



В.А. Гороховатский¹, Т.В. Полякова²

¹Харьковский институт банковского дела Университета банковского дела НБУ,
г. Харьков, Украина, e-mail:gorohovatsky-v@rambler.ru

²Харьковский национальный автомобильно-дорожный университет,
г. Харьков, Украина, e-mail:assiada@list.ru

ОПТИМАЛЬНЫЕ МЕТОДЫ СОПОСТАВЛЕНИЯ СТРУКТУРНЫХ ОПИСАНИЙ ВИДЕООБЪЕКТОВ

Обсуждаются вопросы применения оптимальных подходов при сопоставлении структурных описаний видео-объектов на примере венгерского метода для поиска максимального паросочетания в двудольном взвешенном графе. Результаты экспериментов с использованием признаков аффинных инвариантов, построенных на множестве координат характерных признаков изображения, подтверждают эффективность предложенных методов в плане более высокой достоверности распознавания по сравнению с традиционными методами на основе голосования.

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, ХАРАКТЕРНЫЕ ПРИЗНАКИ, СОПОСТАВЛЕНИЕ СТРУКТУРНЫХ ОПИСАНИЙ, ПАРСОЧЕТАНИЕ, ВЕНГЕРСКИЙ МЕТОД, ВЗАИМНО-ОДНОЗНАЧНОЕ СООТВЕТСТВИЕ

Введение

В настоящее время для решения интеллектуальных задач компьютерного зрения, связанных с распознаванием объектов в условиях действия фона и помех, эффективны структурные технологии, основанные на анализе особенностей изображения в отдельных точках. Оценка свойств описания видео-объекта в виде конечного числа характерных признаков (ХП) позволяет за приемлемое время качественно разделить информацию об объекте и о фоновых образованиях [1-3]. Основой для формирования ХП служит информация, содержащаяся в некотором смысле «значимых» фрагментах. Распознавание на основе структурных описаний в виде множества ХП имеет следующие преимущества: достаточно простое в вычислительном плане формирование признаков, сжатое признаковое пространство с возможностью управления его размерностью в зависимости от задачи, устойчивость к влиянию фоновых искажений и ложных объектов.

В пространстве ХП распознавание (классификация) видео-объектов сводится к построению мер подобия конечных множеств на основе соответствий элементов. Эффективный способ реализации подобия – голосование ХП [2]. Наряду с устойчивостью принятия решения при неполном или избыточном представлении описаний голосование обладает такими важными достоинствами, как высокая вероятность правильного распознавания и простота построения, что делает его конкурентоспособным среди других подходов.

Особый интерес при сопоставлении конечных множеств признаков в задачах искусственного интеллекта представляет применение оптимальных подходов, в основе которых лежит оптимизация значения некоторого критерия (в данном случае величины подобия описаний) путем комбинатор-

ного перебора на конечном множестве вариантов [4]. Одним из эффективных способов структурного анализа данных, связанных с вычислением оптимального сходства конечных множеств (описаний распознаваемых объектов), является венгерский метод (ВМ), названный комбинаторной оптимизацией ограниченных множеств. Его можно рассматривать как решение задачи о назначениях для наиболее общего случая поиска максимального паросочетания в двудольном взвешенном графе [5-7]. Суть венгерского метода состоит в последовательном формировании оптимального решения за конечное число шагов. Решение задачи о назначениях решает проблему установления степени сходства структурных описаний объектов в условиях несоответствия порядка их следования или нарушения целостности описания из-за помех. Одним из достоинств применения ВМ при сопоставлении описаний является возможность оценки близости промежуточной величины сходства к оптимальному варианту решения на каждой из итераций. Это позволяет контролировать процесс и при необходимости (в целях сокращения временных затрат) прекращать его при достижении заданной точности. Такое свойство особенно важно для задач большой размерности, т.к. число векторов ХП в описаниях достигает двух-трех сотен.

Принцип оптимального сопоставления структурных описаний по сравнению с известными решениями (SIFT, SURF [1]), основанными на голосовании ХП, позволяет реализовать схему построения однозначных соответствий ХП из сравниваемых описаний, что в целом улучшает достоверность распознавания.

Цель исследования – анализ эффективности применения оптимальных подходов на примере венгерского метода для сопоставления описаний в виде множеств ХП.

Задачи работы состоят в модернизации ВМ для сопоставления структурных описаний, сравнительной оценке сложности вычислительных затрат с другими методами, экспериментальных исследованиях оптимальных методов для конкретных систем признаков и баз видеоинформации.

