

УДК 004.932.2:004.93'1

**В. А. Гороховатский, Т. В. Полякова, Е. П. Путятин**

Харьковский национальный университет радиоэлектроники  
Проспект Ленина, 14, 61166 Харьков, Украина

## **Модели комплексированных мер подобия структурных описаний изображений**

*Обсуждаются вопросы обработки информации в интеллектуальных системах компьютерного зрения. В целях повышения достоверности классификации изображений для сопоставления структурных описаний предложено построение мер подобия путем комплексирования разнотипных признаков. Приведены результаты экспериментов.*

**Ключевые слова:** компьютерное зрение, структурные методы классификации изображений, характерные признаки, мера подобия, комплексирование, отношения, помехозащищенность.

### **Введение**

Извлечение нужной информации представляет сложную проблему из-за большого объема излишних данных, которые содержат изображения наряду с информацией об интересующих визуальных объектах. Эффективными могут быть структурные технологии, основанные на анализе особенностей изображения в отдельных точках. Анализ свойств описания видеообъекта в виде конечного числа характерных признаков (ХП) позволяет решить практические задачи любой сложности, в частности, связанные с распознаванием объектов на сложном фоне [1–5]. Основой для формирования ХП служит информация, содержащаяся в «значимых» фрагментах. Распознавание на основе описания объектов в виде множества ХП имеет несомненные преимущества: упрощение процедур формирования признаков, сокращение признакового пространства и возможность управления его размерностью, применимость структурного анализа описания для обеспечения устойчивости к влиянию фоновых искажений и ложных объектов.

В пространстве ХП распознавание (классификация) объектов сводится к построению мер подобия конечных мультимножеств на основе соответствий элемент–элемент, элемент–множество, множество–множество. Эффективный способ реализации мер — голосование ХП [1]. Наряду с устойчивостью принятия решения при неполном или избыточном (из-за влияния помех) представлении описаний, голосование обладает такими важными достоинствами как высокая вероят-

ность правильного распознавания и простота построения, что делает его конкурентоспособным среди других подходов.

Одним из путей развития моделей голосования в плане повышения эффективности есть комплексирование разнотипных и разношкальных признаков (атрибутивных, статистических, пространственных, текстурных, логических и др.), полученных различными способами [6–9]. Многоаспектный анализ позволяет с большей точностью осуществить классификацию описаний за счет использования сочетаний признаков из разных пространств.

Один из способов комплексирования — совместное использование атрибутивных (например, значения ХП в соответствии с детектором SIFT) и пространственных (значения аффинных инвариантов) характеристик ХП [4]. При этом аффинные инварианты (АИ) используются при поиске соответствующих им АИ эталона независимо от атрибутов ХП, а при формировании голоса в целях улучшения достоверности дополнительно проверяется величина подобия соответствующих детекторов ХП.

Цель работы — построение и анализ моделей мер подобия для множеств ХП на основе сопоставления структурных описаний видеообъектов путем комплексного использования геометрической и атрибутивной информации.

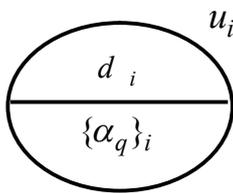
Задачи исследования — формализация мер подобия описаний с применением признаков из разных пространств, анализ вариантов комбинирования признаков, оценка эффективности предложенных моделей путем экспериментального исследования.

### Построение и сопоставление элементов описаний

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

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

где  $i = \overline{1, n}$ ;  $n$  — количество ХП в описании  $U$ ;  $\alpha_q = (\zeta_q, \eta_q)$  — значения АИ [3];  $m_i = 6C_{n-1}^3$  — количество величин  $\alpha_q$  для  $i$ -го ХП;  $d_i \in R^k$  —  $k$ -мерный числовой вектор атрибутивных признаков, инвариантный к геометрическим преобразованиям из некоторой группы, который отражает свойства изображения в локальной окрестности точки  $(x_i, y_i)$  и полученный, например, в соответствии с методом SIFT [1]. В общем случае  $\{d_i\}$ ,  $\{\alpha_q\}$  являются мультимножествами. Структура признака  $u_i$  показана на рисунке.



