

УДК 004.932.2:004.93'1

В.А. Гороховатский, Т.В. Полякова

Харьковский национальный университет радиоэлектроники, Харьков

КРИТЕРИИ И МОДЕЛИ СТРУКТУРНОЙ КЛАССИФИКАЦИИ С ПРИМЕНЕНИЕМ ПРИНЦИПА ГОЛОСОВАНИЯ

Обсуждаются пути построения моделей и критериев для классификации структурных описаний видео-объектов с использованием голосования. Путем выбора моделей для структурных мер подобия с учетом их параметрического определения предлагается построение системы классификации с высоким уровнем устойчивости к аддитивным и пространственным помехам. Анализируются результаты экспериментов.

Ключевые слова: классификация изображений, структурно-иерархические методы, характерные признаки, структурное описание, принцип голосования, критерий классификации, модель сопоставления, достоверность классификации.

Введение и постановка задач

Среди методов структурной классификации, основанных на сопоставлении описаний объектов в виде множеств характерных признаков (ХП), одними из наиболее эффективных с точки зрения полноты или точности структурного соответствия считаются методы разметки и удовлетворения ограниченный [1, 2]. К ним относится анализ контекста формы, который для случая взаимно-однозначного соответствия элементов описаний сводится к задаче

поиска паросочетаний двудольного графа. Однако применение этих методов к задачам классификации видео-объектов, например, при решении проблемы распознавания объектов по фотографиям или в видео-поток, из-за огромного разнообразия описаний и типов ХП и, в то же время, из-за необходимости анализа близости значений ХП внутри описаний приводит к трудоемким вычислениям [1].

Еще более усложняется применение обсуждаемых методов по причине влияния традиционных визуальных искажений в виде аддитивных и простран-

ственных помех. Аддитивные помехи требуют рассмотрения элементов описания вместе с некоторой окрестностью – ограниченным шаром в пространстве структурных признаков. Пространственные искажения в виде локальных помех, мешающих и фоновых объектов приводят, как правило, к потере части признаков или к появлению в составе описания ложных элементов. Применяемые методы должны быть устойчивыми к таким помехам [3].

Развиваемые структурно-иерархические методы классификации (СИМ) [3, 4], основанные на иерархическом анализе с использованием принципа голосования элементов и/или их пространственно-атрибутивных отношений, за счет некоторого снижения степени полноты структурного соответствия значительно упрощают процедуру классификации и обеспечивают решение задач в реальном времени с достаточно высоким уровнем достоверности. Кроме того, эти методы дают возможность управления параметрами для степени эквивалентности ХП и числа ХП, необходимого для обеспечения надежного принятия решения о классе объекта в зависимости от приложений.

Цель исследования – анализ принципов построения и свойств моделей для мер подобия структурных описаний в задаче классификации видеообъектов.

Задачи исследования – изучение моделей подобия структурных описаний и критериев классификации, исследование преимуществ предлагаемых моделей при их реализации, анализ результатов экспериментальных исследований.

Модели классификации

Представим две основные модели классификации с применением СИМ: 1) анализ классов ХП, составляющих описание; 2) анализ классов описаний в виде множеств ХП. Конечная цель классификации непосредственно реализуется моделью 2, которую, в том числе, можно получить и с использованием модели 1. В некоторых вариантах СИМ классы моделей 1 и 2 совпадают [3]. Вид применяемой модели определяет меру подобия и критерий классификации в соответствующем пространстве признаков.

Представим классы объектов эталонными описаниями в виде конечных множеств ХП, которые вычисляются для базы видеоданных, где осуществляется классификация. Пусть классифицируемое описание $Z = \{z_i\}_{i=1}^s$ представлено конечным множеством ХП (s – число элементов), эталонные описания – $Z(j) = \{z_i^j\}_{i=1}^{s_j}$, $j = 1, \dots, J$, J – конечное число классов, s_j – мощность $Z(j)$. ХП – это числовой вектор фиксированной размерности $z_i, z_i^j \in R^n$, определенный в отдельной точке изображения [3].

