

ОЦЕНКА СХОДСТВА СТРУКТУРНЫХ ОБЪЕКТОВ КАК МНОЖЕСТВ КОМПОНЕНТ

В.А. ГОРОХОВАТСКИЙ

Рассматриваются вопросы сопоставления изображений визуальных объектов в системах компьютерного зрения. Предложены меры сходства, учитывающие искажение и появление ложных компонент в структурном описании. Проведен анализ свойств методов. Описаны результаты компьютерного моделирования для реальных изображений.

ВВЕДЕНИЕ

Распознавание образов решает проблемы построения и применения формальных операций над отображениями объектов в некотором пространстве признаков, что отражается в отношениях эквивалентности между объектами. Эти отношения выражают принадлежность к классам, рассматриваемым как самостоятельные семантические единицы. Наиболее часто в целях классификации используются следующие методы интеллектуального анализа данных: ранжирование, сравнение, непосредственное оценивание [1–6].

Путем перехода от изображения к множеству ключевых точек (КТ) или множеству информативных фрагментов удастся достичь существенного снижения объема информации, в то время как характеристики распознавания (достоверность, помехозащищенность) при этом практически не снижаются. При представлении изображений в пространствах КТ визуальные данные можно рассматривать как мультимножества [7], т.к. значения признаков КТ бывают достаточно близкими между собой. КТ представляет собой числовой вектор фиксированной размерности с вещественными значениями, связанный с точкой пространства координат. Одним из наиболее эффективных современных способов построения описания в виде множества КТ есть применение детектора SIFT (Scale-invariant feature transform — масштабно-инвариантное преобразование признаков), который формирует КТ как вектор размерностью 128 [8]. Множество КТ является базой для формирования компонентного представления, например, путем построения отношений между КТ. Будем рассматривать распознаваемый визуальный объект как многоэлементную совокупность (объединение) *компонент*, а значение вектора признаков для компоненты назовем *дескриптором*. Компонентное представление позволяет осуществлять структурный анализ визуальной информации в виде составных объектов. Выбор типов дескрипторов и компонент в значительной мере влияет на свойства классов в плане их разделения между собой. Основной задачей компонентного сопоставления объектов есть надежное функционирование систем в неполном пространстве признаков, а также исключение ложных компонент. Для этих целей необходима разработка специальных мер сходства [4, 7, 8].

Цель работы — формализация и анализ свойств мер сходства и расстояний на множествах КТ применительно к задаче сопоставления изображений в условиях неполного представления и ложных воздействий.

Задачи исследования — построение мер сходства для конечных множеств дескрипторов, изучение особенностей применения мер для визуальных объектов, описание сходства в терминах соответствий и отношений на множестве признаков, анализ и оценка эффективности предлагаемых подходов.

КЛЮЧЕВЫЕ ТОЧКИ И КОМПОНЕНТЫ

Компонента — это либо одиночная КТ, либо некоторое множество КТ, сформированное путем построения отношений. Отношения дают возможность анализировать комбинации признаков, которые в информационном плане часто являются более значимыми для распознавания, чем сами признаки [1]. В этом аспекте не накладывается никаких требований на способ построения компонент, в частности, допускается их пространственное пересечение. В то же время сами КТ характеризуют свойства изображения в конкретных точках. Рис. 1 иллюстрирует схему компонентного представления объектов.

Задача отнесения объекта к некоторому классу решается путем оценки сходства множеств. Сходство отличается тем, что вычисляемые величины сравниваются не со шкалой, а относительно друг друга, это дает возможность наряду с упорядочиванием различать между собой объекты, которые могут включать схожие элементы. Современные технологии для построения сходства-различия в качестве первого этапа предполагают сравнение пар дескрипторов объект-эталон и получение локальных характеристик. Затем локальные значения интегрируются с целью получения глобальной (общей) меры соответствия [7–10].

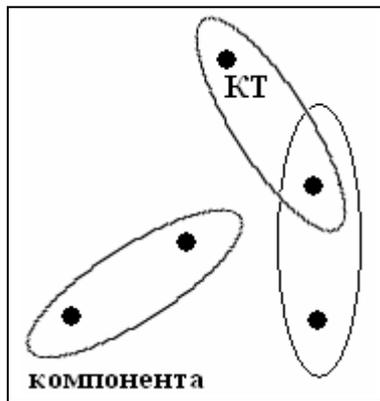


Рис. 1. Объект в виде компонент из пар КТ

Для решения задачи оценки сходства-различия применимы не все классические меры и метрики. Например, метрики Хаусдорфа и ближнего (дальнего) соседа фактически формируются по величине одного из расстояний между элементами и по этой причине не могут считаться устойчивыми к локальным искажениям. С другой стороны, интегрально учитывающая схожесть всех без исключения элементов величина «средней связи», которая вычисляется как среднее арифметическое всевозможных попарных расстояний между элементами, обладает хорошей помехозащищенностью относительно аддитивных помех, однако, к локальным искажениям она также неустойчива. Наиболее близкими к требуемым свойствам являются меры, связанные с операциями пересечения и симметрической разности множеств. Они вычисляются путем подсчета соответствий элементов (голосование). Подсчет эквивалентных компонент реализует аппроксимацию полного

соответствия объектов. Голосование компонент высоких уровней в виде совокупности КТ позволяет учесть внутреннее сходство структур объектов. В соответствии со значениями мер формируются оценки классов для реализации задач распознавания (рис. 2).