Структура характерного признака  $u_i$

Координаты  $c_i = (x_i, y_i)$  для  $u_i \in U$  могут быть представлены через базис  $\{e_0, e_1, e_2\}$  в виде пары АИ  $(\zeta, \eta)$  [8]:

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

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

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

Путем представления (3) после упорядочивания  $(\gamma, \zeta, \eta)$  значение  $m_i$  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 вычислительные затраты.

Рассмотрим расстояние на множестве АИ как евклидову метрику:

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

Расстояние  $\rho_d(d_i, d_j)$  между атрибутивными признаками определим аналогично (4) как евклидову метрику в пространстве векторов.

Эффективным средством устранения влияния ложных или изменившихся под влиянием помех элементов структурного описания есть логическая обработка значений 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  $V$ ;  $\rho_V$  — метрика в этом пространстве;  $\varepsilon_V$  — порог для значения метрики, который задает эквивалентность  $a$  и  $b$  [1].

Для определенности будем рассматривать задачу классификации видеообъектов на конечном множестве классов как вычисление и оптимизацию меры подобия их описаний [1, 3].

### Комплексированные меры подобия описаний

Рассмотрим варианты моделей для комплексированных мер  $\mathcal{G}(U_1, U_2)$  подобия описаний  $U_1, U_2$  с элементами  $u_i$  в виде структуры (1). Для конкретности сосредоточимся на описании мер с использованием принципа голосования элемент–множество, хотя предложенные способы обработки можно распространить и на другие типы голосования.

1. *Последовательная модель с подтверждением.* Вычислим вначале в пространстве геометрических признаков  $\{\alpha\}$  расстояние типа точка–множество

$$\rho(\alpha_{i1}, \{\alpha_p\}_2) = \min_p \rho_\alpha(\alpha_{i1}, \alpha_{p2}) \quad (5)$$

от ХП с номером  $i$  объекта  $U_1$  до множества  $\{\alpha_p\}_2$  ХП объекта  $U_2$ . Зафиксируем номер  $j$  элемента из  $U_2$ , для которого достигается минимум (5). Далее проверяем условие на значимость подобия (5) посредством предиката  $L_\alpha(\rho_\alpha, \varepsilon_\alpha, \alpha_{i1}, \alpha_{j2})$ , где  $\varepsilon_\alpha$  — порог. При выполнении  $L_\alpha(\cdot) = 1$  голос  $h_i$  элемента  $\alpha_{i1}$  засчитываем ( $h_i = 1$ ), если дополнительно выполняется также истинность предиката  $L_d(\rho_d, \varepsilon_d, d_{i1}, d_{j2})$ , где  $d_{i1}, d_{j2}$  — атрибуты точек  $u_{i1}$  и  $u_{j2}$ ;  $\varepsilon_d$  — порог. В противном случае принимаем  $h_i = 0$ .

Аналогично (5) можно вначале найти расстояние  $\rho(d_{i1}, \{d\}_2) = \min_q \rho_d(d_{i1}, d_{q2})$  в пространстве атрибутов от точки  $d_{i1}$  (компонента кортежа  $u_{i1} \in U_1$ ) до множества  $\{d\}_2$  атрибутов объекта  $U_2$ , а затем при условии истинности предиката  $L_d(\rho_d, \varepsilon_d, d_{i1}, d_{j2}) = 1$  определить голос  $h_i$  путем дополнительной проверки истинности  $L_\alpha(\rho_\alpha, \varepsilon_\alpha, \alpha_{i1}, \alpha_{j2})$ .

