

М.А.ДЖЕНЮК, магистр, НТУ «ХПИ»

РАЗРАБОТКА МЕТОДИКИ СЕГМЕНТИРОВАНИЯ ПОТОКОВ ВИДЕОИНФОРМАЦИИ ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ОБРАБОТКИ ВИДЕОФАЙЛОВ

Розглянуті питання автоматичного створення ієрархічної структури відеофайлів, проаналізовані існуючі методи роботи з відеоінформацією і визначені їхні основні алгоритми. На основі проведеного аналізу створені алгоритм і розроблена програма автоматичного створення ієрархічної структури відеофайлів.

The questions of automatic creation of an outline video files are considered, the existing methods of operation with the videoinformation are parsed and their main algorithms are defined. It was developed the algorithm based on analyses and created the program of automated creation outline video files structure.

Рассмотрены вопросы автоматического создания иерархической структуры видеофайлов, проанализированы существующие методы работы с видеоинформацией и определены их основные алгоритмы. На основе проведенного анализа создан алгоритм и разработана программа автоматического создания иерархической структуры видеофайлов.

Анализ литературы. Постановка проблемы. Увеличение объемов видеоинформации и цифровых сообщений привело к проблемам обработки видеоизображений, которые связаны с извлечением содержания этой информации, определением ее характеристик и хранением полученных данных в цифровой компактной и доступной форме. Манипулирование визуальной информацией возможно лишь при осуществлении быстрого доступа к базе данных изображений, видео или документов, преобразованных в цифровую форму. Эти проблемы частично решены, но остаются всегда актуальными, вследствие разнообразия приложений и высокого разнообразия данных.

Цель статьи – разработка методики сегментирования потоков видеоинформации для автоматизации процесса обработки видеофайлов.

Сегментирование потоков видеоинформации. При работе с видеоизображением анализ содержимого визуальной информации осуществляется для выделения интересующих фрагментов информации и ее интерпретации.

Обработка видеоинформации в зависимости от решения определенных задач подразделяется на соответствующие этапы. Каждый из этих этапов характеризуется собственной задачей и методами ее решения [1,2].

Для выделения объекта на изображении или на кадре видео необходимо поместить описание объекта в память компьютера и научиться его сопоставлять с разными участками изображения. Такой подход применим, когда объект поиска известен. Однако, объектов может быть много, поэтому такая про-

цедура окажется слишком трудоемкой. Сегментацию изображения на объекты (до их распознавания), на основе максимально общих объектных свойств (признаков) называют семантической сегментацией. Одним из важнейших свойств объектов является его способность загроаживать другие объекты.

Таким образом, задача семантической сегментации – это, прежде всего, задача обнаружения краев заслонения.

Сегментация видеодокументов подразделяется на:

- выделение объекта на видеоизображении;
- разделение по временным границам.

Выделение (распознавание) объекта на видеоизображении происходит путем сопоставления описания данного объекта с различными участками видеоизображения (в случае, если объект известен).

При сегментации видеоизображения по временным границам происходит процесс разделения видеодокумента (фильм, телевизионные новости, спортивная программа, документальный фильм, и т.д.) с определением временных границ между последовательностями изображений (элементами) подобно содержания. Переход от одного элемента к другому может быть четким, в течение нескольких кадров, или плавным – около десяти кадров (более сложный для определения).

По своей сути определение временных границ видеодокумента сходно с выделением объекта на изображении, так как происходит путем сопоставления ключевых кадров, хранящихся в графической БД, характерных для конкретного видеосегмента, с текущим кадром видео. Данный тип сегментации является более сложным, так как при сравнении различных кадров зачастую используется распознавание объектов, содержащихся в данных кадрах, которые позволяют классифицировать данный кадр и определить, к какому типу сегментов он относится.

Под распознаванием понимается возможность выделения совокупности признаков идентифицирующих объект. Любая задача распознавания сводится к самостоятельной классификации машинного изображения, без помощи человека. При этом распознавание делится на большое количество отдельных задач. В основе решения каждой задачи лежит своя технология, основанная на определенных принципах, различающихся в зависимости от вида распознавания [3-5].

Существует ряд задач, связанных с проблемой распознавания. К ним относятся задачи распознавания сцен, образов, текста. Основной из них является распознавание образов. Речь идет о ситуации, в которой необходимо определить, какой объект изображен, т.е. необходимо отнести изображение к одному из известных классов.

Следующей задачей, связанной с проблемой распознавания, является распознавание сцен. При этом надо не только классифицировать объекты, но сначала выделить на изображении отдельные объекты, которые являются кандидатами на классификацию.

Обе задачи не имеют до сих пор эффективного решения, так как количество возможных объектов, не говоря уж об их сочетаниях, практически бесконечно, что крайне затрудняет возможность программной реализации и создания алгоритмов в этой области.