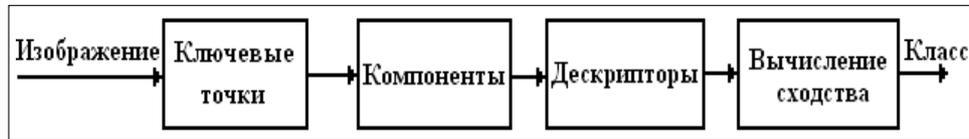


Рис. 2. Схема обработки данных при распознавании

ПРИМЕНЕНИЕ АССОЦИАТИВНЫХ МЕР И МЕТРИК

В задачах анализа двумерных сигналов и оценки биометрической информации различают три типа мер сходства-различия [3–5, 11, 12]: ассоциативные, выражающие отношения числа совпадающих признаков к общему их числу (коэффициенты связи, парные функции); корреляционные («косинусные» меры); показатели расстояния в метрическом пространстве.

Выбор конкретных мер зависит в первую очередь от цели их применения или исследования. Существующие меры тесно связаны между собой в том плане, что оценивают в конечном итоге вероятность того, что сравниваемые объекты будут идентичными. Достаточно простыми в вычислительном плане есть ассоциативные меры, которые сводятся к «подсчету» числа нужных элементов, что отражает меру частичного соответствия объектов. Кроме того, ассоциативные меры легко распространяются на случай анализа компонент. При применении мер для элементов сравниваемых множеств необходимо ввести некоторый критерий эквивалентности.

Пусть U — некоторый универсум дескрипторов $\lambda \in U$. Обозначим Λ^1, Λ^2 — конечные множества дескрипторов, $\Lambda^1, \Lambda^2 \subset U$, $\Lambda^1 = \{\lambda_i^1\}_{i=1}^{\mu(1)}$, $\Lambda^2 = \{\lambda_i^2\}_{i=1}^{\mu(2)}$, $\mu(1), \mu(2)$ — мощности множеств. Введем на U некоторое расстояние $\rho(\lambda, \lambda_0)$ между дескрипторами $\lambda, \lambda_0 \in U$, которое определяет нормированное метрическое пространство $\mathfrak{Z} = (U, \rho)$, $\rho: U \times U \rightarrow [0, 1]$. В качестве ρ может быть использована произвольная метрика, чаще всего для КТ применяют евклидово или манхэттенское расстояние в вещественном векторном пространстве R^n , $U \subseteq R^n$. Распространенными способами нормировки дескрипторов есть деление на норму вектора или на максимальное значение среди компонент [11].

Рассмотрим правило φ , соответствующее попаданию дескриптора λ внутрь шара радиуса δ с центром в λ_0

$$\varphi: \rho(\lambda, \lambda_0) \leq \delta. \quad (1)$$

Правило (1) определяет для каждого из множеств Λ^1, Λ^2 некоторое семейство шаров с параметрами δ, ρ и устанавливает бинарное отношение неразличимости дескрипторов [2]. Принадлежность шару означает эквивалентность λ, λ_0 . Параметр δ определяет степень точности при установле-

нии неразличимости. С увеличением δ растет размер конгломератов эквивалентных КТ. Заметим, что $\delta \in [0,1]$ в силу нормированной метрики ρ . Правилом (1) задаются эквивалентные подмножества элементов как между множествами Λ^1, Λ^2 , так и внутри каждого из них. Выполнение условия (1) проверяется посредством предиката

$$L[\rho(\lambda, \lambda_0), \delta] = \begin{cases} 0, & \rho(\lambda, \lambda_0) \leq \delta, \\ 1, & \text{else,} \end{cases}$$

равного 0 или 1 в зависимости от истинности неравенства и определяющего некоторое подмножество эквивалентных элементов. Применение L означает переход к представлению компоненты в виде дихотомического признака. Вторым способом анализа есть пороговая обработка вида

$$L^\delta[\rho(\lambda, \lambda_0), \delta] = \begin{cases} \rho(\lambda, \lambda_0), & \rho(\lambda, \lambda_0) \leq \delta, \\ 1, & \text{else,} \end{cases}$$

где при выполнении правила (1) сходство элементов равно значению метрики $\rho(\lambda, \lambda_0)$. Области формирования близких значений для L и L^δ совпадают, однако при использовании L^δ сохраняются значения метрики $\rho(\lambda, \lambda_0)$, которые можно использовать как веса элементов.

Применим ассоциативные меры сходства с целью определения величины близости множеств Λ^1, Λ^2 [3–5]. Определим для множеств Λ^1, Λ^2 на основе (1) операции пересечения $C = \Lambda^1 \cap \Lambda^2$ и разности $A_1 = \Lambda^1 \setminus \Lambda^2$, $A_2 = \Lambda^2 \setminus \Lambda^1$. Обозначим мощности полученных множеств как $c = \mu(C)$, $a = \mu(A_1)$, $b = \mu(A_2)$. Наиболее популярны в задачах обработки сложных сигналов меры, полученные на основе множеств C, A_1, A_2 : функция Жаккара $K_1 = c/(a + b + c)$; функция Сьеренсена $K_2 = 2c/(a + b)$; функция Дейка $K_3 = 2c/(a + b + 2c)$; функция Кульчинского $K_4 = (a + b)/(2ab)$; процент несогласия $K_5 = (a + b)/(a + b + c)$; функция Соукала и Снита $K_6 = c/(c + 2(a + b))$; несимметричный трансформированный коэффициент Дейка $K_7 = (c - \min(a, b))/(c + \min(a, b))$ и другие.

