{2}[\rho_{\alpha}(\alpha_{iA}, \alpha_{qB}) + \rho_d(z_{iA}, z_{qB})], \quad (2)$$

которое есть метрикой как линейная комбинация метрик; применение ассоциативной модели предполагает, что для всех геометрических признаков, построенных для одного ХП, используется одно и то же значение ее дескриптора; значения метрик ρ_{α}, ρ_d в (2) обычно нормируются к отрезку $[0, 1]$ для исключения преобладания одной из них.

Следует отметить, что применение сложных комбинированных мер в разнотипных пространствах несколько усложняет процедуру вычисления подобия в целях классификации, однако, как показывают эксперименты, ее применение оправдано в условиях помех [6, 7], т.к. помехозащищенность комбинированных подходов оказывается выше, чем при применении одного из пространств. Каждая из рассмотренных моделей описаний из некоторого пространства признаков допускает трансформации в виде сжатия, фильтрации элементов, перехода в другие более эффективные пространства и т.д. путем преобразования описания вида [8]

$$T: Z \rightarrow Z_T, \quad (3)$$

где T – отображение трансформации; Z_T – трансформированное описание.

Отображение T строится таким образом, чтобы описание Z_T обладало новыми, более совершенными свойствами по сравнению с Z в плане качества распознавания. Обычно качество характеризуется такими показателями, как вероятность правильного распознавания, достоверность, помехозащищенность, быстроедействие [4, 6]. Наши исследования показали, что в отдельных случаях трансформированное описание Z_T может быть получено в виде числового вектора v с целыми компонентами, например, в результате ортогонального разложения первоначального описания в некотором базисе [8].

Рассмотрим преобразование, связанное с разложением по семейству прямоугольных базисных функций, в качестве которых применим дискретные функции Уолша (ФУ). Преобразуем описание путем умножения матриц

$$U = Z * A, \quad (4)$$

где A – матрица Адамара размера $n \times n$, включающая n векторов ФУ w_1, \dots, w_n . Осуществим классификацию элементов Z по максимуму квадрата их значений из матрицы U^2 , в результате получим распределение s элементов описания Z по m априорно заданным классам, связанным с конкретными номерами ФУ. Сформируем сжатое описание u^* в агрегированном виде как вектор целых чисел

$$u^* = (u_1, \dots, u_m), \quad \sum_{j=1}^m u_j = s, \quad (5)$$

где элементы u_j – число векторов ХП, отнесенных к классу j [8]. Учитывая, что эталоны классов могут содержать в своем составе разное число ХП, на практике часто необходимо осуществить нормировку описания (5): $u_j = u_j / s$, что гарантирует выполнение соотношения

$$\sum_{j=1}^m u_j = 1.$$

Последовательность трансформации исходного описания Z в процессе сжатия можно формализовать как $Z \Rightarrow U \Rightarrow u^*$ или в функциональном виде

$$u^* = F_2[F_1[Z]], \quad (6)$$

где F_1 – преобразование к пространству функций Уолша, F_2 – классификация элементов в этом пространстве.

Предложенный метод трансформирует исходное описание в виде матрицы размера $s \times n$ к вектору из m числовых значений, причем выбор $m \ll n$ существенно снижает вычислительные затраты особенно при больших значениях числа s , которое в практических задачах иногда достигает 200 – 1000. Важным преимуществом при этом есть то, что преобразование множества эталонных ХП к сжатому виду выполняется

один раз на предварительном этапе и в целом не влияет на общее время классификации.

Для каждой из рассмотренных моделей описания в процессе классификации необходимо применение соответствующих мер подобия описаний на предмет установления их близости или эквивалентности. Для описаний в виде множества это будет мера сопоставления множеств, для описаний в виде вектора – мера для сопоставления векторов. Для векторов обычно используется подходящая метрика (в частности, для детекторов SIFT и SURF применяется евклидово расстояние), а для множеств – метрика или мера подобия в виде числа голосов эквивалентных элементов [4]. Примерами мер для множеств могут быть метрики Танимото, Хаусдорфа, а также большое разнообразие мер на основе подсчета схожих и несхожих элементов. Для объемных описаний в виде множества ХП часто целесообразно применение преобразований сжатия [4, 6]. Схема построения модификаций структурных описаний, включая разнообразие мер для их сопоставления, показана на рис. 2.