В целом проблема распознавания образов состоит из двух частей: обучения и распознавания. Обучение осуществляется путем показа отдельных объектов с указанием их принадлежности тому или другому образу. В результате обучения распознающая система приобретает способность реагировать одинаковыми реакциями на все объекты одного образа и различными на все объекты различных образов. Очень важно, чтобы процесс обучения завершался только путем показа конечного числа объектов без каких-либо других подсказок. В процессе обучения указываются только сами объекты и их принадлежность образу. За обучением следует процесс распознавания новых объектов, который характеризует действия уже обученной системы [5-8]. Автоматизация этих процедур и составляет проблему обучения распознаванию образов.

Автоматическое структурирование документа состоит из двух основных этапов:

- 1) сегментация – определение границ видеофрагментов и разбиение видео документа на элементы;
- 2) индексации – определение элемента и установление связей между элементами в рамках данного документа (при помощи логических правил).

Разбиение на сегменты и определение типа видеосегмента происходит путем попиксельного сравнения каждого кадра видеодокумента с предыдущим. Для ускорения работы алгоритма это сравнение производится между пикселями вертикальных (или горизонтальных) полос кадра. Таким образом, сравнивая каждую вторую (при коэффициенте частоты сравнения $T = 50\%$), третью полосу (при коэффициенте $T = 33\%$) скорость обработки увеличивается в 2 и 3 раза соответственно, а точность сравнения теряется незначительно. Опытным путем было выявлено оптимальное значение коэффициента T равное 20% , то есть сравнение каждой пятой линии. Таким образом, количество сравниваемых пикселей уменьшается в $\frac{100\%}{T}$ раз.

Математическое описание процесса разбиения на сегменты. Обозначим высоту в пикселях n -ого кадра H_n , а ширину W_n . Тогда количество пикселей кадра будет равно $P_n = H_n \cdot W_n$. Пусть количество совпавших после сравнения пикселей равно x (в случае полного совпадения $x = P_n$). Введем допустимый коэффициент совпадения S , который характеризует принадлежность сравниваемых кадров к одному определенному типу сегментов (при сегментации сегменты нумеруются, а определение и присваивание типа вы-

полняется при индексации).

$$x \geq H_n \cdot W_n \cdot S \text{ или } x \geq P_n \cdot S.$$

Опытным путем было определено значение коэффициента совпадения $S = 0,8$, при котором система работает наиболее эффективно. Другими словами, при совпадении 80 % пикселей данного кадра с предыдущим считают эти кадры принадлежащими одному видеосегменту. То есть если $S_m \geq S$, где S_m – текущее значение коэффициента совпадения, получаемое по формуле

$$S_m = \frac{x}{P_n}.$$

В противном случае (если отношение совпавших пикселей x к общему количеству пикселей P_n меньше S) происходит запись сегмента и его временных параметров (время начала и конца) в структуру видеодокументов.

Данный алгоритм эффективно работает при определении резких переходов между сегментами видеодокумента. В случае плавных переходов процесс сегментации происходит следующим образом: если при сравнении данного кадра с предыдущим, последующего с данным и т.д. текущее значение S_m находится в интервале между коэффициентом совпадения S и коэффициентом перехода S_n (опытным путем получено оптимальное значение $S_n = 0,6$), то это означает начало плавного перехода между сегментами.

$$S \geq S_m \geq S_n.$$

При нахождении первого n -ого кадра, относящегося к какому-либо сегменту, будем считать, что $(n-1)$ -ый кадр является концом плавного перехода.

Выводы. Детально разработан концептуальный подход к решению задачи автоматического создания иерархической структуры документов, для чего было определено место сегментации в общей задаче. Была разработана упрощенная схема программы и описаны ее основные модули. Создание структурной схемы решения задачи позволило разработать методику решения поставленной задачи.

Список литературы: 1. Хармон Л. Распознавание образов при помощи цифровых вычислительных машин / Л. Хармон / М.: Мир, 1974. 2. Харп П. Распознавание образов и анализ сцен / Р. Пурда, П. Харп / М.: Мир, 1970. 3. Mubarak Shah Fundamentals of computer vision // Computer Science Department, University of Central Florida, Orlando, 1997. 4. Forsyth David A., Ponce Jean Computer Vision. A modern approach // Prentice Hall, Upper Saddle River, New Jersey 07458. 5. Закревский А.Д. Логика распознавания / Закревский А.Д. / Мн.: Наука и техника, 1988. – 118 с. 6. Shapiro Linda G., Stockman George C. Computer Vision // Prentice Hall, Upper Saddle River, New Jersey 07458. 7. Козлов В.Н. Элементы математической теории зрительного восприятия / Козлов В.Н. / М.: Изд-во Центра прикл. иссл. при мех.-мат. фак-те МГУ, 2001. – 128 с. 8. Richardson J. M. Pattern recognition and group theory // Frontiers of pattern (Watanabe S., ed.). – New York – London: Academic Press, 1972.

Поступила в редакцию 01.04.2012.