Ассоциативные меры обладают такими достоинствами, как универсальность и быстрое действие вычислений. Кроме того, они, как правило, безразмерны и ограничены отрезком $[0;1]$. Каждая из мер имеет свои особенности. Например, мера K_7 задана отрезком $[-1;1]$, а K_2 ненормирована, т.к. пересечение может превышать величину симметрической разности. Значения отдельных функций, например, K_2, K_4 не определены при совпадении множеств, т.к. при этом выполняется $a = b = 0$, поэтому меры нужно доопределить для этих случаев. Заметим также, что меры K_1, K_5 можно выразить друг через друга. Значение $(a + b)$ — мощность симметрической разности множеств Λ^1, Λ^2 , широко используемой в интеллектуальном ана-

лизе данных, связанном с проблемами грануляции информации [2, 6]. Величина $(a + b)$ есть расстояние между множествами Λ^1, Λ^2 [6]. Значение $(a + b + 2c)$ соответствует сумме мощностей $\mu(1) + \mu(2)$. Одним из критериев выбора конкретной функции есть относительная важность событий совпадения или несовпадения элементов. Например, функция K_3 придает вдвое больший вес совпадающим элементам, а функция K_6 — несовпадающим. Самым естественным вариантом функции сходства есть отношение числа совпадающих признаков к их общему числу.

В работе [3] обсуждается понятие эквивалентности мер в том плане, что две идентичные меры приводят к одной и той же последовательности объектов, упорядоченных по их сходству, близкие объекты остаются близкими и т.д. Например, можно показать, что свойством эквивалентности обладает континуум мер сходства, представленных формулой:

$$K = 2c / [(1 + \alpha)(a + b + 2c) - 2\alpha c], \quad (2)$$

где $-1 < \alpha < \infty$. Нетрудно заметить, что при $\alpha = 0$ из (2) имеем функцию K_2 , при $\alpha = 1$ — функцию K_1 , тогда $\alpha = 3$ соответствует величине K_6 , поэтому споры о том, какой из коэффициентов лучше, считаются беспредметными. Определение эквивалентности мер приводит к необходимости использования на практике именно неэквивалентных мер, оценивающих разносторонние свойства анализируемых объектов. Если, например, выводы, полученные на основе корреляционных мер сходства, совпадут с выводами анализа на основе евклидовой метрики, то с уверенностью можно утверждать, что они действительно основаны на свойствах исходных данных, а не на способе их извлечения [3]. Выражения для функций K_1, \dots, K_7 и (2) могут быть обобщены для количественных признаков, принимающих конечное множество числовых значений [5]. Здесь величины a, b, c для выражений K_1, \dots, K_7 вычисляются следующим образом:

$$a = b = \sum_{i,k} [\max(\lambda_i^1, \lambda_k^2) - \min(\lambda_i^1, \lambda_k^2)], \quad c = \sum_{i,k} \min(\lambda_i^1, \lambda_k^2). \quad (3)$$

В случае фиксированного соответствия между компонентами сравниваемых множеств, как это применено в модификациях корреляционных подходов при анализе изображений [9], реализация мер сходства упрощается и приобретает вид векторно-пространственной модели, когда необходимо осуществить сопоставление и подсчитать число сходных компонент векторов. В то же время для более общего случая в силу наличия геометрических преобразований нет возможности точно установить соответствие компонент в процессе измерений, поэтому требуется осуществлять перебор на некотором подмножестве соответствий. Наиболее полный вариант анализа соответствий реализуется путем перебора всевозможных пар $(\lambda_i^1, \lambda_k^2)$ элементов. Количество сравнений при этом оценивается произведением $\mu_{12} = \mu(1)\mu(2)$.

Схема применения более сложных в вычислительном аспекте корреляционных мер или расстояний для задач компонентного представления прак-

тически не отличается от использования ассоциативных мер. Преимущество метрик состоит в возможности построения на их основе быстродействующих процедур поиска [6]. Принципиальным моментом, позволяющим осуществить анализ структуры объектов, есть включение в процесс обработки отбора близких дескрипторов, что дает возможность оценить частичное сходство. Оценка сходства путем подсчета совпадений элементов приводит к мысли о возможности применения на множестве соответствий расстояния Хемминга [5], в классическом представлении подсчитывающего число различающихся элементов для двух последовательностей одинаковой длины. Для оценки степени соответствия множеств Λ^1, Λ^2 можно предложить следующую модификацию метрики Хемминга

$$\rho_H(\Lambda^1, \Lambda^2) = [\mu_{12}]^{-1} \sum_{i=1}^{\mu(1)} \sum_{k=1}^{\mu(2)} L[\rho(\lambda_i^1, \lambda_k^2), \delta], \quad (4)$$

