

基于粗糙集和模糊聚类的新闻视频镜头 边界检测方法

韩冰 高新波 姬红兵

(西安电子科技大学工程学院, 西安 710071)

摘要 为了将视频分割成镜头,目前的方法都是提取某些特征然后构造不同的相异性函数。然而,太多的特征就会降低镜头分割算法的效率。因此,有必要对每一个镜头检测决策进行特征约简。基于此,提出了基于粗糙集和模糊聚类的分类方法并得到了相应的决策规则。针对新闻场景的特殊性,将镜头分割成突变过渡、渐变过渡以及无场景变化3类。用超过2个小时的新闻视频所做的实验获得了96.5%的查全率和97.9%的准确率。

关键词 镜头边界检测 粗糙集 模糊聚类

中图分类号: TN948.6 **文献标识码**: A **文章编号**: 1006-8961(2007)03-0522-07

A Shot Boundary Detection Method for News Video Based on Rough Sets and Fuzzy Clustering

HAN Bing, GAO Xin-bo, JI Hong-bing

(School of Electronic Engineering, Xidian University, Xi'an 710071)

Abstract As a crucial step in the content-based news video indexing and retrieval system, shot boundary detection attracts much more research interests in recent years. To partition news video into shots, many metrics were constructed to measure the similarity among video frames based on all the available video features. However, too many features will reduce the efficiency of the shot boundary detection. Therefore, it is necessary to perform feature reduction for every decision of the shot boundary. For this purpose, the classification method based on rough sets and fuzzy *c*-means clustering for feature reduction and rule generation is proposed. According to the particularity of news scenes, shot transition can be divided into three types: cut transition, gradual transition and no transition. The efficacy of the proposed method is extensively tested with news programs over 2 hours and 96.5% recall with 97.9% precision have been achieved.

Keywords shot boundary detection, rough set, fuzzy clustering

1 引言

多媒体和视频压缩技术的发展和广泛应用导致了大量视频信息的出现,如何对视频进行分类、管理和检索现在正成为一个值得研究的热门话题。视频信息是海量的,采用人工视频分类方法太耗时且易受人因素为影响。因此,如何研究并实现一个高效、方便且快捷的视频检索系统现在正成为当前多

媒体信息的主要研究内容。在各种视频分析方法中,首要任务是视频分段,即镜头分割。镜头分割的正确与否将直接影响到后续分析步骤(如关键帧提取、场景分析、镜头运动分析等)的准确性。

镜头分割即是要检测镜头间的连接方式。根据镜头连接方式不同可将镜头切换分为突变过渡(abrupt transition)和渐变过渡(gradual transition)两类。现有的算法大都是针对某一种过渡类型检测效果比较好,对于其余类型的镜头过渡性能就比较差。

基金项目:国家自然科学基金项目(60202004);教育部重点项目(104173)

收稿日期:2005-09-09;改回日期:2005-12-05

第一作者简介:韩冰(1978~),女。现为西安电子科技大学模式识别专业博士研究生。主要研究方向为视频检索、模式识别等。已发表论文10余篇。E-mail:bhan@xidian.edu.cn

另外,这些检测标准中大都是使用象素值、亮度或者彩色直方图、边界信息或者压缩域的数据(如DCT系数、宏块类型、运动矢量等),如果测量的相异性函数大于某一阈值的话,就认为发生了镜头过渡^[1-3]。如何利用已有的检测方法和检测特征对所有镜头间的变换进行联合识别和分类已成为当今镜头边界检测的一个研究热点。在保持检测准确率的前提下,为了提高检测效率,有必要对镜头边界的每一个决策进行特征约简。基于此,提出了基于粗糙集和FCM(Fuzzy c-Means)的特征约简和规则生成算法。

首先,提取用作条件属性的视频序列特征,同时给出决策属性。通过计算条件属性之间的相关性,可以得到特征的重要性。最后通过FCM聚类算法对特征进行聚类获得对镜头边界检测有决定意义的特征。对每一个给定的类,用模糊推理产生模糊IF-THEN决策规则。用真实的新闻视频所做的实验结果表明了本文方法的有效性和鲁棒性。