Для модели 1 вводится конечное число q классов $K = \{K_1, \dots, K_q\}$ в пространстве ХП, каждый элемент $z_i \in Z$ относится к одному из классов путем реализации отображения $\Omega: Z \rightarrow K$. В результате описание Z будет представлено объединением $Z = \cup Z_d$ непересекающихся наборов Z_d элементов из K_i . Элементы $z \in K_i$ считаются эквивалентными. На практике модель 1 сложно применить к изображениям, т.к. разные визуальные объекты часто состоят из значительного числа близких ХП. Степень близости напрямую зависит от размера фрагмента, на котором формируются ХП. Построение K затруднено также необходимостью получения конечного числа стабильных классов для неограниченного разнообразия фрагментов реальных изображений [1].

Классификация на основе сопоставления «элемент-множество»

В рамках модели 2 практически целесообразным есть вариант, связанный с интерпретацией ХП объекта как элемента одного из классов $Z(j)$ [1, 5]. Критерием классификации выступает число ХП, отнесенных к $Z(j)$. Механизм оценки связан с голосованием $z_i \in Z$. Описание Z будет отнесено к классу, получившему большинство голосов.

Рассмотрим событие, состоящее в том, что признак $z \in Z$ относится к классу j (проголосует за класс j) и подсчитаем вероятность $p(z, j)$ этого события.

Свяжем вероятностное представление с оценкой меры соответствия признака каждому из классов. Для фиксированной базы описаний $Z(j)$ вероятность $p(z, j)$ можно вычислить через частное двух значений меры, которые выражают число благоприятных случаев к их общему числу:

$$p(z, j) = \xi(z, j) / \sum_{i=0}^J \xi(z, i), \quad (1)$$

где $\xi(z, j) \geq 0$ – неотрицательная функция, отражающая величину соответствия z классу j . Для сопоставления элемент-множество вычислим

$$\xi(z, j) = \exp[-\rho(z, Z(j))],$$

где $\rho(z, Z(j)) = \min_{z_i^j \in Z(j)} \rho_z(z, z_i^j)$ – расстояние от z до

$Z(j)$, а в качестве $\rho_z(z, z_i^j)$ используем евклидову метрику в R^n . Вероятности $p(z, j)$ сведем в табл. 1. Строка таблицы-матрицы $\{p(z_i, j)\}$ содержит вероятности отнесения z_i к классу j . Из вычисления $p(z_i, j)$ следует, что $\sum_{j=1}^J p(z_i, j) = 1$.

Таблица 1
Матрица $\{p(z_i, j)\}$

Номер $z_i \in Z$	Классы объектов			
	1	2	...	J
1	$p(z_1, 1)$	$p(z_1, 2)$...	$p(z_1, J)$
...
i	$p(z_i, 1)$	$p(z_i, 2)$...	$p(z_i, J)$
...
s	$p(z_s, 1)$	$p(z_s, 2)$...	$p(z_s, J)$

Рассмотрим варианты построения критериев классификации и соответствующие им модели для сопоставления элемент-множество.

1. *Независимое голосование XII.* Голос $h(z_i, j)$ признака $z_i \in Z$ отдается классу, которому соответствует максимум в строке $\{p(z_i, j)\}$, т.е.

$$h(z_i, j) = \begin{cases} 1, j = \arg \max_{k \in J} p(z_i, k), \\ 0, j \neq \arg \max_{k \in J} p(z_i, k). \end{cases} \quad (2)$$

Решение (2) соответствует выбору максимальной альтернативы путем $h(z_i, j) = 1$. На основе (2) для всех $z_i \in Z$ подсчитываем количества голосов каждого класса и вычисляем $h(Z, j) = \sum_i h(z_i, j)$. Значения $h(Z, j)$ – целое из диапазона $[0, s]$. Решение о классе объекта с описанием Z принимаем в виде

$$j^* = \arg \max_j h(Z, j), \quad (3)$$

что соответствует максимуму апостериорной вероятности $P(j|Z)$ отнесения объекта к классу, которая в данном случае аппроксимируется посредством суммы величин (2) [5]. Значение вероятности $P(j|Z)$ можно оценить отношением $\mu(Z(j^*)) / s$ числа $\mu(Z(j^*))$ признаков, проголосовавших за класс j^* , к общему числу s признаков объекта.