1. Формализация метода оптимального сопоставления описаний

Представим описание U визуального объекта в виде конечного множества $U = \{u_i\}_{i=1}^m$ элементов u_i , где m – количество ХП в описании, u_i – ХП, который может представлять как дескриптор (значение признака), так и геометрическую информацию (координаты ХП, значения аффинных инвариантов на основе этих координат и т.д.) [3].

Критерием при сопоставлении элементов является значение некоторой метрики $\rho(u_i, u_j)$, оценивающей различия u_i, u_j , взятых из разных описаний $u_i \in U_1, u_j \in U_2$. Методы с использованием голосования определяют число (сумму) голосов элементов, исходя из величин $\rho(u_i, u_j)$ путем оптимизации на конечном множестве или через оценку с помощью некоторого порога. Оптимальные методы на основе $\rho(u_i, u_j)$ формируют некоторое оптимальное решение, например, минимизирующее сумму расстояний между описаниями U_1, U_2 с учетом однозначности соответствий элементов сравниваемых множеств.

Применим ВМ для оптимального установления величины соответствия между двумя структурными описаниями с учетом действия пространственных помех. Результатом выполнения ВМ является формирование максимального паросочетания для элементов двух множеств с минимизацией общей стоимости (весов), которую можно оценить в виде суммы расстояний между парами ХП из сравниваемых описаний. Сопоставление U_1, U_2 на основе ВМ формально сводится к решению оптимизационной задачи:

$$R(x) = \sum_{i=1}^m \sum_{j=1}^n \rho(u_i, u_j) x_{ij} \rightarrow \min, \quad (1)$$

где $u_i \in U_1, u_j \in U_2, m$ и n – количества точек в описаниях U_1, U_2 ; $x_{ij} \in \{0, 1\}$ – бинарный признак, отражающий соответствие i -го и j -го элементов из U_1, U_2 . Решение задачи (1) при ограничении на однозначность соответствия признаков из сопоставляемых множеств

$$\begin{cases} \sum_{j=1}^n x_{ij} = 1 \quad \forall i = \overline{1, m}, \\ \sum_{i=1}^m x_{ij} = 1 \quad \forall j = \overline{1, n}, \end{cases} \quad (2)$$

минимизирует расстояние между U_1, U_2 . Для прикладных задач, где необходима максимизация целевой функции в (1), переход к задаче (1) осу-

ществляют путем вычитания значений $\rho(u_i, u_j)$ из максимального значения метрики.

Как оптимальный, так и традиционный (основанный на голосовании) подходы используют в качестве базовой информации матрицу расстояний

$$P = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots \\ \rho_{m1} & \rho_{m2} & \dots & \rho_{mn} \end{bmatrix}$$

между всевозможными парами ХП $u_i \in U_1, u_j \in U_2$.

Реализация принципа однозначного голосования элементов u_i связана с анализом отдельных строк матрицы P на предмет формирования голоса, равного 0 или 1. По полученному набору соответствий (голосов) можно вычислить значение критерия подобия двух описаний в виде суммы $S_1 = \sum_i L(i)$, где $L(i)$ – предикат, равный 1, если голос отдается (соответствие установлено).

Реализация ВМ предполагает целенаправленный выбор такой последовательности из $\{\rho_{ij}\}$, сумма значений которой $S_2 = \sum \rho_{ij}$ будет минимальна на множестве возможных соответствий элементов описаний, и при этом из каждой строки и столбца матрицы P будет выбран только один элемент, что в результате обеспечивает выбор оптимального по критерию S_2 однозначного соответствия множеств элементов описаний (паросочетание).

Реализация ВМ осуществляется путем целенаправленного эквивалентного преобразования матрицы P с последовательным анализом строк и столбцов для получения матрицы с неотрицательными элементами и системой m независимых нулей, из которых никакие два не принадлежат одной и той же строке или одному и тому же столбцу. При достижении ситуации с m независимыми нулями проблема выбора считается решенной, оптимальный вариант назначений определяется позициями независимых нулей в преобразованной матрице [5]. Пример формирования паросочетания показан на рис. 1.

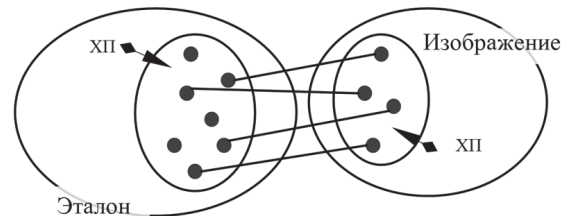


Рис. 1. Паросочетания для множеств ХП

Известно, что задача поиска паросочетаний с использованием ВМ решается за время, пропорциональное величине d^3 , где d – это размер сравниваемых описаний. Видим, что требуемые затраты вычислений существенно возрастают с увеличением d . В то же время для метода с голосованием затраты прямо пропорциональны d , что не так критично к росту объема описаний.