2 粗糙集基本概念

粗糙集(RS)理论^[4-6]是由波兰科学家Pawlak在1982年提出的一种处理含糊和不确定问题的有效的数学工具。它的优点是不需要预先知道某些特征或属性的数量描述,直接从给定问题的描述集合出发,通过不可分辨关系和不可分辨类确定给定问题的近似域(上、下逼近)来描述,揭示出条件属性对于决策属性的重要性,从而明确有用信息和冗余信息,发现隐含在大量数据背后的内在规律。该理论已经渗透到人工智能的各个分支,在机器学习、知识获取、决策分析、知识发现、专家系统、模式识别等方面也都已有成功的应用。

粗糙集 $R(A) = \langle \underline{R}(A), \overline{R}(A) \rangle$ 是由给定集合 A 的下近似 $\underline{R}(A)$ 和上近似 $\overline{R}(A)$ 表示。集合 $BN(A) = \overline{R}(A) - \underline{R}(A)$ 是 A 的粗糙不确定边界。其中 R 是论域 X 上的一个相容关系。 X/R 为等价类族。

定义1 设 U 是对象的非空有限集合,称为论域; A 是属性的非空有限集合; $V = \bigcup_{a \in A} V_a$, V_a 是属性 a 的值域; f 是信息函数, $f: U \times A \rightarrow V$,对于每一个 $x \in U$ 和 $a \in A$, $f(x, a) \in V_a$ 。则决策表可由信息系统来定义, $A = (U, V, f, C \cup D)$ 。 D 的 C 正域定义为

$$POS_C(D) = \bigcup_{X \in U/IND(D)} \underline{C}(X) \quad (1)$$

IND代表不可区分关系。

如果决策值矢量能够通过它相应的条件属性值矢量被完全地预测,则距离属于 $POS_C(D)$,也就是说, $POS_C(D)$ 是系统论域确定性的部分。

在条件属性中,一些属性对于分类结果起着重要的作用,而另外一些可能对分类结果影响不大。因此,属性的重要性或者依赖性可以用来衡量属性对于分类结果的重要性,从而可以对不必要的属性进行约简。

定义2 设 P 和 Q 是属性集。 $P \subseteq R, Q \subseteq R$:
 $U/P = \{X_1, X_2, \dots, X_l\}, U/Q = \{Y_1, Y_2, \dots, Y_l\}$ 。则称知识 Q 是 $k(0 \leq k \leq 1)$ 度依赖于知识 P 的,如果

$$k = \gamma_P(Q) = \frac{|POS_P(Q)|}{|U|} \quad (2)$$

其中, $|\cdot|$ 表示一个给定集合的基数。因此,系数 k 表示应用知识 P 能够确切地划入分类 $U/I(Q)$ 的对象百分比。同时也反映了不同的属性对论域对象的分类能力。

3 基于粗糙集和FCM聚类的镜头过渡检测

3.1 属性选择

在现有的算法中,最常用的就是直接应用各种颜色空间中的色彩值或者直方图特征以及统计特征。根据大量的文献^{[1~3], [7], [8]},总结出了以下12种常用特征。从这些特征中找到镜头检测的重要特征,根据这些重要特征来构造检测相异性函数,从而达到有效的检测。第1类是标准的RGB空间,在镜头过渡的过程当中度量3种基本颜色的变化;第2类是HSI空间,它是衡量帧间亮度、色度和饱和度的变化,在具体计算每一帧时,对于RGB和HSI分量计算每一帧的均值;第3类和第4类是直方图特征,从灰度以及不同的颜色空间来度量镜头间的变化;最后一类就是统计特征,计算每一帧的亮度分量 I 的均值、方差和偏态。具体特征选取和表示如下:

- (1) 图像的红色、绿色和蓝色分量 R, G, B ;
- (2) 图像的色度、饱和度和亮度分量 H, S, I ;
- (3) 灰度直方图 H_{grey} 和颜色直方图:RGB颜色直方图 H_{rgb} 及HSI颜色直方图 H_{hsi} ;
- (4) 统计特征:均值 M 、方差 St 和偏态 P ;