2. *Интегрированное голосование с весами.* Вычислим для каждого класса показатель $\gamma(Z, j)$ как сумму элементов столбца матрицы $\{p(z_i, j)\}$ по всем $z_i \in Z$ для фиксированного j : $\gamma(Z, j) = \sum_{i=1}^s p(z_i, j)$. Он отражает вес класса j в анализируемом описании. В общем случае $\gamma(Z, j) \in R^+$, однако, за счет введения бинарной пороговой обработки вида

$$L[p(z_i, j), \epsilon_p] = \begin{cases} 1, p(z_i, j) \geq \epsilon_p, \\ 0, p(z_i, j) < \epsilon_p, \end{cases} \quad (4)$$

с порогом значимости ϵ_p значения $\gamma(Z, j) = \sum_{i=1}^s L[p(z_i, j), \epsilon_p]$ могут быть сведены к

дискретной величине из $[0, s]$ и после нормировки интерпретированы как аппроксимация апостериорной вероятности $P(j|Z)$ с учетом независимости элементов $z_i \in Z$ [6]. Решение о классе принимаем в виде

$$j^* = \arg \max_j \gamma(Z, j). \quad (5)$$

Важным моментом при классификации в соответствии с (3),(5) и при обработке (2) есть значение m_1 глобального и последующих по величине локальных максимумов, в частности, величина второго (ближайшего) максимума m_2 . Величина m_2 отражает значение следующей возможной альтернативы (класса). Качество классификации, характеризуемое уровнем достоверности, можно оценить отношением m_2 / m_1 локального и глобального максимумов.

3. *Анализ подмножеств XII.* Учитывая то обстоятельство, что значения $p(z_i, j)$ отражают подобие $z_i \in Z$ к классам-множествам, набор критериев и схем обработки может быть расширен за счет применения методов анализа множеств, например, принципа «ближних соседей» или оценки подобия подмножеств [1]. В качестве значения альтернативы класса можно применить сумму нескольких наибольших элементов столбца, для чего нужно предварительно осуществить сортировку или пороговую обработку (4). Другим примером может быть учет при вычислении $h(Z, j)$ голосов не одного, а нескольких максимумов в строке или пороговая обработка вида

$$Q[p(z_i, j), \epsilon_p] = \begin{cases} p(z_i, j), p(z_i, j) \geq \epsilon_p, \\ 0, p(z_i, j) < \epsilon_p. \end{cases} \quad (6)$$

Такие схемы более гибкие и устойчивые к разным типам помех по сравнению с независимым или весовым голосованием, т.к. в них совместно анализируется несколько альтернатив. В то же время усложнение обработки приводит к увеличению времени классификации.

На практике анализ множества $\{p(z_i, j)\}$ во всех приведенных схемах можно значительно упростить, предварительно исключив из рассмотрения близкие к нулю значения $p(z_i, j)$, которые можно отсеять логической обработкой вида (4), (6). Кроме того, рассмотренные варианты обработки можно было бы осуществить с меньшими вычислительными затратами, оперируя лишь величинами числителя из (1), т.е. значениями $\xi(z, j)$. Нормировка в (1) позволяет фиксировать диапазон изменения $p(z_i, j)$ в виде отрезка $[0, 1]$ и интерпретировать величины $p(z_i, j)$ как вероятности. Более того, $p(z_i, j)$ можно сформировать непосредственно из значений $\rho(z_i, j)$

(или $\rho(z_i, Z(j))$). При этом класс (голос) элемента z_i будет определяться как минимум в строке матрицы: $j = \arg \min_{k \in J} \rho(z_i, k)$, веса классов как $\gamma(Z, j) = \sum_{i=1}^S \rho(z_i, j)$, а решение – в виде $k = \arg \max_j h(Z, j)$ или $k = \arg \min_j \gamma(Z, j)$.

Важным преимуществом вероятностной интерпретации величин $\rho(z_i, j)$ есть возможность перехода от анализа одиночных признаков к их сочетаниям, например, в виде пар $[z_i, z_k]$, $k \neq i$. Предполагая независимость z_i, z_k , можно вычислить $\rho([z_i, z_k], j) = \rho(z_i, j)\rho(z_k, j)$ как произведение. Аналогичным образом можно определить характеристики для сочетаний трех и более признаков. Учтывая, что $\rho(z_i, j)$ зависят от номера i , здесь напрямую неприменимы схемы Бернулли для оценки вероятностных характеристик пар, поэтому $\rho([z_i, z_k], j)$ нужно вычислять путем полного перебора. Общие выражения можно получить лишь для оценочных вероятностей, если выбрать характеристикой отдельного ХП, например, максимальное, минимальное или среднее значение в столбце $\{\rho(z_i, j)\}$. Сравнительный анализ матриц $\{\rho(z_i, j)\}$ для одиночных ХП и $\{\rho([z_i, z_k], j)\}$ для пар признаков показывает, что при сопоставлении со «своим» эталоном отношение m_2 / m_1 максимумов для пар всегда меньше, чем для одиночных ХП.