При применении ВМ для сопоставления описаний U_1, U_2 их размерности m и n в общем случае могут отличаться. В этом случае в традиционном ВМ размер матрицы P увеличивают до максимального среди m и n , а возникшие элементы заполняют нулями [5-7]. Если исходная матрица не является квадратной, то дополнительно вводят необходимое количество строк (или столбцов), а их элементам присваивают значения, определяемые условиями задачи. Для задачи определения подобия описаний полученным за счет расширения элементам необходимо присвоить некоторое максимальное значение среди возможных значений $\rho(u_i, u_j)$. Заметим, что для нашей задачи нецелесообразно усекать матрицу весов до квадратной традиционным способом в виде $\min(m, n)$, т.к. при этом может быть потеряна ценная информация об видео-объекте. С другой стороны, дальнейшее усложнение обработки за счет применения ВМ требует минимизации размеров матрицы на этапе предварительной обработки. Это достигается посредством применения специальных методов сжатия структурных описаний [2]. Другим способом снижения затрат есть отбор в некотором смысле «значимых» строк и столбцов, например, содержащих большие по значению расстояния.

2. Геометрические признаки аффинных инвариантов

Особый интерес представляет применение оптимальных методов для геометрических признаков, построенных на основе координат ХП и отражающих пространственное расположение ХП объекта. Любое подмножество трех неколлинеарных точек объекта определяет аффинный базис, который задает на объекте систему координат. После выбора системы координат ХП можно представить с помощью аффинных инвариантов (АИ). После воздействия аффинного преобразования на объект значения АИ не изменяются [1]. Существуют различные способы пространственной группировки АИ для сопоставления описаний. В работе [3] для отдельного ХП формируется подмножество АИ путем рассмотрения всевозможных базисов (разные системы аффинных координат). Другим способом является формирование подмножества АИ для всех ХП в одной и той же системе координат, т.е. для отдельного базиса. Такой способ представления имеет вид значительного числа подмножеств (количество базисов) с небольшим количеством элементов. Данное представление удобно для применения ВМ из-за небольших по размеру матриц P . Кроме того, оно дает возможность управлять количеством подмножеств (базисов) и принимать решение по неполному описанию.

Рассмотрим множество АИ, отображающих геометрические свойства описания объекта и представленных в виде конечного множества числовых

троек $\{\alpha_q\}_{q=1}^r$ [3], где r – число АИ. Расстояние между признаками α_q и α_p эталона и изображения для матрицы P можно вычислить как

$$\rho_{qp} = \sqrt{\sum_{i=1}^3 (\alpha_{qi} - \alpha_{pi})^2}.$$

Взаимно-однозначное соответствие элементов описаний может быть установлено путем решения задачи поиска взвешенного паросочетания. Решение о принадлежности объекта конкретному эталону принимается на основе максимального числа установленных соответствий между подмножествами АИ или на основе минимума величины S_2 .

3. Результаты экспериментов

В качестве критерия, характеризующего достоверность распознавания в конечной базе данных, рассмотрим величину $\theta = h_1 / h_2$, где h_2 – максимум гистограммы голосов, h_1 – ближайший к нему максимум. Значение критерия θ показывает, насколько уверенно осуществляется принятие решения на основе полученного максимума числа голосов, отданных за конкретный эталон. Чем меньше значение θ , тем в большей степени глобальное решение значимо по отношению к локальному, соответствующему наиболее близкому из остальных (возможно, ложных) классов.

Для базы объектов видео-данных Coil-20 [1] вычислено множество АИ $\{\alpha_q\}$. Эксперимент для традиционного подхода на основе голосования ХП состоял в том, что на вход системы распознавания поступало описание одного из эталонов, которое последовательно сравнивалось со всеми описаниями базы, в результате принималось решение в соответствии с максимумом нормированного количества голосов [2]. Для предложенного оптимального метода сопоставление осуществлялось по принципу «множество-множество» с достижением минимума в выражении (1) при условии однозначного соответствия элементов множеств. Для базы Coil-20 с применением структурных описаний в виде множеств АИ достигнуто безошибочное распознавание.

Сравнительные эксперименты показали, что в традиционном подходе величина θ для правильно найденного класса относительно других эталонов достигает значения $\theta = 0,8$, в то время как для оптимального метода на основе ВМ максимальное из значений $\theta = 0,16$. Это непосредственно доказывает, что достоверность распознавания для оптимального метода существенно лучше.