3.2 基于 RS 和 FCM 的镜头边界检测

FCM 算法是目前比较流行的一种模糊聚类算法,它已经在诸多领域获得了广泛的应用,并取得了满意的效果。另外,对于 RS 和模糊聚类的研究,有学者已将 RS 与 K 均值聚类相结合进行图像分割^[9];文献^[10]将 RS 与模糊聚类方法结合对超谱波段进行约简,达到了良好的效果。在此在前人的基础之上将 FCM 与 RS 进行结合应用到镜头分割之中。下面列出了一种通用的 FCM 聚类算法^[11]。

假设 $X = \{x_1, x_2, \dots, x_n\} \subset \mathbf{R}^n$ 为待聚类分析的样本集, $w_{i,j}$ 表示样本 x_i 属于第 j 类的模糊隶属度, $j = 1, 2, \dots, c$ 为给定的聚类类别数。令 v_j 为第 j 类的聚类中心, $d_{i,j} = \|x_i - v_j\|$ 表示样本数据 x_i 到聚类中心 v_j 的距离,是特征间的相似性度量。则基于 FCM 的模糊聚类分析算法可归纳如下,其中 $\|\cdot\|$ 为 \mathbf{R}^n 上任意一种内积范数, m 为模糊加权指数。

初始化:设定迭代停止阈值 ε , 初始化原型模式 $P^{(0)}$, 设置迭代计数器 $b = 0$;

(1) 计算或更新分类矩阵 $W^{(b)}$

对于 $\forall i, k$, 如果 $\exists d_{i,k}^{(b)} > 0$, 则有

$$w_{i,k}^{(b)} = \left\{ \sum_{j=1}^c \left[\frac{d_{i,j}^{(b)}}{d_{i,k}^{(b)}} \right]^{\frac{2}{m-1}} \right\}^{-1} \quad (3)$$

如果 $\forall i, r$, 使得 $d_{i,r}^{(b)} = 0$, 则有

$$w_{i,r}^{(b)} = 1 \quad \text{且对 } j \neq r, w_{i,j}^{(b)} = 0 \quad (4)$$

(2) 更新聚类原型模式矩阵 $P^{(b+1)}$

$$p_i^{(b+1)} = \frac{\sum_{k=1}^n (w_{i,k}^{(b)})^\alpha \cdot x_k}{\sum_{k=1}^n (w_{i,k}^{(b)})^\alpha}, i = 1, 2, \dots, c \quad (5)$$

(3) 如果 $\|P^{(b)} - P^{(b+1)}\| < \varepsilon$, 则算法停止并输出划分矩阵和聚类原型, 否则令 $b = b + 1$, 转向步骤 1。

将视频的各种特征作为条件属性集合 $C = \{c_1, c_2, \dots, c_m\}$ ($m \in \{1, 2, \dots, 12\}$), 镜头的过渡类型为决策属性 $Q = d$; 在 FCM 算法中特征的特征向量则可定义为各种特征的分类能力, 即

$$x_k = \gamma_{c,k}(d), k = 1, 2, \dots, n \quad (6)$$

通过聚类, 选择离类中心最近的特征作为主要特征, 但是在此过程中发现, 有的特征虽然离类中心的距离最近, 但是它的分类能力并不是最优的, 因为其受到类数的影响。再者, $\gamma_{c,k}(d)$ 反映的是一个特征总的分类能力, 而不同的特征对不同镜头间过渡形式的分类也是有差异的, 因此, 直接应用 $\gamma_{c,k}(d)$ 来度

量特征之间的相似性或者重要性是不够准确的。将上述定义进行改进,

$$x_k = \{x_{1,m}, x_{2,m}, \dots, x_{k,m}\} \quad (7)$$

其中, $x_{k,j} = |c_k(Y_j)|/Y_j, j = 1, 2, \dots, m$ 。其中 $c_k = \gamma_{c,k}(d)$ 表示根据条件属性 c_k ($k \in \{1, 2, \dots, 12\}$) 的分类能力。 Y_j 表示知识 Q 对论域的划分, 即 $U/Q = \{Y_1, Y_2, \dots, Y_j\}$ 。 $x_{k,j}$ 反映了每一类镜头过渡能够被第 k 个特征识别出来的比例。对相同的镜头过渡具有近似相同分类效果的特征被认为是相似的, 通过对 x_k ($k = 1, 2, \dots, n$) 进行聚类将具有相似分类能力的特征划为一组, 然后根据最大隶属度原则考虑在相近的特征中相关性更强的特性, 从各组中只选其一, 得到了原始高维特征的一个近似等价约简。