где в качестве нормировочного коэффициента используется произведение мощностей μ_{12} . Значение ρ_H принадлежит отрезку $[0,1]$. Метрика (4) по сравнению с другими обладает важным свойством структурного анализа данных внутри сравниваемых множеств. Запишем выражение для метрики $\rho_\Lambda(\Lambda^1, \Lambda^2)$ на основе обработки L^δ в следующем виде:

$$\rho_\Lambda(\Lambda^1, \Lambda^2) = [\mu_{12}]^{-1} \sum_{i=1}^{\mu(1)} \sum_{k=1}^{\mu(2)} L^\delta[\rho(\lambda_i^1, \lambda_k^2), \delta]. \quad (5)$$

Соотношение (5) как линейную комбинацию метрик можно считать метрикой [11]. Метрика (5) является модификацией расстояния «средней связи», однако, в отличие от (4) не отражает напрямую числа различающихся элементов множеств и поэтому носит символический характер «средней температуры». Сумма значений в пределах порога δ при фиксированном μ_{12} может дать лишь приближенную оценку числа таких элементов. Заметим, что в случае, когда для всех элементов первого множества выполняется правило (1), величина (5) равна значению метрики средней связи. В частном случае ненулевое значение метрики (5) может быть получено даже по одной паре схожих компонент, что может быть недопустимым с точки зрения достоверности решения. В таком случае метрику (5) нужно дополнить логическим условием вида

$$\rho_\Lambda^\varepsilon(\Lambda^1, \Lambda^2) = \begin{cases} \rho_\Lambda(\Lambda^1, \Lambda^2), & N_L > N^\varepsilon, \\ \rho^\varepsilon, & \text{else,} \end{cases} \quad (6)$$

где N^ε — некоторое пороговое значение для минимально допустимого числа соответствий, N_L — реальное число соответствий, полученное в процессе вычисления (5) и равное числу неединичных значений предиката L ; ρ^ε — некоторое символическое значение метрики, означающее ситуацию «полного отсутствия» сходства при введенных ограничениях. Логический анализ вида (6) дает возможность обеспечить заданную достоверность принятия решения о сходстве, определяемую числом имеющихся соответствий. Еще один вариант модификации метрики средней связи можно получить нормировкой на число N_L близких элементов

$$\rho_{\Lambda}(\Lambda^1, \Lambda^2) = [N_L]^{-1} \sum_{i=1}^{\mu(1)} \sum_{k=1}^{\mu(2)} L^N[\rho(\lambda_i^1, \lambda_k^2), \delta], \quad (7)$$

где предикат L^N отличается от L^δ заменой 1 на 0. Принципы, использованные при построении (4)–(7), можно распространить и на другие метрики.

СХОДСТВО В ТЕРМИНАХ АНАЛИЗА СООТВЕТСТВИЙ

Опишем множество θ_i^{12} соответствий элемента $\lambda_i^1 \in \Lambda^1$ во множестве Λ^2 в виде $\theta_i^{12} = \{\lambda_k^2 \in \Lambda^2 : L^*[\rho(\lambda_i^1, \lambda_k^2), \delta]\}$, где L^* — один из рассмотренных предикатов. В частных случаях множество θ_i^{12} может включать лишь один элемент (однозначное соответствие) или вообще не содержать элементов $\theta_i^{12} = \emptyset$ (пустое соответствие). Аналогичным образом описываются множества соответствий θ_i^{11} и θ_i^{22} внутри объектов Λ^1, Λ^2 . Совместный анализ множеств соответствий $\theta_i^{11}, \theta_i^{22}$ может использоваться для оценки потенциальных возможностей различения объектов в пространстве дескрипторов. Учитывая, что мощности множеств соответствий и несоответствий для конечных множеств связаны между собой однозначным образом, меру множества соответствий считаем мерой сходства.

Основой для вычисления сходства (4), (5) компонентных объектов является прямоугольная матрица расстояний $\rho(i, k) = \rho(\lambda_i^1, \lambda_k^2)$, $i = 1, \dots, \mu(1)$; $k = 1, \dots, \mu(2)$. Строка матрицы $\rho(i, k)$ соответствует расстояниям i -го дескриптора первого объекта до каждого из дескрипторов второго. Перейдем от матрицы расстояний к ее представлению в виде матрицы соответствий $\Theta(i, k)$ тех же размеров путем реализации отображения $\Omega : \rho(i, k) \rightarrow \Theta(i, k)$, и на основе $\Theta(i, k)$ определим функцию сходства. Построение матрицы $\Theta(i, k)$ и ее анализ сводится к нескольким возможным вариантам. Во-первых, должен быть выбран однозначный или множественный тип соответствия θ_i^{12} дескрипторов, который определяется количеством элементов множества Λ^2 , которые могут считаться эквивалентными элементу из Λ^1 . При однозначном типе множество θ_i^{12} содержит максимум один элемент, тогда строка матрицы $\Theta(i, k)$ может содержать только один неединичный элемент (единица означает отсутствие соответствия). При множественном соответствии в строке $\Theta(i, k)$ может быть произвольное или строго фиксированное число элементов (в пределах от 0 до $\mu(2)$).

Во-вторых, в зависимости от применяемого предиката задается бинарный или многозначный вид соответствия. Бинарная обработка реализуется предикатом L и матрица $\Theta(i, k)$ приобретает двоичный вид. При многозначном представлении матрица $\Theta(i, k)$, кроме единиц, содержит элементы $\rho(i, k)$, т.е. $\Theta(i, k) = [\rho(i, k) | L^\delta[\rho(i, k), \delta]]$, где отличие от 1 означает соответствие.

