第37卷 第17期 计 算 机 工 程 2011年9月
Vol.37 No.17 Computer Engineering September 2011

·开发研究与设计技术· 文章编号：1000—3428(2011)17—0265—03 文献标识码：A 中图分类号：TP301.6

基于粗糙集的压缩域关键帧提取算法

史丽春[1,2]，蔡静之[1]，张明新[2]

(1. 北方民族大学电气信息工程学院，宁夏 银川 750021；2. 常熟理工学院计算机科学与工程学院，江苏 常熟 215500)

摘 要：提出一种基于粗糙集(RS)的压缩域关键帧提取算法。该算法从视频的压缩域数据流中提取I帧，以相邻I帧的差值为行、多个特征属性为列建立信息系统，对该系统进行归一化、离散化，采用RS的属性约简理论从中得到没有冗余的属性核，利用不可分辨关系进行关键帧提取。与像素差法、直流系数法的性能比较结果表明，该算法的计算复杂度更低，且适用于不同类型的视频。

关键词：压缩域；关键帧提取；粗糙集；信息系统；不可分辨关系

Key Frame Extraction Algorithm Based on Rough Set in Compressed Domain

SHI Li-chun[1,2], CAI Jing-zhi[1], ZHANG Ming-xin[2]

(1. College of Electronics and Information Engineering, North Ethnic University, Yinchuan 750021, China;
2. College of Computer Science and Engineering, Changshu Institute of Technology, Changshu 215500, China)

【Abstract】This paper proposes a key frame extraction algorithm based on Rough Set(RS) in compressed domain. It extracts I frame from compressed domain data sequence, and constructs information system with the difference between two adjacent I frames in column and attributes sets which are extracted from decompressed I frames in row, then the established information system is normalized and discredited. It simplifies information system by using attribute reduce theory of RS to obtain attribute cores without redundancy, and key frame is extracted by using the indiscernibility relation of RS. Experimental result shows that, competed with pixel difference algorithm and direct current coefficient algorithm, the algorithm proposed can reduce the complexity of computing and can be used for different types of video effectively.

【Key words】compressed domain; key frame extraction; Rough Set(RS); information system; indiscernibility relation

DOI: 10.3969/j.issn.1000-3428.2011.17.089

1 概述

在基于内容的视频检索[1]中，为了建立索引，往往需要提取镜头的关键帧来反映一个镜头的主要内容。关键帧提取是基于内容的视频检索中的关键技术，目前出现了基于镜头、纹理、形状、颜色、运动等各种检索方法[2]。例如，固定间隔抽取法、颜色特征法、镜头边界法、帧平均法、直方图平均法、文献[3]的光流分析法、文献[4]的镜头运动逼近方法、文献[5]的无监督聚类方法。但这些算法采用的视频特征过于单一，不能用于不同类型的视频。

粗糙集(Rough Set, RS)理论不需要预先给定检测对象的数学描述，直接从知识分类出发，导出问题的决策。利用该特点，本文提出一种基于RS理论的压缩域关键帧提取算法，充分利用视频中多种特征的特征信息，使算法适用于不同类型的视频。

2 粗糙集理论

RS理论是[6]一种处理含糊和不确定问题的有效的数学工具。下面介绍RS中的一些基本概念。

定义1(不可分辨关系) 给定一个论域U和U上的一簇等价关系S，若$P\subseteq S$且$P\neq\varnothing$，则$\cap P$仍是论域U上的一簇等价关系，称为P上的不可分辨关系，记作$IND(P)$，常简记为P。对$\forall x\in U$，$[x]_{IND(P)}=[x]_P=\cap[x]_R$。则$U/IND(P)=\{[x]_{IND(P)}\mid\forall x\in U\}$表示与等价关系$IND(P)$相关的知识，常用$U/P$表示。

定义2(约简和核) 令R为一簇等价关系，$P\in R$，如果$ind(R)=ind(R-\{P\})$，则称P为R中不必要的；否则，称P为R中必要的。如果每一个$P\in R$都为R中必要的，则称R为独立的；否则，称R为依赖的。