3.3 规则生成

经过大量的观察实验发现, 在新闻视频中, 一般镜头过渡不会超过 4s, 即不超过 120 帧, 因此, 仅仅考虑每段新闻视频片断中镜头过渡前后各 60 帧左右范围的属性变换。每一种新闻示例中都有一种镜头过渡或者是无场景变化。存在的过渡类型有突变、溶解、划变、摄像机摇镜头以及无场景变化五类; 决策集包括突变、渐变、无场景变化 3 类。计算过渡过程中各种特征的变化值, 并做归一化处理。另外, 构造各个特征的隶属度, 首先考虑各个过渡形式出现的几率是差不多的且为了便于计算, 因而将归一化后的属性值大致 5 等分。由于对于一些渐变过渡具有较大的变化范围, 因此采用等腰梯形隶属度函数; 对于介于突变和渐变以及无镜头变化和渐变之间的过渡不是很明显, 因而采用三角形隶属度函数; 针对无镜头过渡和突变过渡具有明显的变化这种特点特征, 因而采用直角梯形的隶属度函数。据此建立了如图 1 所示的各个特征的隶属度。根据图 1 以及最大隶属度原则将其转化为相应的模糊语言量。

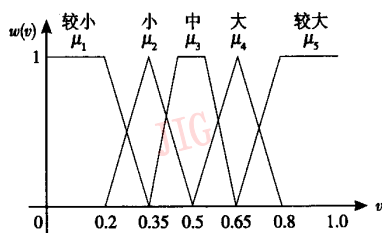


图 1 表中各个特征方法的隶属度

Fig. 1 Definition of the fuzzy variable in Table 1

表 1 为 30 个测试视频 (包括突变、溶解、划变等

表 1 镜头过渡信息表
Tab. 1 The decision table for shot transition

示例	属性集												决策集	
	R	G	B	H	S	V	H_{grey}	H_{rgb}	H_{hsi}	M	St	P	存在的过渡类型	决策
1	大	大	大	较小	大	大	中	中	中	大	较小	较小	镜头拉伸	N
2	较大	较大	较大	较大	较大	较大	较大	较大	较大	大	大	大	突变	C
3	较小	较小	较小	较小	大	较小	较小	较小	较小	小	较小	较小	摇镜头	N
4	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	无过渡	N
5	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
6	中	中	大	中	大	中	中	中	中	中	中	中	划变	G
7	较小	较小	较小	较小	较小	大	较小	较小	大	较小	较小	较小	无过渡	N
8	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
9	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
10	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
11	中	中	中	中	较小	中	中	小	小	中	中	较小	无过渡	N
12	中	中	中	中	中	中	中	中	中	中	中	中	划变	G
13	较大	较大	较大	小	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
14	中	中	中	大	大	中	中	中	中	中	中	中	划变	G
15	较大	较大	较大	较大	小	较大	较大	较大	较大	较大	小	较大	突变	C
16	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
17	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	较小	镜头拉伸	N
18	中	中	小	较小	较小	较小	较小	较小	较小	中	较小	较小	镜头拉伸	N
19	较小	小	小	较小	较小	较小	较小	较小	较小	较小	较小	较小	无过渡	N
20	较大	较大	较大	中	较大	较大	较大	较大	较大	大	较大	较大	突变	C
21	小	小	较小	较小	较小	中	较小	较小	较小	小	大	大	镜头拉伸	N
22	较小	较小	较小	较小	较小	较小	大	中	中	小	较小	较小	摇镜头	N
23	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C
24	中	中	中	中	中	小	中	中	中	中	中	中	溶解	G
25	中	中	大	中	中	中	中	小	较小	小	小	小	镜头拉伸	N
26	较小	较小	较小	较小	较小	较小	大	较小	大	较小	较小	较小	无过渡	N
27	中	中	小	中	中	中	中	中	中	小	中	中	溶解	G
28	中	中	中	中	中	中	中	中	中	中	中	中	溶解	G
29	大	大	大	大	大	大	大	较大	较大	大	大	大	突变	C
30	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	较大	突变	C