Если действовать аналогично предыдущему способу, то подтверждение соответствия может быть выполнено в нескольких вариантах, так как каждый из атрибутов  $d_{i1}, d_{j2}$  здесь рассматривается вместе со своим множеством АИ. Можно рассмотреть такие условия: расстояние между соответствующими множествами АИ меньше некоторого порога; расстояние (5) от одного из АИ до уже зафиксированного эталона меньше порога; расстояние от выбранного подмножества АИ до множества АИ зафиксированного эталона меньше порога и др. По объему вычислений такой способ последовательного анализа признаков может оказаться эффективнее схемы (5), т.к. объем данных для элемента из пространства  $\{d\}$  значительно меньше, чем для элемента из  $\{\alpha\}$ .

Оба способа в предложенном варианте обработки отличаются тем, что один из типов признаков (атрибуты или геометрия) получает преимущество перед другим, который, в свою очередь, играет роль контроля или подтверждения.

Результирующее значение меры вычисляется в виде некоторой функции  $\psi$  от суммы голосов  $\mathcal{G}(U_1, U_2) = \psi(\sum h_i)$ , связанной с нормировкой на их максимально возможное число.

2. *Ассоциативная модель.* Совместное использование в одной мере пространственно-атрибутивных признаков можно реализовать в виде ассоциативной меры для  $u_{i1} \in U_1, u_{j2} \in U_2$ :

$$\mathcal{G}(u_{i1}, u_{j2}) = \frac{1}{2}[\rho_\alpha(\alpha_{i1}, \alpha_{j2}) + \rho_d(d_{i1}, d_{j2})], \quad (6)$$

причем (6) является метрикой как линейная комбинация метрик [6, 7]. Отметим, что для всех АИ, относящихся к одному ХП, используется одно значение дескриптора ХП. Минимум (6) определяется на множестве эталонных описаний и затем проверяется на превышение порога  $\varepsilon_g$  для значения метрики. Величины метрик в (6) обычно нормируются к отрезку  $[0, 1]$  для исключения преобладающего

влияния одной из них. Как показали наши эксперименты, диапазоны изменения признаков  $\alpha$  и  $d$  достаточно близки. Ассоциативная мера (6) учитывает значения признаков из разных пространств равноценным образом, здесь возможно применение также взвешенных характеристик [6].

3. *Логическая модель.* Применяя бинарную обработку  $L_\nu(\cdot)$  в соответствующих пространствах признаков совместно с логической операцией  $\Theta$ , можно построить меры подобия в виде функции

$$\mathcal{G}(u_i, u_j) = \phi[\rho_\alpha, \rho_d, \Theta[L_\alpha(\cdot), L_d(\cdot)]] \quad (7)$$

где  $\Theta$  задает логическую взаимосвязь признаков из разных пространств;  $\rho_\alpha, \rho_d$  — метрики;  $\varepsilon_\alpha, \varepsilon_d$  — пороги в предикатах  $L_\alpha, L_d$  для соответствующих типов признаков. Если, например, в (7) выбрать в качестве  $\Theta$  конъюнкцию, то имеем меру на базе проверки истинности обоих предикатов. Если  $\Theta$  — дизъюнкция, то имеем меру на основе истинности одного из предикатов. Логическая модель (7) больше подходит для установления соответствия описаний в виде множеств. Один из вариантов: если минимум по  $\{d\}$  и минимум связанного с ним  $\{\alpha\}$  находятся внутри одного и того же эталона, то засчитывается голос за этот эталон. Более практичным есть применение конъюнкции, другие варианты приводят к последовательной модели.

Правило определения голоса  $h_j(u_i)$  элемента  $u_i$  при вычислении (7) на основе  $\Theta$  в виде конъюнкции можно получить в виде

$$h_j(u_i) = \begin{cases} 1, & [i_j(\alpha_i)] \neq \emptyset \& [i_j(d_i)] \neq \emptyset, \\ 0 & otherwise, \end{cases} \quad (8)$$

где  $i_j(d) = \{d \in \Lambda(j) \mid L_d\}$ ,  $i_j(\alpha) = \{\alpha \in \{\alpha\}_j \mid L_\alpha\}$  — множества соответствий атрибута  $d$  и АИ  $\alpha$  в компоненте  $j$ -го эталона; символ  $\&$  означает одновременное выполнение условий. Голос (8) засчитывается, если множества соответствий каждого из типов признаков в эталонном множестве не пусты.