При практическом применении методов структурной классификации возникает необходимость рассмотрения ситуации отказа от классификации. Введение отказов в целом направлено на повышение достоверности за счет уменьшения вероятности ложной тревоги. Отказы реализуются путем проверки значимости решений локального (для элементов) и глобального (для классов объектов) уровней. Как правило, ситуации отказа можно характеризовать конечным числом управляемых параметров. Причиной отказа на локальном уровне есть, например, относительно низкое значение $m_1 = \max_j \rho(z_i, j)$. В таком случае по признаку z_i вычисление голоса не производится. На глобальном уровне принятие решения о классе может опираться на недостаточное значение числа элементов, проголосовавших за наиболее подходящий класс. Анализ отказов приводит к необходимости оперировать отдельными классами K_0 для ХП и $Z(0)$ для описаний. Характеристики K_0 и $Z(0)$ непосредственно не находят отражения в матрицах вероятностей и задаются параметрическим способом, т.к. в явном виде эталона для класса отказов нет.

Распределения или их характеристики, полученные на основе $\rho(z_i, j)$, могут быть положены в основу принятия решения о классе. В частности, можно использовать различия гистограмм или оценок математических ожиданий, дисперсий.

Разновидности моделей сопоставления

1. Модель «элемент-элемент» есть основой всех типов сопоставления с использованием голосования для классификации. В то же время она может быть применена и самостоятельно, например, для формирования множества соответствий $\theta(z_i, z_k^0)$ признаков «эталон – объект» с последующим поиском максимального числа либо доли соответствий среди эталонов [3].

2. Модель «множество-множество» основана на одноименном расстоянии или мере подобия и анализирует групповое сходство элементов описаний. Приведенные рассуждения для вычисления $\rho(z, j)$ на основе расстояния $\rho(z, Z(j))$, где в основе лежат классы описаний, могут быть непосредственно распространены и на уровень сопоставления $Z, Z(j)$ по типу множество-множество. Сопоставление реализуется вычислением меры подобия или метрики $\rho(Z, Z(j))$, затем в соответствии с (1) вычисляются вероятности $\rho(Z, j)$ отнесения Z к классу j . Построение $\rho(Z, Z(j))$ всегда связано с анализом матрицы расстояний $\{\rho(z_i, z_k)\}$, где $z_i \in Z, z_k \in Z(j)$. Как показали наши исследования, тип сопоставления множество-множество на основе подсчета эквивалентных элементов имеет лучшую помехозащищенность по сравнению с типом элемент-множество. Нормировку при реализации модели множество-множество лучше выполнять относительно значения s_j (число элементов $Z(j)$), т.к. значение s меняется в широком диапазоне из-за действия помех. Модель «сочетание элементов - сочетание элементов», например, в виде анализа пар ХП, занимает промежуточное положение между моделями «элемент-элемент» и «множество-множество».

3. Модель «элемент - класс элемента» оперирует множеством K классов ХП и конструирует описания Z из элементов K . Можно рассматривать как индивидуальные классы ХП для каждого из эталонов $Z(j)$, так и некоторую общую систему $K = \{K_1, \dots, K_q\}$ классов ХП. Основа построения этой модели – взаимодействие двух систем: классов ХП и классов описаний. Для каждой из систем может быть применена формула (1) для подсчета соответствующих коэффициентов. Одним из способов реализации модели есть применение кластеризации для построения классов ХП.

Результаты экспериментов и выводы

Основными факторами, определяющими построение структурной классификации, можно считать модели мер подобия при сопоставлении элементов и описаний в целом, а также способы логической обработки значений мер и критерии для определения класса описания.