Результаты экспериментальных исследований. Как показали результаты компьютерного моделирования, предложенные модификации признаков пространств обеспечивают преимущественные показатели при классификации [7]. Эффективность проведения классификации для конечного числа классов в конкретном пространстве признаков (контрольная выборка) может быть оценена на основе значения критерия достоверности

$$\theta = \vartheta_m / \vartheta_0, \quad (7)$$

где ϑ_0 – значение максимального элемента матрицы подобия для всех эталонов, которое соответствует априорно заданному эталону, ϑ_m – максимум значения подобия с другими эталонами (недиагональные элементы матрицы подобия). Величина (7) непосредственно отображает степень различимости объектов в рассматриваемой базе изображений с применением выбранной системы признаков и связана с оценкой вероятности правильного распознавания. Чем меньше значение θ , тем увереннее работает метод распознавания с конкретными данными. Как показало компьютерное моделирование для фиксированной базы изображений, значение критерия θ при применении традиционного метода голосования (независимые ХП) достигает значения $\theta = 0,8$, в то время как для предложенного метода на основе ПС величина $\theta = 0,02$. Такое низкое значение критерия θ для разработанного метода

подтверждает его высокий показатель правильной классификации в эксперименте.



Рис. 2. Схема построения модификаций структурных описаний

В другом эксперименте исследовалась устойчивость методов с использованием ПС относительно действия пространственных помех, модель формирования которых состояла в преобразовании координат ХП в пределах окрестности ± 3 с некоторой вероятностью p . Эксперименты показали, что вероятность правильной классификации, которая подсчитывалась на основе отношения числа правильных ответов к общему числу экспериментов, для модифицированного метода с использованием ПС снижается не так быстро, как для традиционного. Так при $p = 0,35$ она составила значения 0,9 для разработанного метода и 0,8 для традиционного. Этот факт говорит о достаточно высоких показателях функционирования предложенных подходов на основе применения ПС.

Проведен сравнительный эксперимент по классификации для базы домашних животных [8], в котором признаковое описание представлено в пространстве 8 ФУ. Для сопоставления проведены также эксперименты

и для метода голосования, порог эквивалентности двух ХП выбран равным 0,16 (1% от максимума значения евклидовой метрики). Предложенный метод построения информативного описания на основе ортогонального пространства ФУ (4) – (6) имеет несомненное преимущество в плане помехозащищенности перед традиционным голосованием. До уровня аддитивной помехи, характеризуемой значением дисперсии $\sigma^2 = 0,02$, эффективность методов примерно одинакова (вероятность правильного распознавания выше 0,94), далее с увеличением σ^2 кривая вероятности для разработанного метода выше, и при $\sigma^2 = 0,03$ вероятности составляют 0,8 и 0,6 соответственно. Высокая устойчивость к шуму обусловлена прежде всего интегральными свойствами ФУ. В то же время быстродействие разработанного метода существенно выше: сравнительное время классификации одного изображения составило 0,03 с и 0,95 с соответственно, т.е. более чем в 30 раз лучше (компьютер Intel Core2 Duo 2,67 ГГц).

Выводы. Модифицированные методы построения и сопоставления структурных описаний визуальных объектов за счет учета пространственных геометрических соотношений и преобразования к векторному описанию обеспечивают лучшую достоверность распознавания и существенное сокращение объема вычислений по сравнению с традиционным подходом на основе голосования. Применение комбинированного анализа на основе объединения разных признаков пространств еще в большей степени улучшает характеристики распознавания в базах изображений.

Предложенные модифицированные системы признаков целесообразно применять для прикладных задач, требующих тщательного анализа особенностей изображений, когда традиционные методы не дают результата в плане достоверности. Другим направлением есть использование структурных методов в системах реального времени.

Впервые проведен анализ принципов построения и свойств систем структурных признаков, в основе которых лежит применение пространственных связей между ХП и агрегирование описания в векторное представление. На примерах моделирования процесса классификации с применением разработанных систем признаков для реальных изображений получено экспериментальное подтверждение эффективности предложенных методов. Практически важным есть получение предпочтительных характеристик распознавания в плане достоверности и быстродействия по сравнению с известными методами,

что говорит о целесообразности развития и применения предложенных модификаций в задачах компьютерного зрения.

Перспективы исследования состоят в анализе применений разработанных методов для объемных баз изображений.