В общем случае меры (5)–(8) применимы и для варианта множественной обработки соответствий ХП с порогами  $\varepsilon_\alpha, \varepsilon_d$ , когда одному ХП объекта может соответствовать несколько ХП из эталонного описания. Множественные соответствия имеют более высокую помехозащищенность, чем однозначные [3].

Если в (7) считать  $\phi$  способом комплексирования, то эта модель имеет общий характер и учитывает произвольный тип совместной обработки пространств признаков. Выражение (7) можно охарактеризовать как меру на пространственно-атрибутивных отношениях элементов  $\alpha, d$ , т.к. именно отношения элементов выступают аппаратом при построении таких мер.

Применение мер вида (7) в задаче классификации сводится к подсчету и оптимизации числа голосов  $h_l$ , отданных за эталон класса  $l$ . Результат определяется функцией  $\psi(\sum h_l)$  в виде нормированной на отрезке  $[0,1]$  величины путем нор-

мировки: на число  $n_l$  ХП эталона, на число  $n$  ХП объекта или на произведение  $n_l n$ . Результат  $h_l^*$ , соответствующий оптимуму значений меры на множестве классов  $l = \overline{1, M}$ , в целях обеспечения достоверности решения проходит верификацию вида  $h^*(l) \geq \varepsilon_g$ , где  $\varepsilon_g$  — порог для минимально значимой величины подобия, напрямую связанной с числом ХП, которое может представлять распознаваемый объект в решаемой прикладной задаче.

### Результаты экспериментов и выводы

Комплексированные методы сопоставления структурных описаний за счет многостороннего анализа свойств объектов обеспечивают более высокую достоверность распознавания по сравнению с традиционными подходами.

На результат распознавания непосредственно влияют значения порогов  $\varepsilon_\alpha$ ,  $\varepsilon_d$ ,  $\varepsilon_g$ . На практике их выбор зависит от базы видеоданных, а также от уровня аддитивных помех, определяющих допустимое отклонение значений признаков. Обычно величина отклонения задается в виде процента (1–10 %) от максимального разброса в значениях признаков или от меры их близости [1, 3]. Экспериментальные исследования по выбору порогов для конкретной базы видеоданных Coil-20 [8] показывают, что подавляющее большинство значений признаков  $\alpha_q$ , сформированных путем анализа координат ХП, принадлежат диапазону  $[-15; 15]$ , и только отдельные значения не входят в этот интервал. Из этих соображений выберем порог 0,3 как 1 % от максимального разброса. Порог  $\varepsilon_d$  для эквивалентности атрибутов ХП (вектор размерности 128 с нормированными значениями на отрезке  $[0, 1]$ ) определялся как 1 % от разброса значений евклидовой метрики и выбран  $\varepsilon_d = 0,11$ . Значение порога для ассоциативной меры  $\varepsilon_g = 0,2$  вычислено как среднее арифметическое  $\varepsilon_\alpha$ ,  $\varepsilon_d$ .

Основная цель экспериментов — сравнительная оценка эффективности разных видов комплексирования и традиционных подходов в условиях помех.

Первая задача состояла в оценке работоспособности метода SIFT при действии на каждую компоненту детектора аддитивной помехи с нормальным распределением (матожидание — ноль, дисперсия —  $\sigma^2$ ). Табл. 1 содержит вероятности  $P$  правильной классификации объектов для базы Coil-20 в зависимости от уровня помехи, задаваемого значением  $\sigma$  среднеквадратического отклонения (СКО). Как видим из табл. 1, до уровня  $\sigma = 0,012$  метод SIFT обеспечивает вероятность  $P = 1$ , а при  $\sigma = 0,0128$  значение  $P \approx 0,9$ .