如果$Q\subseteq R$是独立的，且$ind(Q)=ind(R)$，则称Q为R的一个约简，显然R可以有很多约简，R中所有必要关系组成的集合称为R的核，记作$core(R)$。约简和核的关系为：$core(R)=\cap red(R)$，其中，$red(R)$表示R的所有约简。

定义3(信息系统) 形式上，四元组$S=(U,A,V,f)$是一个知识表达系统，其中，U为对象的非空有限集合，称为论域；A为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；$f:U\times A\to V$是一个信息函数，它为每个对象的每个属性值赋予了一个信息值，即$\forall a\in A,x\in U,f(x,a)\in V_a$，信息表常用$S=(U,A)$代替$S=(U,A,V,f)$。

3 关键帧I帧

MPEG视频中包含3种帧：I帧，P帧和B帧。I帧利用自身的相关性压缩提供压缩数据流中的随机存取点，采用基于DCT的编码技术。I帧是MPEG视频编码的关键帧，采用帧内编码，记录了视频流的主要信息。基于I帧进行视频处理不仅可以保留基本的视频原始信息，且大大减少了数据量。

基金项目：甘肃省自然科学基金资助项目(3ZS051-A25-047)

作者简介：史丽春(1984－)，女，硕士研究生，主研方向：图像与视频检索；蔡静之，教授；张明新，教授、博士

收稿日期：2011-01-20 **E-mail：**shilichun6@163.com

4 基于 RS 的压缩域关键帧提取算法

本文的压缩域关键帧提取算法步骤如下：

第 1 步 从 MPEG2 二进制数据流中提取一个镜头的所有 I 帧，如图 1 为一个镜头的 I 帧序列。

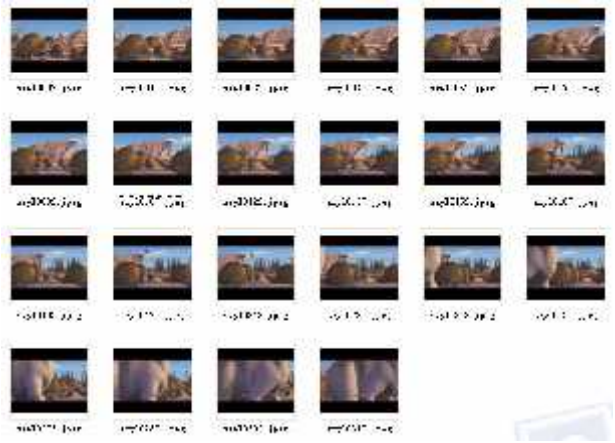


图 1 一个镜头的 I 帧序列

第 2 步 建立镜头的信息系统 S，其中，行为属性集合 A={亮度直流系数(*DC*), 颜色直方图差值(*DOH*), 差值直方图(*HOD*), 像素差(*PIXCEL*), 边缘变化率(*ECR*)}；列为论域 U，以相邻像素差 I_i 为元素，将缺省属性值补 0，如表 1 所示。

表 1 信息系统

I_i	*DC*	*DOH*	*HOD*	*PIXCEL*	*ECR*
I_1	1.139 5	7 320	275 325.2	557 618	0.024 752
I_2	2.125 5	5 840	580 209.3	508 642	0.010 044
I_3	0.702 2	7 254	1 402 172.0	488 892	0.051 093
I_4	9.190 9	6 468	437 063.9	546 052	0.028 004
I_5	1.681 8	7 600	397 383.4	557 762	0.039 252
I_6	3.913 9	8 068	1 178 621.0	477 787	0.081 712
I_7	1.443 5	8 284	306 050.5	575 953	0.027 046
I_8	5.273 8	9 088	2 715 389.0	471 864	0.042 371
I_9	8.142 5	8 518	1 709 158.0	577 216	0.087 084
I_{10}	1.815 5	7 726	1 269 871.0	490 806	0.057 052
I_{11}	1.082 0	7 374	756 754.9	488 328	0.067 365
I_{12}	0.156 2	7 956	1 210 610.0	499 311	0.002 104
I_{13}	3.622 7	11 742	1 083 071.0	526 578	0.044 156
I_{14}	1.897 1	10 290	3 168 855.0	562 285	0.038 598
I_{15}	14.082 6	9 114	770 756.9	490 060	0.026 383
I_{16}	7.185 0	9 512	1 289 995.0	565 999	0.050 399
I_{17}	16.539 8	19 154	951 221.0	873 111	0.068 341
I_{18}	0.538 2	19 496	448 907.8	716 074	0.041 623
I_{19}	1.699 1	14 644	4 219 860.0	972 259	0.147 042
I_{20}	47.998 1	28 446	0.0	1 380 764	0.117 372

