

УДК 681.5

Х. ВЕРН, В. И. ВАСИЛЬЕВ, К. В. САФРОНОВ, И. М. ЧУШЕНКОВ

ИЕРАРХИЧЕСКИЙ ИТЕРАЦИОННЫЙ МЕТОД РАСПОЗНАВАНИЯ ОБРАЗОВ ПРИ АНАЛИЗЕ МНОГОМЕРНЫХ ДАННЫХ

Задача распознавания образов является одной из наиболее известных задач компьютерного зрения, но, несмотря на довольно высокий уровень исследований в этой области, на данный момент не существует универсального алгоритма распознавания, который был бы способен эффективно решить некоторые практические задачи, обладающие определенной спецификой. В данной статье предлагается новый иерархический итерационный метод распознавания образов, который базируется на методах сопоставления с образцом и использует алгоритмы оптимизации. Результаты экспериментов, проведенных в течение фазы ввода в эксплуатацию систем распознавания, доказали эффективность представленного метода. Разработанный метод был применен в различных проектах, а именно для решения задач регистрации медицинских томограмм, распознавания текста в промышленных системах, а также при визуальном сервоуправлении. Успешно проведенные в различных областях эксперименты демонстрируют универсальность разработанного метода. *Распознавание образов; иерархическая оптимизация; многомерные данные*

Задача распознавания образов является одной из наиболее известных задач, направленных на создание систем компьютерного зрения, которые находят широкое применение в различных областях деятельности человека (например, в биометрии, при обработке речи, обработке текста, анализе изображений и т.п.). Вместе с тем, несмотря на довольно высокий уровень исследований в этой области, построение систем распознавания образов сегодня все еще остается сложной теоретической и технической проблемой. На практике всегда возникают задачи, обладающие определенной спецификой, которые не могут быть эффективно решены с помощью существующих алгоритмов распознавания, поскольку на сегодняшний день универсального алгоритма решения таких задач не существует.

1. СТАНДАРТНЫЕ ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ

В литературе, посвященной данной проблеме распознавания образов, выделяется 2 класса альтернативных методов [1]:

1. **Методы, основанные на теории принятия решений.** Этот класс методов используется для распознавания образов (объектов) ко-

личественные дескрипторы, такие, как длина, площадь, текстура и т.п.

К методам, основанным на теории принятия решений, относятся [1–5]:

- **методы сопоставления с образцом.** Данные методы представляют каждый класс объектов в виде вектора признаков образа, который принимается в качестве прототипа этого класса. При распознавании незнакомого образа объект относится к тому классу, прототип которого оказывается «ближайшим» при использовании заранее заданной дискриминантной функции. Наиболее известные методы, относящиеся к данному классу — *метод классификации по минимуму расстояния* и *метод корреляционного сопоставления* [1];

- **статистические методы** основаны на построении функции плотности распределения вероятностей образов для каждого класса объектов и классификации появления каждого из классов по какому-либо правилу (распределение можно построить на этапе обучения, используя обучающую выборку). Наиболее часто используемыми классификаторами являются *классификатор Байеса*, *минимаксный классификатор* и *классификатор Неймана–Пирсона* [2].

- **методы, использующие нейронные сети,** основаны на построении нейронной сети, на входы (рецепторы) которой подается век-

тор признаков распознаваемого образа (объекта), а на выходах формируются сигналы эффекторов, соответствующих заданным образам. Тот из эффекторов, который выдает сигнал максимальной амплитуды, отображает класс, к которому относится распознаваемый объект. Дискриминантные функции такой нейронной сети строятся непосредственно в ходе ее обучения [3].

2. Структурные методы. Данный класс методов использует качественные дескрипторы (признаки), например, реляционные дескрипторы. Распознавание образов происходит на основе анализа присущих им структурных связей, которые, как правило, не учитываются при использовании методов, основанных на теории принятия решений.

В классе структурных методов распознавания образов можно выделить следующие группы методов [1–5]:

- **методы сопоставления номеров фигур или строк символов** основаны на сравнении границ областей, описанных в терминах номеров фигур или строк символов;

- **методы синтаксического распознавания строк символов или деревьев** основаны на задании множества производных элементов, из которых состоят образы, набора правил в форме грамматики, управляющих объединением этих элементов, а также и построением распознавателя (автомата), структура которого определяется набором правил грамматики.

3. Алгебраические методы. К данному классу методов распознавания образов относятся методы, основанные на алгебраической теории распознавания, решающие задачу выделения локально-симметричных компонентов или примитивов из изображений при использовании преобразований, относительно которых искомый примитив инвариантен. Одним из таких методов является **резонансный метод выделения геометрических примитивов** [2].

2. ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ

Как показывает анализ, существующие методы распознавания, упомянутые выше, не позволяют эффективно решать ряд распространенных задач, возникающих на практике. К таким задачам относятся:

- регистрация (совмещение) медицинских изображений, которая заключается в анализе томограмм какого-либо органа и поиске за приемлемое время аналогичных томограмм в атласе (базе данных томограмм) для

последующей постановки диагноза или совмещении томограмм другого типа для этого же органа (например, компьютерной и магниторезонансной томограмм);

- распознавание текстов в промышленных системах, которое осложняется сильными помехами, дефектами и искажениями символов;

- распознавание отдельных объектов и определение их положения на множестве перекрывающихся друг друга объектов, относящихся к разным классам.

Специфика данных задач состоит в следующем:

- необходимо использовать алгоритм, не нуждающийся в постоянном обучении и модификации даже в случае изменения набора распознаваемых объектов;

- необходимо распознавать не только объекты, но и определять их положение в пространстве, ориентацию и взаимное расположение.

Часто требования к алгоритмам распознавания дополняются условиями их конкретного применения, например, в промышленности. Поскольку при построении систем распознавания приходится учитывать специфику обрабатываемых данных и объектов, то при изменении этих данных либо распознаваемых объектов приходится решать задачу распознавания заново, что приводит к определенным осложнениям.

Целью исследований, представленных в настоящей статье, является разработка такого метода распознавания образов, основанных на анализе многомерных данных, который бы позволил:

- минимизировать (по времени и сложности) процедуру построения набора признаков;

- минимизировать (по времени и сложности) процедуру обучения алгоритма распознавания;

- распознавать различные образы, их положение и взаимное расположение;

- распознавать образы, подвергшиеся искажениям (таким, как изменение масштаба, яркостные искажения, аффинные преобразования и т. п.).

3. МЕТОДИКА ИССЛЕДОВАНИЙ

Прежде всего отметим, что рассмотренные в разд. 1 методы распознавания образов неприменимы для решения поставленной задачи в силу следующих обстоятельств:

а) *методы сопоставления с образцом* чувствительны к искажениям распознаваемых образов и связаны с трудоемкими вычислениями, однако, легко обучаемы;

б) *статистические методы* встречаются с рядом трудностей при решении практических задач, так как выборки образов из каждого класса часто оказываются недостаточно представительными, а функции плотности распределения вероятностей сложны для построения;

в) *нейронные сети* имеют сложную многослойную структуру и требуют длительного обучения с использованием большого числа примеров;

г) *структурные методы* чувствительны к искажениям распознаваемых образов и требуют сложной процедуры построения набора признаков;

д) *алгебраические методы* пока недостаточно исследованы и изучены.

Предлагаемый авторами альтернативный иерархический итерационный метод распознавания образов базируется на подходе, положенном в основу методов сопоставления с образцом, но лишен присущих им недостатков вследствие использования некоторых особенностей, которые будут описаны ниже.

3.1. Математическая модель

Допустим, что требуется отнести каждый из M распознаваемых образов Y_j , $j = 0, 1, \dots, M$ к одному из N классов образов X_1, X_2, \dots, X_N .

Описание образа Y_j можно представить с помощью иерархической модели, включающей в себя группы признаков образа (структурные, вероятностные и т.д.) A, B, C, D, \dots, Z . Например, это могут быть: $A = \{x, y, z, \alpha, \beta, \gamma\}$, где x, y, z — перемещение по осям, α, β, γ — вращение вокруг осей; $B = \{s, q, \dots\}$, где s — масштабирование, q — деформация образа и т. п.

В целом, модель i -го образа (прототипа i -го класса) X_i представляет собой совокупность наборов групп признаков:

$$X_i = \{ \{A_1^i, B_1^i, C_1^i, \dots\}, \{A_2^i, B_2^i, C_2^i, \dots\}, \dots, \{A_k^i, B_k^i, C_k^i, \dots\} \}. \quad (1)$$

Тогда формально задачу распознавания образов можно представить следующим образом.