В то же время отметим, что для оптимального метода затраты времени на решение задачи распознавания несколько больше, чем для традиционного метода. При этом временные затраты значительно увеличиваются с увеличением размерности сравниваемых описаний, а применение аппарата АИ, как известно, по сравнению с обычным пред-

ставлением в виде множества ХП еще более усиливает временные различия. Например, для $m=6$ с использованием пространства АИ временные затраты оптимального метода при распознавании оказались в 1,5 раза выше, а при $m=13$ – в десятки раз больше. Отметим, что эти затраты могут регулироваться за счет процедуры предварительной обработки структурных описаний для конкретно решаемой прикладной задачи.

Выводы

Оптимальные методы для сопоставления структурных описаний объектов за счет более тщательного отбора соответствий обеспечивают лучшую достоверность распознавания по сравнению с традиционными подходами на основе голосования. Проведен анализ и исследование оптимального метода сопоставления на основе ВМ для описаний, представленных множеством аффинных инвариантов.

В целом полученные характеристики для оптимального метода могут быть оценены как предельно достижимые, в то время как традиционный подход с голосованием отдельных элементов описания можно считать некоторой практической аппроксимацией. Оптимальные методы целесообразно применять для задач, требующих наиболее тщательно го анализа соответствия множеств описаний.

Впервые показано, что применение оптимальных методов определения максимального паросочетания в двудольном графе для задачи сопоставления структурных описаний видео-объектов повышает достоверность распознавания за счет однозначного и более точного учета подобия подмножеств элементов описаний. На примере геометрических признаков изображений получено экспериментальное подтверждение эффективности предложенных методов.

Практически важным является получение предпочтительных характеристик распознавания по сравнению с известными методами, что говорит о целесообразности развития и применения оптимальных методов сопоставления в задачах компьютерного зрения.

Перспективы исследования состоят в разработке подходов к построению мер подобия на основе структурного анализа подмножеств значений аффинных инвариантов, сгруппированных для отдельных базисов.

Список литературы: 1. Шапиро, Л. Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с. 2. Gorokhovatskiy V.A. Compression of Descriptions in the Structural Image Recognition / V.A. Gorokhovatskiy // Telecommunications and Radio Engineering. – 2011, Vol. 70, No 15. – P. 1363-1371. 3. Гороховатский, В.А. Модели комплексированных мер подобия структурных описаний изображений [Текст] / В.А. Гороховатский, Т.В. Полякова, Е.П. Путятин // Реєстрація, зберігання і обробка даних. – 2011. – Т. 13, № 1. – С. 21–28. 4. Рассел, С. Искусственный интеллект: современный подход; 2-е изд. [Текст] / С. Рассел, П. Норвиг; пер. с англ. – М.: Изд. дом «Вильямс», 2006. – 1408 с. 5. Гольштейн, Е.Г. Задачи линейного программирования транспортного типа [Текст] / Е.Г. Гольштейн, Д.Б. Юдин, – М.: Наука, 1969. – 382 с. 6. Соськин, Д.М. Адаптивный алгоритм распределения заказов на обслуживание автомобилей такси [Текст] / Д.М. Соськин // Известия Томского политехнического университета. – 2009. – Т. 315, № 5. – С.65-69. 7. Пападимитриу, Х. Комбинаторная оптимизация. Алгоритмы и сложность [Текст] / Х. Пападимитриу, К. Стайглиц // М.: Мир, 1985. – 512 с.

Поступила в редколлегию 19.05.2011

УДК 004.932.2:004.93'1

Оптимальні методи зіставлення структурних описів відеооб'єктів / В.О. Гороховатський, Т.В. Полякова // Біоніка інтелекту: наук.-техн. журнал. – 2011. – № 3 (77). – С. 85-88.

Обговорюються питання застосування оптимальних підходів при зіставленні структурних описів відеооб'єктів на прикладі угорського методу для пошуку максимального паросполучення у дводольному зваженому графі. Результати експериментів з використанням ознак аффінних інваріантів, побудованих на множині координат характерних ознак зображення, підтверджують ефективність запропонованих методів в плані більш високої достовірності розпізнавання в порівнянні з традиційними методами на основі голосування.

Бібліогр.: 7 найм.

UDC 004.932.2:004.93'1

Optimal methods of structural descriptions of video objects comparison / V.O. Gorokhovatsky, T.V. Polyakova // Bionics of Intelligense: Sci. Mag. – 2011. – № 3 (77). – P. 85-88.

The application of optimal approaches for structural descriptions of video-objects comparison are discussed, e.g. Hungarian method for finding maximum matching in the bipartite weighted graph. Experimental results based on affine invariant features which are built on coordinate set of image characteristic features confirm efficiency of proposed methods at higher recognition rate in comparison with traditional voting methods.

Ref.: 7 items.