系统建立过程如下：

(1)选择属性集合 A：$A=\{a_1,a_2,\cdots,a_n\}$，a_i，$1\leqslant i$ n 为第 i 个属性。例：$A=\{a_1,a_2,a_3,a_4,a_5\}$={直流系数, 灰度直方图, 灰度差分直方图, 像素差, 边缘变化率}。

(2)确定论域 U：$U=\{u_1,u_2,\cdots,u_m\}$ 为对象论域，其中，$u_i=I_i-I_{i-1}$，$1\leqslant i$ m，I_i 为第 i 个 I 帧。以图 1 所示的视频片断为例，其对象论域为 $U=\{u_1,u_2,\cdots,u_{20}\}$。

(3)确定信息系统值域 V：V 为信息系统的值域，V_a 为属性 a 的值域。如对于图 1 所示的视频片断，其信息系统值域为 $V=\bigcup_{a\in A}V_a$，V_a 为大于 0 小于正无穷的实数。

(4)建立信息函数 f：$f:U\times A\to V$ 是每一个属性与其值域对应的信息函数。如对于图 1 所示的视频片断，其对 A 中所选属性建立的信息函数 f 如下：

1) $f_{a_1}:U\times a_1\to V_{a_1}$：将第 i 个 I 帧图像分为 8×8 矩阵并进行离散余弦变换(Discrete Cosine Transform,DCT)。提取第 i 个 I 帧图像每个 8×8 矩阵的直流系数得到 36×44 矩阵 $\boldsymbol{H}_i$。

$$f_{a_1}(u_i)=\frac{1}{36\times 44}\sum_{i=0}^{35}\sum_{j=0}^{43}\left|\boldsymbol{H}_i(i,j)-\boldsymbol{H}_{i-1}(i,j)\right|,u\in U$$

2) $f_{a_2}:U\times a_2\to V_{a_2}$：$K_i(j)$ 为第 i 个 I 帧落在第 j 个颜色区间的像素数：$f_{a_2}(u_i)=\sum_{j=0}^{n-1}\left|K_i(j)-K_{i-1}(j)\right|,u\in U$。

3) $f_{a_3}:U\times a_3\to V_{a_3}$：$L_{i,i+1}(j)$ 为第 i 个 I 帧差值直方图落在第 j 个颜色区间的像素数。$F(i)$ 为第 i 个 I 帧的差分。$f_{a_3}(u_i)=\left|F(i)-F(i-1)\right|$。

4) $f_{a_4}:U\times a_4\to V_{a_4}$：$O_i(i,j)$ 为第 i 个 I 帧像素灰度值。$f_{a_4}(u_i)=\sum_{i=0}^{287}\sum_{j=0}^{351}\left|O_i(i,j)-O_{i-1}(i,j)\right|$。

5) $f_{a_5}:U\times a_5\to V_{a_5}$：$P(i)$ 为统计的第 i 个 I 帧的图像边缘像素数。$f_{a_5}(u_i)=\dfrac{\left|P(i)-P(i-1)\right|}{P(i-1)}$。

第 3 步 确定不可分辨关系：

$$R=\left\{(u_i,u_j)\middle|u_i,u_j\in U\wedge f_a(u_i)\equiv f_a(u_j)\right\},a\in A$$