На основе результата функции F_T распознаваемому образу Y_j ставится в соответствие

модель X_k , имеющая наибольшую схожесть с данным образом:

$$X_k = F_T(f(Y_j, X_i), f(Y_j, X_l)), \quad (2)$$

$$i, l = 0, \dots, M; \quad j = 0, \dots, N; \quad i \neq l,$$

где $f(Y_j, X_i)$ — дискриминантная функция, отображающая степень схожести распознаваемого образа Y_j и модели образа X_i и которую можно представить следующим образом:

$$f(Y_j, X_i) = \{d_0(Y_j, X_i), d_1(Y_j, X_i), \dots, d_m(Y_j, X_i)\}, \quad (3)$$

где d_0, d_1, \dots, d_m — дискриминантные функции более низких порядков, использующие только определенные (ограниченные) группы признаков образа Y_j для его классификации (например, только структурные признаки или только вероятностные). Примеры таких дискриминантных функций будут приведены ниже (5, 6, 7).

Для описания предлагаемого метода распознавания образов необходимо ввести еще одну (обратную) функцию:

$$Y_j = F_P(f(Y_j, X_k), f(Y_t, X_k)), \quad (4)$$

$$j, t = 0, \dots, M; \quad j = 0, \dots, M; \quad j \neq t.$$

Функция F_P ставит в соответствие модели X_k образ Y_j , имеющий наибольшую схожесть с данной моделью.

3.2. Описание метода

Идея разрабатываемого метода заключается в поиске на множестве многомерных исходных данных определенных образов, заданных специальными моделями (см. рис. 1). Например, при решении задачи распознавания символов, исходными данными является двухмерное изображение, а специальными моделями являются шаблоны каждого из символов.

Преимуществами данного подхода являются:

- простота обучения алгоритма распознавания (обучение заключается в выборе функций качества и их взаимосвязи);
- простота процедуры построения набора признаков (в случае распознавания текста, набором признаков являются наборы пикселей шаблона символа).

Естественно, что поиск на множестве многомерных исходных данных связан с высокими вычислительными затратами.