注: C 表示 Cut transition, 突变过渡; G 表示 Gradual transition, 渐变过渡; N 表示 No transition, 无镜头过渡

过渡形式和摇镜头、镜头拉伸等无场景变化以及诸如运动、闪光、字幕的出现等影响)的特征变化表。

从表 1 中可以看出, R、G、B 分别对单一的颜色敏感, 若场景中突然出现或消失(或者是部分的出现或者消失) R 或者 G 或者 B 中的任意一种, 则会有突然的变化而被认为是突变过渡; 而对于 H, 它对闪光灯和字幕的出现非常敏感。而在第 29 段测试

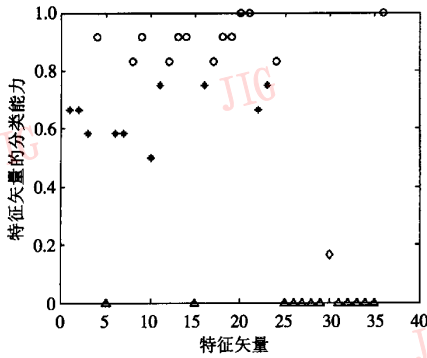
视频中, 含有大量剧烈的运动, 仅有 H_{rgb} 特征和 H_{hsi} 特征对运动抑制较好。

根据上述的决策表及式(1)、(2), 可以得出每一个特征的下近似 $\underline{C}(x)$ 。再根据式(3)、(4)、(9), 可以分别计算出聚类的特征矢量, 即每一特征对不同的镜头过渡的分类能力。 $x_{1,1} = 8/12, x_{2,1} = 8/12, x_{3,1} = 7/12, x_{4,1} = 11/12, x_{5,1} = 0, x_{6,1} = 7/12,$

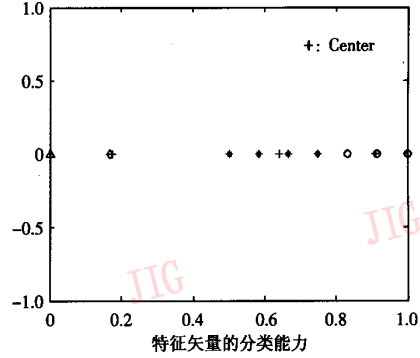
$x_{7,1} = 7/12, x_{8,1} = 10/12, x_{9,1} = 11/12, x_{10,1} = 8/12,$
 $x_{11,1} = 9/12, x_{12,1} = 10/12; x_{1,2} = 11/12, x_{2,2} =$
 $11/12, x_{3,2} = 0, x_{4,2} = 10/12, x_{5,2} = 11/12, x_{6,2} =$
 $11/12, x_{7,2} = 11/12, x_{8,2} = 12/12, x_{9,2} = 12/12, x_{10,2} =$
 $8/12, x_{11,2} = 9/12, x_{12,2} = 10/12; x_{1,3} = 0, x_{2,3} = 0,$
 $x_{3,3} = 10/12, x_{4,3} = 0, x_{5,3} = 0, x_{6,3} = 1/6, x_{7,3} = 0,$

$x_{8,3} = 0, x_{9,3} = 0, x_{10,3} = 0, x_{11,3} = 0, x_{12,3} = 6/6。$

如上,得到 36 个对于不同镜头的特征矢量。首先将已有的特征矢量分类,如果特征选取的不够则影响算法的准确度,而选择过多又会增加算法的复杂度。因而设类数 $c = 4$ 。聚类结果如图 2 所示。