第 4 步 信息系统的数据归一化，归一化后的信息系统如表 2 所示。

表 2 归一化后的信息系统

I_i	*DC*	*DOH*	*HOD*	*PIXCEL*	*ECR*
I_1	0.020 55	0.065 47	0.000 00	0.094 35	0.156 26
I_2	0.041 16	0.000 00	0.077 29	0.040 46	0.054 78
I_3	0.011 41	0.062 55	0.285 67	0.018 74	0.338 00
I_4	0.188 85	0.027 78	0.041 00	0.081 62	0.178 70
I_5	0.031 89	0.077 86	0.030 94	0.094 51	0.256 30
I_6	0.078 54	0.098 56	0.228 10	0.006 52	0.549 26
I_7	0.026 91	0.108 11	0.007 79	0.114 52	0.172 09
I_8	0.106 97	0.143 68	0.618 59	0.000 00	0.277 82
I_9	0.166 93	0.118 46	0.363 50	0.115 91	0.586 32
I_{10}	0.034 68	0.083 43	0.252 13	0.020 84	0.379 11
I_{11}	0.019 35	0.067 86	0.122 05	0.018 11	0.450 27
I_{12}	0.000 00	0.093 60	0.237 11	0.030 20	0.000 00
I_{13}	0.072 46	0.261 08	0.204 78	0.060 20	0.290 14
I_{14}	0.036 39	0.196 85	0.733 56	0.099 48	0.251 79
I_{15}	0.291 09	0.144 83	0.125 60	0.020 02	0.167 51
I_{16}	0.146 92	0.162 44	0.257 23	0.103 57	0.333 21
I_{17}	0.342 45	0.588 96	0.171 35	0.441 46	0.457 00
I_{18}	0.007 99	0.604 09	0.044 00	0.268 69	0.272 66
I_{19}	0.032 25	0.389 45	1.000 00	0.550 55	1.000 00
I_{20}	1.000 00	1.000 00	0.000 00	1.000 00	0.795 29

第 5 步 信息系统的数据离散化，数据离散化后的信息系统如表 3 所示。

表 3 数据离散化后的信息系统

I_i	*DC*	*DOH*	*HOD*	*PIXCEL*	*ECR*
I_1	0	0	0	0	0
I_2	0	0	0	0	0
I_3	0	0	0	0	0
I_4	0	0	0	0	0
I_5	0	0	0	0	0
I_6	0	0	0	0	0
I_7	0	0	0	0	0
I_8	0	0	0	0	0
I_9	0	0	0	0	0
I_{10}	0	0	0	0	0
I_{11}	0	0	0	0	0
I_{12}	0	0	0	0	0
I_{13}	0	0	0	0	0
I_{14}	0	0	1	0	0
I_{15}	0	0	0	0	0
I_{16}	0	0	0	0	0
I_{17}	0	0	0	0	0
I_{18}	0	0	0	0	0
I_{19}	0	0	1	0	1
I_{20}	1	1	0	1	1

第 6 步 信息系统的属性约简。

本文采用基于 Pawlak 属性重要度的属性约简算法。得到范例的属性核 $core(A)=\{a_3,a_5\}$。

第 7 步 提取镜头的关键帧，结果如图 2 所示。



图 2　提取的镜头关键帧序列

5　实验结果与分析

实验以 Windows XP 为操作平台，以 Matlab7 为工具，利用电影片段、动画片段、风景片段 3 种不同的视频数据对算法进行验证。

5.1　实验性能指标

本文采用查全率和准确率测试模型来衡量算法性能。目前对关键帧提取算法并没有统一的性能评价标准，因此，以采用人工的方法进行关键帧提取作为评价标准，查全率和查准率为性能指标：

$$\text{查全率} = \frac{\text{正确提取数}}{\text{正确提取数} + \text{漏提取数}} \times 100\%$$

$$\text{查准率} = \frac{\text{正确提取数}}{\text{正确提取数} + \text{误提取数}} \times 100\%$$

视频序列的详细信息如表 4 所示。

表 4　视频序列详细信息

视频片段	帧数	镜头数	关键帧数
电影片段	42 960	195	275
动画片段	2 655	17	91
风景片段	44 640	345	413

5.2　实验结果

本文实验依次采用像素域的像素差法、直流系数法和本文的关键帧提取算法对实验视频片段提取关键帧。电影片段、动画片片段及风景片段的提取结果分别如表 5~表 7 所示。实验结果证明，像素差法对动画类视频的提取效果优于风景类和电影类片段，直流系数法对风景类视频的提取效果优于动画类和电影类视频，本文算法对于 3 类视频的检索效果都能很好地符合人的视觉特性，且算法提取效果稳定。