Каждый из вариантов обработки имеет свои особенности применения. Например, при множественном многозначном представлении имеем матрицу, каждая строка которой содержит определенное число элементов со значениями из отрезка $[0, \delta]$, что при построении общего сходства требует дополнительных действий по выбору одного из них либо нормировки. Учет множественных соответствий предполагает при формировании общего сходства выбор одного из вариантов для веса соответствия: все соответствия равноценны с весом 1; соответствия равноценны с весом, равным $1/m$, где m — их число в строке; вес соответствия определяется величиной $\rho(i, k)$ и т.д. Однако множественные соответствия в целом оказываются устойчивее к действию помех, чем однозначные [7, 10]. Кроме того, при получении однозначных соответствий по сравнению с множественными необходимы дополнительные вычисления, связанные с поиском оптимумов, ранжированием и т.д. В методах SIFT получил применение вариант однозначного способа установления соответствий с дополнительной проверкой надежности такого решения путем анализа величины второго оптимума [8].

Опишем теперь сходство как функцию $\Theta(i, k)$. Одним из вариантов есть выражение

$$\rho_{\Lambda}(\Lambda^1, \Lambda^2) = [\mu_{12}]^{-1} \sum_{i=1}^{\mu(1)} \sum_{k=1}^{\mu(2)} \Theta(i, k), \quad (8)$$

представляющие собой значение метрики (4), (5) с учетом типа соответствия. Для случая однозначных бинарных соответствий значение суммы в (8) есть число строк, содержащих ноль, что соответствует числу «одинаковых» элементов множеств. В случае многозначных бинарных соответствий сумма в соотношении (8) означает общее число всех не совпавших элементов (с учетом повторений). Для других типов соответствия значение (8) можно отнести к модификациям метрик: ближнего соседа (однозначные соответствия), m ближайших соседей (фиксированное число m наиболее сходных элементов), метрики средней связи (многозначные). Во всех перечисленных модификациях учитываются близкие элементы множеств, для сходства которых выполнена пороговая обработка L^{δ} .

Заметим, что нормировка в (8) играет важную роль для анализа сходства. Например, при бинарных однозначных соответствиях для двух одинаковых множеств с различающимися элементами (без повторений) значение суммы в (8) равно $\mu_{12} - \mu(1)$, или $\mu(1)[\mu(2) - 1]$, а величина (8) для этой ситуации равна $\rho_{\Lambda} = 1 - 1/\mu(2)$. В то же время для идентичных множеств $\rho_{\Lambda} = 0$. Это говорит о необходимости применения различных подходов для этих ситуаций, по-своему отражающих событие совпадения множеств. Например, более практичным вариантом для случая неповторяющихся элементов вместо метрики (8) может оказаться сходство

$$K = [\min[\mu(1), \mu(2)]]^{-1} \sum_{i=1}^{\mu(1)} \sum_{k=1}^{\mu(2)} [1 - \Theta(i, k)], \quad (9)$$

равное 1 при полном совпадении. Выражение (9) точнее отражает суть сопоставления множеств компонент как подсчет относительной доли одинаковых элементов. В то же время (8) более универсально и не требует дополнительной проверки повторяемости элементов.

При построении мер сходства в задачах распознавания, когда анализируемый объект (например, множество Λ^2) поочередно сравнивается с базой эталонов, нормировка при определении доли однозначных соответствий может быть реализована путем деления на $\mu(1)$, которое в данном случае отражает максимально возможное их число. Тогда мера сходства есть доля элементов эталона, которые нашли свое соответствие в объекте, и вычисляется как $K = \eta / \mu(1)$, где числитель η равен числу установленных соответствий.

Важным для задач анализа визуальных данных является также расположение нулей в матрице соответствий. Понятно, что если в каждой строке и каждом столбце содержится только один ноль, то мы имеем дело с одновременным соответствием нескольких пар элементов разных множеств, что должно отражаться на значении метрики. Этим качеством описанные метрики и меры сходства не обладают, что говорит о необходимости их дальнейшего усовершенствования. В некоторой мере это учтено в мерах на основе отношений, изложенных ниже.

На базе значений i -той строки матриц $\rho(i, k)$ или $\Theta(i, k)$ может быть сформирована промежуточная мера сходства типа элемент-множество, на основе которой в дальнейшем можно вычислить сходство двух множеств. Особенно актуален такой вид сходства для тех задач распознавания, где для каждого структурного элемента объекта вначале осуществляется оценка класса, а затем по совокупности полученных оценок принимается глобальное решение о классе объекта. С другой стороны, практика задач компьютерного зрения показывает, что результат распознавания по сходству множеств часто оказывается более достоверным в условиях помех, чем принятие решения по совокупности локальных решений.

МЕРЫ НА МНОЖЕСТВЕ ОТНОШЕНИЙ

Введем на универсуме U дескрипторов некоторое отношение R_U , $R_U \subseteq U \times U \times \dots \times U$. Для конкретности рассмотрим бинарные отношения $R_U \subseteq U \times U$ и сформулируем принципы построения и сравнения отношений из множества R_U . Наиболее применимы в анализе данных отношения неразличимости (равенство), предшествования (ранжирование) и соседства [2]. При анализе изображений дополнительно рассматривают два типа отношений: амплитудные и пространственные [7]. Амплитудные отношения характеризуют связь на множестве значений дескрипторов (неразличимость), пространственные — связь между их координатами, отражающую важные для визуальных данных геометрические аспекты (соседство).