(a) 数据空间聚类图



(b) 聚类中心在坐标轴上的位置图

图 2 聚类结果图

Fig. 2 The result of clustering

由图 2(a)中可以看出,以上特征矢量分为 4 类,其中第 30 个特征矢量,即 $x_{6,3}$ 被单独划成一类,则说明该特征是非常重要的,且必不可少。由此也可以得出是第 6 个特征对第 3 种过渡是必不可少的。该特征为亮度特征 V ,它对第 3 种过渡——渐变过渡起着重要的作用。从图 2(b)可以得到相应的聚类中心的位置,离 4 个聚类中心最近的特征矢量分别为 $x_{4,1}, x_{9,1}, x_{10,1}, x_{6,2}, x_{7,2}, x_{10,2}, x_{6,3}, x_{12,3}$ 。由此,可以得到如下的镜头过渡检测决策:

- (1) If $H = \mu_5$ or $H = \mu_4$ and $H_{hsv} = \mu_5$ or $H_{hsv} = \mu_4$ and $m = \mu_5$ or $m = \mu_4$, Then 无镜头过渡;
- (2) If $V = \mu_1$ or $V = \mu_2$ and $H_{grey} = \mu_1$ or $H_{grey} = \mu_2$ and $m = \mu_1$ or $m = \mu_2$, Then 突变过渡;
- (3) If $V = \mu_3$ and $P = \mu_3$, Then 渐变过渡。

随后,根据检测到的特征,利用帧间差构造相异性函数,再对镜头边界进行检测。

4 实验结果与分析

用中央电视台新闻联播中共 5 段节目对算法进行了测试,详细信息如表 2 所示。为了验证所提出新闻镜头分割算法的有效性,采用两个标准:查全率和准确率,其定义如下:

$$\begin{aligned} \text{查全率} &= \text{正确检出数} / (\text{正确检出数} + \text{漏检数}) \\ \text{准确率} &= \text{正确检出数} / (\text{正确检出数} + \text{误检数}) \end{aligned} \quad (8)$$

其中,正确检出数 = 检出数 - 误检数。

表 2 新闻节目的详细信息

Tab. 2 Detailed information on the news video programs

新闻视频	1	2	3	4	5	总共
长度(s)	2 095	1 802	2 186	1 487	944	8 514
镜头边界数	374	323	396	247	187	1 527
突变镜头数	339	315	390	237	178	1 459
渐变镜头数	35	8	6	10	9	68

首先用人工方法对视频进行切分,结果作为检测到的正确镜头个数,由于视频序列的数据量很大,一般认为误差在 10 帧以内都是允许的。本文方法首先自适应的从众多的检测特征中选取了最适应于各种过渡检测的特征,并且形成了镜头边界检测的一般规则。由于现有方法大都是针对一种过渡进行检测,而直方图方法^[12]通过设置门限可以同时用来检测突变和渐变过渡,与本文方法具有可比性,结果列于表 3。

5 段新闻视频共有 1 527 个镜头边界,本文方法检测到 1 505 个镜头边界,其中包括 53 个漏检和 31

表 3 本文方法与直方图方法对比检测结果

Tab.3 The comparison of our method with the histogram method(DOH)

测试视频	本文方法					直方图方法 ^[12]				
	检出数	漏检数	误检数	查全率(%)	准确率(%)	检出数	漏检数	误检数	查全率(%)	准确率(%)
新闻 1	370	12	8	96.8	97.9	380	30	36	92.0	90.5
新闻 2	320	10	7	96.9	97.8	350	15	42	95.4	88
新闻 3	388	16	8	96.0	98.0	400	46	50	88.4	87.5
新闻 4	245	7	5	97.2	98.0	250	22	25	91.1	90.0
新闻 5	182	8	3	95.7	98.3	190	17	20	90.9	89.5
平均	1 505	53	31	96.5	97.9	1570	130	173	91.5	89.0

个误检。准确率和查全率分别达到 97.9% 和 96.5%。检测过程中的漏检主要是由于在一些镜头边界中帧对间小的内容变化,而大多数错检是由于在渐变过程中不规则的摄像机运动造成的。而对于直方图方法,选择了多个门限,经过多次的实验,选取了检测率最好的。如此下来,为了获得较高的检测结果,其检测的复杂度是相当高的。