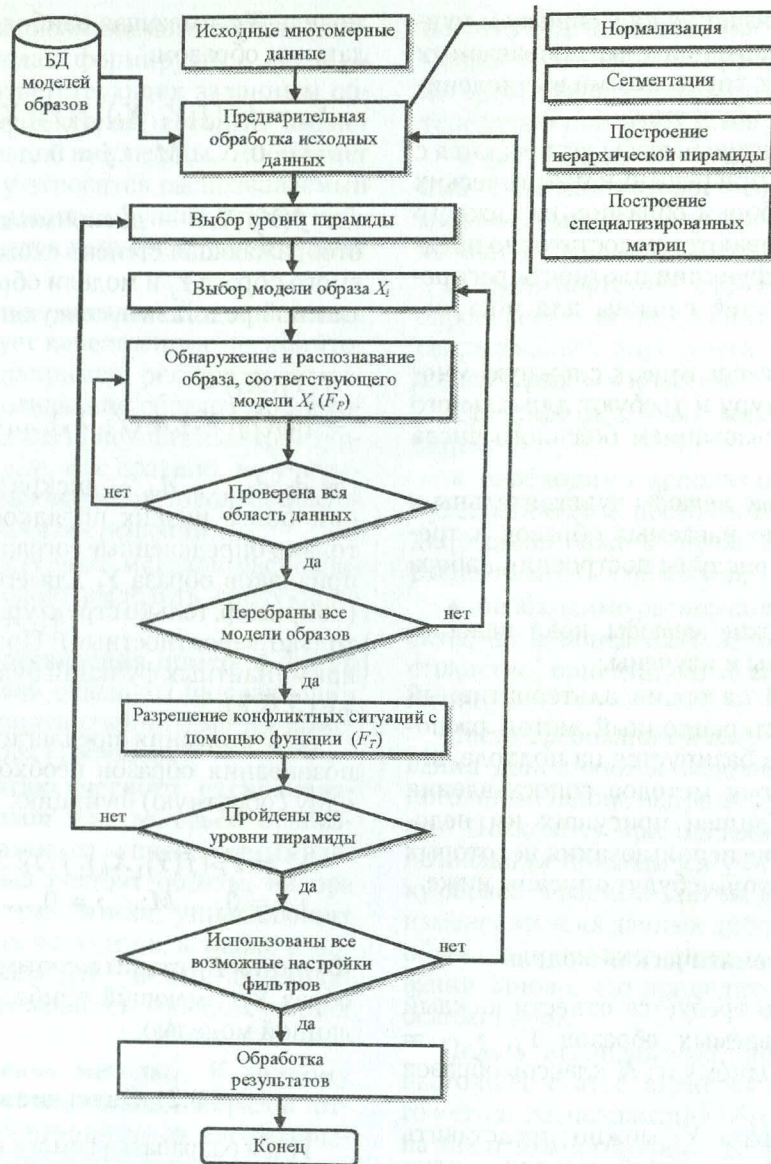


Рис. 1. Блок-схема алгоритма распознавания образов

Естественно, что поиск на множестве многомерных исходных данных связан с высокими вычислительными затратами, которые можно минимизировать, используя алгоритмы оптимизации (например, метод Пауэлла, метод Нелдера–Мида, симплекс-метод и др.) [6]. Такой алгоритм оптимизирует функцию качества, которая описывает степень соответствия модели распознаваемого образу. В качестве алгоритма оптимизации, используемого в предлагаемом алгоритме распознавания, был выбран метод Пауэлла. Критерий качества (дискриминантная функция), на основе значения которого выбиралось положение шаблона символа X_i относительно изображения Y_j для следующей ите-

рации, рассчитывается по следующим формулам:

$$f(Y_j, X_i) = \{d_1(Y_j, X_i), d_2(Y_j, X_i)\} \quad (5)$$

$$d_1 = \sum_{i=0}^L l_i, \quad (6)$$

где L — количество точек шаблона, l_i — расстояние от i -й точки шаблона до ближайшей точки распознаваемого символа;

$$d_2 = \frac{L_e}{L_t} \cdot \frac{L_e}{L_c}, \quad (7)$$

где L_e — количество точек шаблона, совпадающих с точками распознаваемого символа,

L_t — общее количество точек шаблона, L_c — количество точек области изображения, ограниченного наложенным шаблоном.

В целях увеличения точности, а также скорости распознавания образов были предложены следующие усовершенствования процедуры поиска:

- *использование расширенного алгоритма оптимизации Пауэлла.* Данный алгоритм оптимизирует функцию качества в три этапа (глобальный поиск, локальный поиск и точный поиск);

- *иерархическая оптимизация.* При этом исходные данные представляются в виде пирамиды, каждый уровень которой содержит сжатые каким-либо образом данные (например, с помощью фильтра Гаусса при распознавании символов) и осуществляет распознавание образов сначала на более «грубом» уровне представления, постепенно переходя к более «точному»;

- *применение специальных матриц* (матриц расстояний и указателей) при расчете критерия качества, основанного на определении расстояния от модели до распознаваемого образа;

- *использование деформируемых моделей.* В процессе сопоставления моделей и распознаваемых объектов, при условии наличия искажения, модели деформируются различными способами;

- *использование нескольких критериев качества.* При разработке алгоритма распознавания необходимо выбрать такие критерии качества, которые имели бы разную природу, что позволило бы с большей вероятностью найти оптимум;

- *иерархическая предобработка исходных данных.* Данный метод основан на том предположении, что различные образы могут быть верно распознаны при предварительной обработке исходных данных фильтрами с различными параметрами;

- *применение алгоритма взвешивания составных частей модели распознаваемого образа.* При распознавании перекрытых и искаженных образов считается целесообразным назначить каждой части модели определенный коэффициент, который бы учитывался при расчете критерия качества сопоставления.

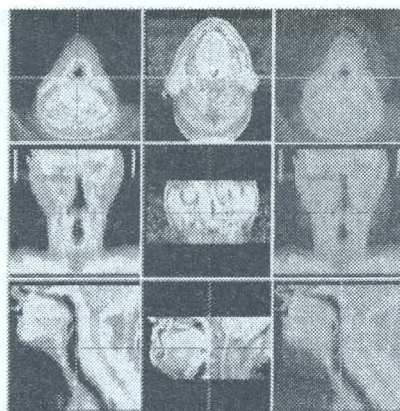
4. АНАЛИЗ РЕЗУЛЬТАТОВ

Предложенный алгоритм распознавания образов показал высокую эффективность при решении следующих практических задач:

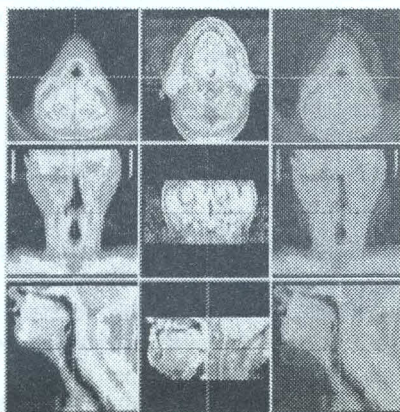
- *регистрация (совмещение) компьютерных и магниторезонансных томограмм* [6, 7] (рис. 2);

- *распознавание текстов в промышленных системах* [8] (рис. 3).