Бинарное амплитудное отношение $r^{11} \in R_U$ для элементов $\lambda_i^1, \lambda_j^1 \in \Lambda^1$ множества Λ^1 может быть описано в виде совокупности $r^{11} = \{(\lambda_i^1, \lambda_j^1) \mid R(\lambda_i^1, \lambda_j^1)\}$, где $R(\lambda_i^1, \lambda_j^1)$ (или $R[r^{11}]$) — правило, формирующее отношение. Примером $R(\lambda_i^1, \lambda_j^1)$ могут быть предикаты L, L^δ .

Конкретно для элементов первого множества отношение имеет вид: $r^{11} = \{(\lambda_i^1, \lambda_j^1) \mid \rho(\lambda_i^1, \lambda_j^1) \leq \delta\}$ или $r^{11} = \{(\lambda_i^1, \lambda_j^1) \mid L\}$. Аналогично можно описать отношение r^{12} и между элементами множеств Λ^1, Λ^2 . Пространственное бинарное отношение для элементов $\lambda_i^1, \lambda_j^1 \in \Lambda^1$ имеет вид $r_c^{11} = \{(c_i^1, c_j^1) \mid R_c(c_i^1, c_j^1)\}$ или конкретно $r_c^{11} = \{(c_i^1, c_j^1) \mid \rho_c(c_i^1, c_j^1) \leq \delta_c\}$, где c_i^1, c_j^1 — координаты элементов λ_i^1, λ_j^1 , ρ_c — метрика в координатном пиксельном пространстве, δ_c — порог, определяющий размер пространственной окрестности, в пределах которой элементы λ_i^1, λ_j^1 считаются удовлетворяющими отношению R_c . Как правило, координатные отношения имеют смысл близости или соседства. Величина δ_c должна быть согласована с соответствующим значением для множества Λ^2 с учетом допустимых геометрических преобразований и определяет возможность «склеивания» (в плане совместного анализа) для пары компонент между собой. Другой пример правила R_c может быть связан с заданием бинарных пространственных отношений в виде дуальных графов [12], отражающих связь пар компонент.

Правило для формирования отношения r^{12} представим в виде конъюнкции

$$R[r^{12}] : (\rho(\lambda_i^1, \lambda_j^2) \leq \delta) \& (\rho(\lambda_k^1, \lambda_l^2) \leq \delta).$$

Установим теперь правило $R \& R_c$ для одновременного применения $r^{12}, r_c^{11}, r_c^{22}$ в виде

$$R \& R_c[r^{12}, r_c^{11}, r_c^{22}] : (\rho(\lambda_i^1, \lambda_j^2) \leq \delta) \& (\rho(\lambda_k^1, \lambda_l^2) \leq \delta) \& (\rho_c(c_i^1, c_k^1) \leq \delta_c) \& (\rho_c(c_j^2, c_l^2) \leq \delta_c), \quad (10)$$

которое означает выполнение (1) для каждой из пар элементов (амплитудное отношение r^{12}) и двух условий для соответствующих координат, реализующих пространственное отношение $r_c^{kk}, k = 1, 2$. В целом выражение (10) можно представить как новое отношение r_4^{12} для четверки:

$$r_4^{12} = \{(\lambda_i^1, c_i^1), (\lambda_k^1, c_k^1), (\lambda_j^2, c_j^2), (\lambda_l^2, c_l^2) \mid r^{12}(\lambda_i^1, \lambda_j^2), r^{12}(\lambda_k^1, \lambda_l^2), r_c^{11}(c_i^1, c_k^1), r_c^{22}(c_j^2, c_l^2)\}. \quad (11)$$

Отношение (11) можно трактовать как установление соответствия для пары точек двух разных объектов, причем координаты точек, в свою очередь, включены в пространственные отношения r_c^{11}, r_c^{22} внутри каждого из объектов. Выражение (11) можно использовать для построения новых компонент. Оно легко распространяется на произвольное число пар компонент, пространственное отношение которых фиксирует некоторую геометрию визуального объекта.

В целом отношение (11) построено на метриках ρ , ρ_c и может быть использовано в выражениях для ассоциативных мер K_1, \dots, K_7 и метриках типа (4)–(7). При этом фактически будет реализован подсчет количества отношений, установленных для элементов двух множеств. Здесь вместо предиката L необходимо использовать новый предикат L' , проверяющий выполнение отношения вида (11). В метриках на основе обработки L^δ и сходствах, связанных с ранжированием, дополнительно можно численно оценивать степень выполнения отношения (11). Примером такой оценки есть сумма $\rho(\lambda_i^1, \lambda_j^2) + \rho(\lambda_k^1, \lambda_l^2)$.

Некоторые координатные подмножества точек, применяемые в ряде задач компьютерного зрения, дают возможность построить геометрические инварианты относительно аффинных преобразований [10]. В этом случае координатные отношения точек в виде значений геометрических инвариантов могут быть дополнены значениями соответствий их дескрипторов, устанавливаемых посредством (10), (11), что в целом даст возможность усовершенствовать существующие подходы и повысить достоверность решений.