就计算量而言,本文方法运算复杂度略高于直方图方法。因为直方图方法只应用一种检测特征,这显然不符合视频多变的特性,因此其检测率较低。本文方法虽然多用了一些特征,但是这些特征并不是随意选取的,而是自适应选择的,同时,检测率是相当高的。

粗糙集部分属性离散化的复杂度为 $O(|U|)$ 。求属性对论域划分最坏的复杂度为 $O(|C||U|^2)$,其中, C 为属性集合, U 为对象集合。本文中 C 每

次只有一个属性,因此复杂度为 $O(|U|^2)$ 。若是对对象进行排序之后再行划分,它的复杂度就变为 $O(|U|\lg|U|)$ 。在给定划分的复杂度为 $O(|U|)$,所以求一个决策类的下近似的整个复杂度就是 $O(|U|\lg|U|)$ 。因此,对于 m 个决策类分别由 n 个条件属性所决定的下近似的复杂度为 $O(mn|U|\lg|U|)$ 。模糊聚类部分的复杂度为 $O(c|U|)$, c 为聚类数。在此, $m=3, n=12, c=3$ 。由上面的分析可以看出,本文算法基本上复杂性只与 U (即样本集合) 有关,即随着测试视频数量的增多,本文方法复杂度才会有明显增加。

如果采用每一特征的总的分类 $\gamma_{c_k}(d)$ 作为聚类特征矢量,即 $x_1 = 19/30, x_2 = 19/30, x_3 = 7/30, x_4 = 20/30, x_5 = 11/30, x_6 = 19/30, x_7 = 18/30, x_8 = 19/30, x_9 = 23/30, x_{10} = 14/30, x_{11} = 18/30, x_{12} = 27/30$ 。从而得到聚类结果如图 3 所示。

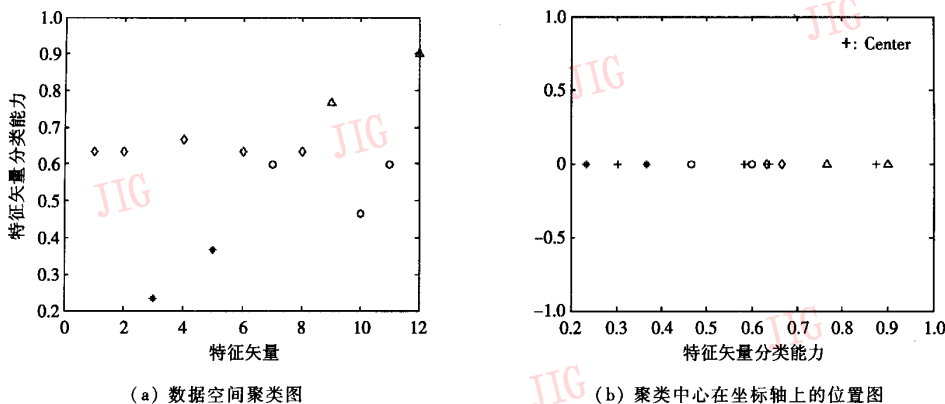


图 3 聚类结果图

Fig.3 The result of clustering

进而选择的决策特征为 B, H, S, H_{grey}, S, P 。即得到的决策为

(1) If $B = \mu_5$ or $B = \mu_4$ and $H = \mu_5$ or $H = \mu_4$ and $S = \mu_5$ or $S = \mu_4$ and $H_{grey} = \mu_5$ or $H_{grey} = \mu_4$ and $S = \mu_5$

or $S = \mu_4, P = \mu_5$ or $P = \mu_4$ Then 无镜头过渡;

(2) If $B = \mu_1$ or $B = \mu_2$ and $H = \mu_1$ or $H = \mu_2$ and $S = \mu_1$ or $S = \mu_2$ and $H_{grey} = \mu_1$ or $H_{grey} = \mu_2$ and $S = \mu_1$ or $S = \mu_2$, and $P = \mu_1$ or $P = \mu_2$ Then 突变过渡;