表 5　电影片段提取结果

提取算法	正确提取数	误提取数	漏提取数	查全率/(%)	查准率/(%)
像素差法	219	84	19	79.6	72.2
直流系数法	218	104	44	79.3	67.7
本文算法	253	107	22	92.0	69.6

表 6　动画片片段提取结果

提取算法	正确提取数	误提取数	漏提取数	查全率/(%)	查准率/(%)
像素差法	85	51	6	93.4	76.2
直流系数法	71	97	20	78.0	60.7
本文算法	83	37	8	91.2	69.1

表 7　风景片段提取结果

提取算法	正确提取数	误提取数	漏提取数	查全率/(%)	查准率/(%)
像素差法	352	221	61	85.4	61.4
直流系数法	335	188	76	81.1	64.0
本文算法	368	219	45	89.1	62.7

6　结束语

目前，将粗糙集理论应用于视频检索系统的研究成果还不多见。为此，本文提出一种基于粗糙集的压缩域关键帧提取算法。该算法通过构造 RS 理论的信息系统，并将处理含糊与不确定问题的 RS 经典算法应用于视频检索系统的关键帧提取中。将本文算法与直流系数法、像素差法对相同视频片断进行关键帧提取实验，对比结果证明，本文算法适用于不同类型的视频，且提取效果较好。

参考文献

[1] Horswiil H, Burgess J K. VIRS: A Content-based Video Indexing and Retrieval Systems[D]. Tasmania, Australia: University of Tasmania, 2001.

[2] 丁洪丽，陈怀新. 基于镜头内容变化率的关键帧提取算法[J]. 计算机工程, 2009, 35(13): 225-227.

[3] Wolf W. Key Frame Selection by Motion Analysis[C]//Proc. of IEEE Conference on Acoustics, Speech, and Signal Processing. [S. l.]: IEEE Press, 1996: 1228-1231.

[4] Zhuang Yueting, Rui Yong, Huang T S, et a1. Adaptive Key Frame Extraction Using Unsupervised Clustering[C]//Proc. of 1998 International Conference on Image Processing. Urbana, USA: [s. n.], 1998: 866-870.

[5] Hanjalie A, Zhang Hongjiang. An Integrated Scheme for Automated Video Abstraction Based on Unsupervised Cluster——Validity Analysis[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(8): 1280-1289.

[6] Pawlak Z. Rough Set[J]. International Journal of Computer and Information Science, 1982, 11(5): 341-356.

编辑　张　帆

(上接第 264 页)

需求，本文将单应急点资源调度模型推广到多应急点的情形，并给出模型的有效解法，解决了多应急点应急时间最早、参与应急点及应急线路最少的调度方案求取问题。

与已有方法比较，本文模型更便于利用计算机求解，对应急管理实际的适应面更广，仿真实例表明了模型的合理性和算法的有效性。当问题规模较大时，模型的计算机求解程序运行时间较长，这一方面应对实际问题进行整合、分解以缩小规模；另一方面有待于算法的改进，这也是下一步的研究工作。

参考文献

[1] 刘春林，盛昭翰，何建敏. 基于连续消耗应急系统的多出救点选择问题[J]. 管理工程学报, 1999, 13(3): 12-16.

[2] 戴更新，达庆利. 多资源组合应急调度问题的研究[J]. 系统工程的理论与实践, 2000, 20(9): 52-55.

[3] 刘春林，何建敏，施建军. 一类应急物资调度问题的优化模型研究[J]. 中国管理科学, 2001, 9(3): 29-36.

[4] 杨继君，许维胜，冯云生，等. 基于多模式分层网络的应急资源调度模型[J]. 计算机工程, 2009, 35(10): 21-24.

[5] ReVelle C S, Eiselt H A. Location Analysis: A Synthesis and Survey[J]. European Journal of Operational Research, 2005, 165(1): 1-19.

[6] Tavakoli A, Lightner C. Implementing a Mathematical Model for Locating EMS Vehicles in Fayetteville, NC[J]. Computers & Operations Research, 2004, 31(9): 1549-1563.

[7] Gong Qiang, Batta R. Allocation and Reallocation of Ambulances to Casualty Clusters in a Disaster Relief Operation[J]. IIE Transactions, 2007, 39(1): 27-39.

编辑　金胡考