Применение отношений по сравнению с использованием одиночных дескрипторов в плане достоверности распознавания имеет существенно лучшие показатели. Заметим, что предлагаемая схема анализа визуальных данных путем построения отношений может быть успешно расширена на любую арность (число элементов), которые они содержат. С другой стороны, увеличение числа элементов приводит к усилению интегральных качеств системы сопоставления, одновременно снижая устойчивость к локальным искажениям элементов множеств. Учитывая направленность таких схем обработки на устранение локальных искажений, целесообразно применять только двух- или трехместные отношения КТ, причем в целях ускорения вычислений использовать не полные, а целенаправленно сформированные усеченные подмножества отношений [10].

ЭКСПЕРИМЕНТЫ

Предварительный расчетный анализ рассмотренного разнообразия мер структурного соответствия показал следующее. Меры K_2 , K_4 в силу своих особенностей не обладают свойством непрерывности при изменении объема информации, выражающемся в числе дескрипторов, в меньшую или большую сторону. Мера K_7 также слабо реагирует на изменения структуры объекта. Метрика (5) при удачно выбранном пороге δ практически не отличается от метрики (4), и ее целесообразно применять лишь с нормировкой на число «близких» элементов аналогично (7). Среди остальных мер из набора K_1, \dots, K_7 следует выделить K_1 , K_3 , K_5 , в одинаковой степени реагирующие на добавление и удаление дескрипторов списка, а при необходимости выбора предпочтение следует все-таки отдать мере K_1 , более точно отражающей уровень структурных изменений по сравнению с K_3 .

В целях более тщательного качественного анализа обсуждаемых мер для практических наборов признаков проведены компьютерные экспери-

менты на реальных изображениях. В качестве компонент использованы множества КТ, полученные применением варианта детектора SIFT для базы данных, содержащей 30 полутоновых изображений аквариумных рыбок размером 100×100 пикселей, описанной в [7]. На рис. 3 показано изображение одного из эталонов, а также размещенные в оверлейном режиме на нем и еще на трех эталонах дескрипторы (темные точки). Число дескрипторов на выделенном эталоне равно 42.

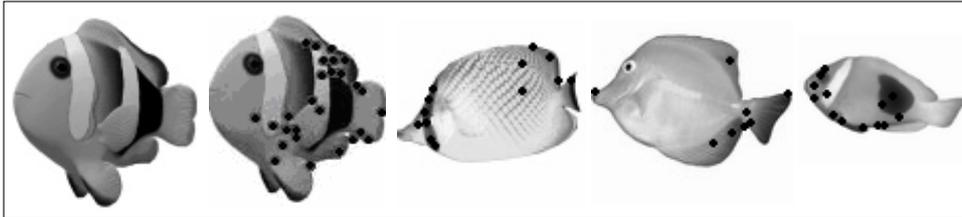


Рис. 3. Эталоны и характерные признаки

По результатам моделирования представлена таблица значений мер при сравнении выбранного эталона в пространстве признаков с собой (столбец 1) и с другими 4-мя эталонами с примерно таким же числом признаков. Наряду со значениями мер K_1, \dots, K_7 и метрик (4), (5) в таблицу включены значения меры для бинарных соответствий на базе отношения r^{12} , вычисленной как $K_9 = \eta_r / \mu_r$, где η_r — число полученных соответствий пар, μ_r — общее число пар. Рассмотрены две разновидности: K_9^1 — однозначные соответствия с нормировкой на число пар в Λ^1 и K_9^2 — многозначные соответствия (порог) с нормировкой на произведение числа пар. В целях сравнения приведены также аналогичные значения мер K_8^1, K_8^2 для однозначных соответствий, установленных по оптимуму сходства. Для сопоставления дескрипторов использовано евклидово расстояние векторов, а значение порога выбрано $\delta = 0,113$, что соответствует 1% от максимального значения метрики, равного $\sqrt{128}$. Как видим из таблицы, значения мер (4) и (5) отличаются, только начиная со 2-го знака после запятой, что говорит об удачно выбранном пороге δ , отражающем близость дескрипторов. При этом меры (4) и (5) на «своем» эталоне дают величину сходства около 0,88, что несколько отклоняется от минимально возможного значения, равного 0. Несмотря на это, с применением всех обсуждаемых мер без помех достигается безошибочное распознавание на анализируемой базе из 30 эталонов. Значения мер для реальных изображений, приведенные в таблице, в целом подтверждают предварительные оценки.

Оценим теперь свойства мер с точки зрения влияния на их величину искажений типа исключения КТ или появления ложных КТ. Уровень помехи добавления ложного элемента задавался вероятностью β_1 , а вероятность β_2 отражала событие исключения признака из списка КТ. Заметим, что для ряда мер, где нормировка происходит путем деления на величину μ_{12} , действие этих двух типов помех имеет некоторые различия. В частности, бывают ситуации, когда исключение общих дескрипторов не изменяет значения меры, что означает ее нечувствительность к таким искажениям.