(3) If $B = \mu_3$ and $H = \mu_3$ and $S = \mu_3$ and $H_{grey} = \mu_3$ and $S = \mu_3$ and $P = \mu_3$ Then 渐变过渡。

用以上的决策对相同的视频片断进行检测,得到 96% 的查全率和 94% 的查准率。虽然准确率是可以接受的,但是从中可以发现直接用 $\gamma_{c_k}(d)$ 所用的特征要多,从而计算复杂度明显高于用 $x_{k,j}$ 作为特征矢量的运算复杂度。这是由于采用每一种特征对于不同过渡的分类能力作为特征矢量,这比采用每一特征总的分类能力作为特征矢量更符合实际情况。

5 结 论

本文提出基于 RS 和 FCM 聚类的镜头转换检测特征约简方法,根据约简结果得出检测不同镜头过渡的规则。视频图像中可提取的特征很多,这些特征间的相关性不同,且对于不同镜头的过渡,其分类能力也不同。针对不同的特征对不同镜头间过渡形式分类效果各不相同这一特点,粗糙集理论能够用属性的分类能力(或属性间的依赖性)定义特征之间的相关性,再通过 FCM 聚类,从而得到了最优的特征约简,然后根据约简结果得出镜头过渡类型检测的决策。实验结果证明,该方法不但简单有效而且减少数据维数的同时又最大程度的保留了原始视频图像的信息。

参考文献 (References)

- 1 John S Boreczky, Lawrence A Rowe. Comparison of video shot boundary detection techniques[A]. In: SPIE Conference of Storage & Retrieval for Image & Video Databases[C], San Jose, California, USA, 1996: 170 ~ 179.
- 2 Gargi U, Kasturi R, Strayer S H. Performance characterization of video-shot-change detection methods [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(1): 1 ~ 13.
- 3 Ford R M, Robson C, Temple D, et al. Metrics for shot boundary detection in digital video sequences[J]. Multimedia System, 2000, 8(1): 37 ~ 46.
- 4 Pawlak Z. Rough Set [J]. International Journal of Computer and Information Science, 1982, 11(5): 341 ~ 356.
- 5 Pawlak Z. Vagueness and uncertainty: a rough set perspective[J]. International Journal of Computational Intelligence, 1995, 11(2): 227 ~ 232.
- 6 Wang Guo-yin, Zhao Jun, An Jiu-jiang, et al. Theoretical study on attribute reduction of rough set theory: comparison of algebra and information views [A]. In: Proceedings of the Third IEEE International Conference on Cognitive Informatics [C], Victoria, British Columbia, Canada, 2004: 148 ~ 155.
- 7 Gao Xin-bo, Tang Xiao-ou. Unsupervised model-free news video segmentation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(9): 765 ~ 776.
- 8 Han Bing, Gao Xin-bo, Ji Hong-bing. An efficient algorithm of gradual transition for shot boundary segmentation [A]. In: SPIE Conference on Multispectral Image Processing and Pattern Recognition [C], Beijing, China, 2003: 956 ~ 961.
- 9 Liu Yan, Yue Ying-juan, Li Yan-jun, et al. Application of rough set and K-means clustering in image segmentation [J]. Infrared and Laser Engineering, 2004, 33(3): 300 ~ 302. [刘岩, 岳应娟, 李言俊等. 基于粗糙集的图像聚类分割方法研究[J]. 红外与激光工程, 2004, 33(3): 300 ~ 302.]
- 10 Shi Hong, Shen Yi, Liu Zhi-yan. Hyperspectral band reduction based on RS and fuzzy c-means clustering[J]. Journal of Electronics & Information Technology, 2004, 26(4): 619 ~ 624. [石红, 沈毅, 刘志言. 基于粗糙集和模糊聚类的超谱波段约简[J]. 电子与信息学报, 2004, 26(4): 619 ~ 624.]
- 11 Bezdek J C. Clustering validity with fuzzy sets [J]. Mathematical Biology, 1974, (1): 57 ~ 71.
- 12 Zhang H J, Kankanhalli A. Automatic partitioning of full-motion video[J]. ACM Multimedia Systems, 1993, 1(1): 10 ~ 28.