При решении последней задачи для оценки эффективности предлагаемого метода были проведены эксперименты с использованием существующих программных систем распознавания образов (SimpleOCR, GOCR, TopOCR, Readiris Pro, ScreenOCR, OpenOCR, Finereader, TOCR), которые охватывают весь спектр методов распознавания образов, разработанных на данный момент. Эксперименты показали, что только предлагаемый метод способен распознавать тексты в промышленных системах с вероятностью 99% [8].



а



б

Рис. 2. Пример регистрации томограмм: положение томограмм до (а) и после (б) регистрации



Рис. 3. Пример распознаваемых в промышленных системах текстов

Другая задача, исследуемая в данный момент с использованием иерархического итерационного метода распознавания образов, заключается в распознавании отдельных объектов различных классов и их положения в трехмерном пространстве исходных данных, полученных с помощью стереокамеры и камеры, произведенной по PMD-технологии.

ЗАКЛЮЧЕНИЕ

В данной статье представлен предложенный авторами иерархический итерационный метод распознавания образов на множестве многомерных исходных данных, основанный на алгоритмах сопоставления. Разработанный алгоритм обладает следующими преимуществами:

- способность распознавать не только образ, но и его положение в пространстве;
- отсутствие необходимости дообучения или модификации алгоритма при изменении исходных данных или распознаваемых образов;
- инвариантность к распознаванию искаженного образа;
- универсальность метода, заключающаяся в возможности решения задачи распознаванию образов на множестве двухмерных, трехмерных и многомерных данных.

При решении ряда практических задач, рассмотренных в данной статье, было установлено, что иерархический итерационный метод распознавания позволяет эффективно решать задачи распознавания образов в различных приложениях.

СПИСОК ЛИТЕРАТУРЫ

1. **Гонсалес, Р.** Мир цифровой обработки / Р. Гонсалес, Р. Вудс. М.: Техносфера, 2005. 1072 с.
2. **Сойфер, В. А.** Методы компьютерной обработки изображений / В. А. Сойфер. М.: Физматлит, 2003. 784 с.
3. **Форсайт, Д. А.** Компьютерное зрение. Современный подход: пер. с англ. / Д. А. Форсайт, Ж. Понс. М.: Вильямс, 2004. 928 с.
4. **Theodoridis, S.** Pattern Recognition / S. Theodoridis, K. Koutroubas. Elsevier, 2003. 689 с.
5. **Горелик, А. Л.** Методы распознавания / А. Л. Горелик, В. А. Скрипкин. М.: Высш. шк., 1984. 208 с.
6. **Safronov, K.** Automated registration of medical head volume images using gradient based method / K. Safronov, J. Graf, I. Tchouchenkov, H. Woern // Proc. of the 7th Int. Workshop on Computer Science and Information Technologies CSIT'2005. Ufa, Russia, 2005. Vol. 1. P. 174–179.
7. **Safronov, K.** Combined medical image registration using both mutual and gradient informaiton / K. Safronov, I. Tchouchenkov, H. Woern // Proc. of the 8th Int. Workshop on Computer Science and Information Technologies CSIT'2006. Karlsruhe, Germany, 2006. Vol. 1. P. 7–10.
8. **Safronov, K.** Optical character recognition using optimisation algorithms / K. Safronov, I. Tchouchenkov, H. Woern // Proc. of the 9th Int. Workshop on Computer Science and Information Technologies CSIT'2007. Krasnounsolsk, Russia, 2007. Vol. 1. P. 85–89.

ОБ АВТОРАХ



Вёрн Хайнц, проф., дир. IPR (Ун-т Карлсруэ, Германия). Дипл. инж. по электротехн. (Ун-т Штутгарта, 1973). PhD по электротехн. (там же, 1978). Иссл. в обл. распознавания образов.



Васильев Владимир Иванович, проф., зав. каф. выч. техн. и защ. инф. Дипл. инж. по промэлектронике (УГАТУ, 1970). Д-р техн. наук по сист. анализу и автом. управлению (ЦИАМ, 1990). Иссл. в обл. много-связн., многофункц. и интел. систем.



Сафронов Кирилл Валерьевич, инж. IPR (Ун-т Карлсруэ, Германия). Дипл. инж. по организации и технологии защиты информации (УГАТУ, 2003). Иссл. в обл. распознавания образов.



Чушенков Игорь Михайлович, рук. иссл. группы IPR (Ун-т Карлсруэ, Германия). Дипл. инж. по радиофизике (БелГУ, 1984). Канд. техн. наук по радиофизике (там же, 1988). Иссл. в обл. распознавания образов.