Координаты и значения ложных дескрипторов в эксперименте формировались с использованием равномерного распределения. В таблице в качестве примера приведены величины сходства исходного и искаженного множеств дескрипторов объекта на рис. 3 при значениях $\beta_1 = 0,25$ и $\beta_2 = 0,25$, что соответствует изменению примерно четвертой части состава дескрипторов. Из таблицы путем сравнения с первым столбцом видим, что все рассматриваемые меры в той или иной степени реагируют на локальные искажения, однако более точно отражают действие таких помех меры $K_1(K_5)$, K_8 , K_9 и метрики (4), (5). Мера K_7 не реагирует на искажения такого типа. Анализ пар в мерах K_8^2 , K_9^2 по отношению к одиночным соответствиям, как видим, обеспечивает большую чувствительность к помехам исключения, а также более высокую достоверность решений относительно других эталонов базы, т.к. локальные оптимумы мер относительно глобального оптимума здесь менее значительны, чем в мерах K_8^1 , K_9^1 . Отметим, что для мер (4), (5) при помехе исключения наблюдаются значения ниже, а у мер K_8^2 , K_9^2 выше, чем у эталона. Это объясняется изменением базы суммирования при помехе исключения и говорит о сложности использования в этих мерах пороговых способов для принятия решений. Несмотря на эти особенности, все обсуждаемые меры обеспечивают правильное распознавание изображения на рис. 3 для используемой базы данных при заданном уровне помех.

Таблица. Значения мер для множеств дескрипторов

Функция	1	2	3	4	5	$\beta_1 = 0,25$	$\beta_2 = 0,25$
K_1	1	0,68	0,64	0,54	0,76	0,85	0,89
K_2	84	4,21	3,53	2,38	6,33	5,5	6,8
K_3	1	0,81	0,78	0,70	0,86	0,92	0,94
K_4	0	0,28	0,14	0,15	0,19	0,5	0,5
K_5	0	0,31	0,36	0,46	0,24	0,15	0,11
K_6	1	0,51	0,47	0,37	0,61	0,73	0,79
K_7	1	0,90	0,71	0,72	0,81	1	1
(4)	0,8775	0,9494	0,9324	0,9259	0,9365	0,8992	0,874
(5)	0,8826	0,9520	0,9355	0,9296	0,9400	0,9034	0,8792
K_8^1	1	0,7142	0,9523	0,5952	0,9047	0,8461	1
K_8^2	0,1224	0,0676	0,0505	0,0741	0,0663	0,1007	0,1260
K_9^1	1	0,53	0,74	0,46	0,62	0,87	1
K_9^2	0,0107	0,0044	0,0028	0,0046	0,0043	0,0073	0,0115

ВЫВОДЫ

Методы вычисления мер эквивалентности для компонентного представления визуальных объектов в условиях искажений компонент сводятся к по-

строению частичных мер для множеств, учитывающих модель искажений. Основная идея состоит в построении модификаций, связанных с отдельным анализом сходства-различия комбинаций компонент. Применение аппарата соответствий сосредотачивает решение проблемы на выделении наиболее значимых данных, а построение отношений для дескрипторов позволяет обеспечить высокую достоверность решений. С точки зрения вычислительных затрат и возможности реагирования на определенный уровень частичных искажений предпочтение следует отдать мерам K_1 , K_8 , K_9 , (4), (5).

Впервые показано, как формализовать и применить метрические подходы для сопоставления структурных визуальных объектов в признаковом пространстве, изучены особенности известных подходов и обсужден синтез мер с новыми свойствами.

Практически важным есть сравнение принципов построения и нормировки мер, получение конкретных характеристик на используемой экспериментальной базе изображений, что подтверждает целесообразность применения предложенных модификаций в задачах компьютерного зрения.

Перспективы исследования состоят в разработке теоретических основ для построения интегрированного сходства по совокупности разнотипных признаков, включающих как значения, так и отношения разной арности, а также оценка предельных характеристик методов в задаче распознавания объектов при действии фоновых и локальных мешающих воздействий.

ЛИТЕРАТУРА

1. *Поспелов Д.А.* Искусственный интеллект. Модели и методы: Справочник. — В 3-х кн. Кн. 2. — М.: Радио и связь, 1990. — 304 с.
2. *Миркин Б.Г.* Анализ качественных признаков и структур. — М.: Статистика, 1980. — 319 с.
3. *Шитиков В.К.* Количественная гидроэкология: методы системной идентификации. — Тольятти: ИЭВБ РАН, 2003. — 463 с.
4. *Баклицкий В.К.* Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. — М.: Радио и связь, 1986. — 216 с.
5. *Елисеева И.И.* Группировка, корреляция, распознавание образов (Статистические методы классификации и измерения связей). — М.: Статистика, 1977. — 144 с.
6. *Каграманян А.Г.* Метрические свойства грануляции информации // Бионика интеллекта. — 2007. — № 1 (66). — С. 17–24.
7. *Гороховатский В.А.* Структурное распознавание изображений на основе моделей голосования признаков характерных точек // Реєстрація, зберігання і обробка даних. — 2008. — Т. 10, № 4. — С. 75–85.
8. *Lowe D.G.* Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. — 2004. — **60**, № 2. — P. 91–110.
9. *Путятін Є.П.* Методи та алгоритми комп'ютерного зору: навч. посіб. — Х.: ТОВ «Компанія СМІТ», 2006. — 236 с.
10. *Гороховатский В.А.* Иерархия пространственных отношений структурных признаков в задачах сопоставления визуальных объектов // Системи управління, навігації та зв'язку. — Київ: Центральний наук.-досл. ін-т навігації і управління, 2008. — Вип. 3 (7). — С. 85–89.
11. *Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Классификация и снижение размерности: справ. изд. — М.: Финансы и статистика, 1989. — 607 с.
12. *Дуда Р.* Распознавание образов и анализ сцен. — М.: Мир, 1976. — 512 с.

Поступила 26.01.2009