Таблица 1. Вероятность распознавания для аддитивных помех

$\sigma$	0,012	0,0125	0,0127	0,0128	0,0129	0,013	0,0135
$P$	1	0,98	0,92	0,88	0,82	0,67	0,59

Моделирование воздействия пространственных помех осуществлялось путем изменения координат отдельных ХП объекта, которые с вероятностью 0,5 преобразовывались добавлением значений  $\pm 1$ ,  $\pm 2$ ,  $\pm 3$ . Величина окрестности измене-

ния координат отражает уровень помехи. Табл. 2 содержит вероятности правильного распознавания для разных методов: традиционного голосования на основе АИ, голосования на основе АИ с пороговой проверкой значения атрибута (комплексирование (5)), применением ассоциативной метрики (6). При этом для первого метода применялась только модель пространственных помех, т.к. аддитивная помеха не изменяет координат ХП, для остальных двух — одновременное действие аддитивной и пространственной помех. Для экспериментов выбран уровень аддитивной помехи  $\sigma = 0,0127$ , чтобы оценить устойчивость комплексирования в наиболее сложной ситуации влияния помех разных типов.

Таблица 2. Вероятности распознавания для разных методов

Окрестность	0	$\pm 1$	$\pm 2$	$\pm 3$
Голосование $\alpha_q$	1	0,85	0,76	0,67
Комплексирование (5)	0,92	0,91	0,91	0,89
Ассоциативная мера (6)	0,99	0,99	0,99	0,95

Как видим из табл. 2, применение комплексирования имеет существенные преимущества по сравнению с традиционными методами, основанными на раздельном голосовании атрибутов и геометрических признаков АИ. Вероятность классификации для традиционного метода на основе  $\alpha_q$  резко снижается до 0,67 с увеличением окрестности, в то время как комплексированные методы сохраняют высокую вероятность 0,95 при значительных искажениях. Установлено, что ассоциативная мера более устойчива к пространственным искажениям по сравнению с комплексированием (5). Время обработки для рассмотренных в табл. 2 методов практически совпадает.

Впервые показано, что объединение пространств представления структурных признаков в одной мере при сопоставлении описаний улучшает гибкость, универсальность анализа и повышает достоверность классификации за счет более точного и многостороннего анализа свойств описаний. Изучены варианты комплексированных мер на основе пространств геометрического и атрибутивного плана, свойственных системам компьютерного зрения.

Практически важным является получение предпочтительных характеристик распознавания в условиях аддитивных и пространственных помех с применением комплексированных мер по сравнению с известными методами, что говорит о целесообразности их развития и применения в задачах компьютерного зрения.

Перспективы исследования состоят в разработке подходов к построению мер подобия, где исключается дублирование пространственной информации для элементов эталонного описания в базе данных.

1. Гороховатский В.А. Структурное распознавание изображений на основе моделей голосования признаков характерных точек / В.А. Гороховатский, Е.П. Путьгин // Реєстрація, зберігання і оброб. даних. — 2008. — Т. 10, № 4. — С. 75–85.

2. *Rahtu E.* Object Classification with Multi-Scale Autoconvolution / E. Rahtu, J. Heikkila // In Proceedings of ICPR04. — Cambridge (England), 2004. — Vol. 3. — P. 37–40.
3. *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.
4. *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-к: ДНУ, 25–27 листопада 2009. — С. 71–72.
5. *Баклицкий В.К.* Методы фильтрации сигналов в корреляционно-экстремальных системах навигации / В.К. Баклицкий, А.М. Бочкарев, М.П. Мусьяков. — М.: Радио и связь, 1986. — 216 с.
6. *Айвазян С.А.* Прикладная статистика: Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. — М.: Финансы и статистика, 1989. — 607 с.
7. *Раушенбах Г.В.* Меры близости и сходства / В кн.: Анализ нечисловой информации в социологических исследованиях. — М.: Наука, 1986. — С. 169–203.
8. *Шапиро Л.* Компьютерное зрение / Л. Шапиро, Дж. Стокман; пер. с англ. — М.: БИНОМ. Лаборатория знаний, 2006. — 752 с.
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 редакцию 15.01.2011