В результате компьютерного моделирования вариантов СИМ для баз реальных видео-объектов установлено следующее [3, 6]. Применение схемы независимого голосования более устойчиво к действию структурных помех, чем интегрированное голосование. Применение способа интегрированного голосования, в свою очередь, более устойчиво к аддитивным помехам по сравнению с независимым голосованием. В целях построения универсального метода, в нужной мере устойчивого как к аддитивным, так и к пространственным помехам, следует изучить возможность комбинирования этих моделей.

Сравнительный анализ вероятностей $p(z_i, j)$ для отдельных элементов описания и $p([z_i, z_k], j)$ для пар признаков показывает, что как без помех, так и с учетом влияния помех достоверность классификации по совокупностям (сочетаниям) признаков выше, так как отношение m_2 / m_1 локального и глобального максимумов в строке матрицы для пар всегда меньше аналогичного отношения для одиночных ХП. Диапазон значений соотношения для базы аквариумных рыб с использованием детектора SIFT составил: для пар – 0,4-0,58, для одиночных ХП – 0,87-0,95.

Впервые показано, что путем выбора моделей для мер локального и глобального подобия, а также способов логической обработки и критерия принятия решения о классе с учетом параметрического определения можно добиться необходимых качеств системы структурной классификации видео-объектов с точки зрения уровня устойчивости к ад-

дитивным и пространственным помехам. Формализовано и изучено семейство мер подобия и связанных с ними критериев классификации.

Практически важным есть экспериментальное подтверждение теоретических принципов построения моделей классификации для реальных баз видеоданных при решении прикладных задач компьютерного зрения.

Перспективы исследования состоят в разработке комбинированного подхода при анализе голосов структурных признаков, где наряду с робастным выбором подобных элементов используется интегрированная информация о подмножестве наиболее близких элементов структурных описаний

Список литературы

1. Шапиро Л. Компьютерное зрение: пер. с англ. / Л. Шапиро, Дж. Стокман. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
2. Рассел С. Искусственный интеллект: современный подход; 2-е изд.; пер. с англ. / С. Рассел, П. Норвиг. – М.: Изд. дом «Вильямс», 2006. – 1408 с.
3. 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.
4. Гороховатский В.А. Модели локально-пространственного подобия структурных описаний визуальных объектов / В.А. Гороховатский // Системи управління, навігації та зв'язку: зб. наук. пр.– К.: ЦНДІ навігації і управління. – 2010. – Вип. 3(15). – С. 82-85.
5. Kim S. Object Recognition Using a Generalized Robust Invariant Feature and Gestalt's Law of Proximity and Similarity / S. Kim, K.-J. Yoon, I.-S. Kweon // Pattern Recognition. – 2008. – 41, No. 2. – P. 726-741.
6. Гороховатский В.А. Структурно-иерархические методы распознавания изображений при влиянии пространственных помех / В.А. Гороховатский // Бионика интеллекта. – 2010. – №3(74). – С. 58-61.

Поступила в редколлегию 10.02.2011

Рецензент: д-р физ.-мат. наук, ст. научн. сотр. О.В. Сытник, Институт радиофизики и электроники НАН Украины, Харьков.

КРИТЕРІЇ ТА МОДЕЛІ СТРУКТУРНОЇ КЛАСИФІКАЦІЇ ІЗ ЗАСТОСУВАННЯМ ПРИНЦИПУ ГОЛОСУВАННЯ

В.О. Гороховатський, Т.В. Полякова

Обговорюються шляхи побудови моделей та критеріїв для класифікації структурних описів відео-об'єктів із використанням голосування. Шляхом вибору моделей для структурних мір схожості з урахуванням їх параметричного визначення пропонується побудова системи класифікації з високим рівнем стійкості до адитивних і просторових перешкод. Аналізуються результати експериментів.

Ключові слова: класифікація зображень, структурно-ієрархічні методи, характерні ознаки, структурний опис, принцип голосування, критерій класифікації, модель зіставлення, достовірність класифікації.

CRITERIA AND MODELS OF STRUCTURAL CLASSIFICATION USING VOTING PRINCIPLE

V.O. Gorohovatsky, T.V. Polyakova

The ways of constructing models and criteria for classification of video objects structural descriptions based on voting are discussed. The construction system of classification with high immunity to additive and spatial interference is offered by selection models for structural similarity measures, given their parametric definition. Results of experiments are analyzed.

Keywords: image classification, structurally-hierarchical methods, characteristic features, structural description, voting principle, criterion of classification, model of comparison, reliability of classification.