Список литературы: 1. *Шапиро Л.* Компьютерное зрение / *Л. Шапиро, Дж. Стокман.* – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 2. *By H.* SURF: Speeded Up Robust Features / *H. By, A. Ess, T. Tuytelaars, L. Gool* // Computer Vision and Image Understanding (CVIU). – 2008. – № 110 (3). – P. 346-359. 3. *Szeliski R.* Computer Vision: Algorithms and Applications / *R. Szeliski.* – Springer. – 2010. – 979 p. 4. *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. 5. *Колмогоров А.Н.* Элементы теории функций и функционального анализа / *А.Н. Колмогоров, С.В. Фомин.* – М.: Наука, 1976. – 544 с. 6. *Gorokhovatskiy V.A.* Geometrical Invariant Features Peculiar for the Methods of Structural Classification of Images / *V.A. Gorokhovatskiy, T.V. Poliakova* // Telecommunications and Radio Engineering. – 2012. – № 71 (17). – P. 1557-1564. 7. *Полякова Т.В.* Анализ и сжатие структурных описаний на основе геометрических инвариантных признаков / *Т.В. Полякова* // Прикладная радиоэлектроника. – 2012. – Т. 11. – № 3. – С. 417-420. 8. *Власенко Н.В.* Построение информативных компактных описаний и классификация объектов путем представления в ортогональном базисе / *Н.В. Власенко* // Бионика интеллекта. – 2013. – №1 (80). – С. 93-98.

Bibliography (transliterated): 1. *Shapiro L.* Komp'yuternoe zrenie / *L. Shapiro, Dzh. Stokman.* – М.: BINOM. Laboratorija znaniy, 2006. – 752 s. 2. *By H.* SURF: Speeded Up Robust Features / *H. By, A. Ess, T. Tuytelaars, L. Gool* // Computer Vision and Image Understanding (CVIU). – 2008. – № 110 (3). – P. 346-359. 3. *Szeliski R.* Computer Vision: Algorithms and Applications / *R. Szeliski.* – Springer. – 2010. – 979 p. 4. *Gorokhovatskiy V.A.* Strukturnoe raspoznavanie izobrazhenij na osnove modelej golosovaniya priznakov harakternyh toчек / *V.A. Gorokhovatskiy, E.P. Putjatin* // Reєstracija, zberigannja i obrobka danih. – 2008. – Т. 10. – № 4. – С. 75-85. 5. *Kolmogorov A.N.* Jelementy teorii funkcij i funkcional'nogo analiza / *A.N. Kolmogorov, S.V. Fomin.* – М.: Nauka, 1976. – 544 s. 6. *Gorokhovatskiy V.A.* Geometrical Invariant Features Peculiar for the Methods of Structural Classification of Images / *V.A. Gorokhovatskiy, T.V. Poliakova* // Telecommunications and Radio Engineering. – 2012. – № 71 (17). – P. 1557-1564. 7. *Poljakova T.V.* Analiz i szhatie strukturnyh opisanij na osnove geometricheskih invariantnyh priznakov / *T.V. Poljakova* // Prikladnaja radiojelektronika. – 2012. – Т. 11. – № 3. – С. 417-420. 8. *Vlasenko N.V.* Postroenie informativnyh kompaktnyh opisanij i klassifikacija ob'ektov putem predstavlenija v ortogonal'nom bazise / *N.V. Vlasenko* // Bionika intellekta. – 2013. – № 1 (80). – С. 93-98.

Поступила (received) 26.03.2014

Статью представил д-р техн. наук, проф., зав. каф. информационных технологий и мехатроники ХНАДУ Никонов О.Я.

Gorokhovatskiy Vladimir, Dr.Sci.Tech, Professor
Kharkiv Institute of banking of the University of banking of the National bank of Ukraine
Ave. Peremohy 55, Kharkiv, Ukraine, 61174
tel./phone: (057) 336-05-64, e-mail: gorokhovatsky-v@rambler.ru
ORCID ID: 0000-0002-7839-6223

Vlasenko Natalia
Kharkiv National University of Economics

Ave. Lenina 9, Kharkiv, Ukraine, 61166
tel./phone: +38 (057) 702-06-74, e-mail: gorohovatskaja@gmail.com
ORCID ID: 0000-0002-2560-9854

Poliakova Tatiana, asst.
Kharkiv National Automobile and Highway University
Str. Petrovskogo 25, Kharkiv, Ukraine, 61002
tel./phone: +38 057 7073774, e-mail: assiada@list.ru
ORCID ID: 0000-0002-2857-8897