Features Extraction and Selection Based on Rough Set and SVM in Abrupt Shot Detection

Yu Wu   Jialiang Han   Hongbo Li   Man Li

Institute of Artificial Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China

Abstract

Rough Set based reduction algorithm is adopted as a method for feature selection in abrupt shot boundary detection. First, some information of macro blocks is extracted in P frame of MPEG video sequence. Then the information of the motion-activity, the type of macro blocks and the motion-distribution are obtained by analyzing MPEG compressed-domain. Combined these information with the difference of pixel and the difference of histogram the abrupt shot detection can be achieved. The simulation experimental results show that the detection model combined Rough Set with SVM is effective in features selection. Some useful features for abrupt shot detection are discovered.

Keywords: Rough Set, Motion information, Abrupt shot, SVM

1. Introduction

The extensive amount of video coverage today, generates difficulties in identifying and selecting desired information. Obviously, traditional retrieval based on text can not meet these demands, so content-based video retrieval has been proposed as a solution to address this problem [1]. Shot-Boundary detection technique is the foundation of video analysis and content-based retrieval which impact its accuracy. A lot of research works have been done in this field recent years, but there are few research works on video features selection especially in shot boundary detection.

An assumption often made is that the content should remain nearly the same from one frame to the next within one camera shot. So, in general, shot boundaries can be detected by employing a difference metric to measure the change between two consecutive frames. A shot boundary is declared if the difference exceeds a certain threshold. Pixel- or block-based temporal image difference [2]-[3], or difference of gray and color histograms [4]-[5] were supposed to be the measure metric. Histograms are robust to object motion. And they are easy to compute. So they have been widely used in shot-based video analysis. And several authors claimed that this measure could achieve good trade-off between accuracy and speed. Unfortunately there is a problem of segmenting the film into a sequence of shots based on difference of histograms when illumination varies.

Video parsing on MPEG compressed data has been reported by making use of DCT blocks and motion vectors information [4]. A count of non-zero motion vectors was used to detect scene discontinuity. However, this method failed to handle special effects. Meng et al. [6] used the variance of DC coefficients in I and P frames and motion vectors information to characterize scene changes. Sethi and Pate1 [5] used only the DC coefficients of I frame to perform hypothesis testing via luminance histogram. It was assumed that distance between two I frames was fixed and small. The exact location of abrupt changes cannot be located with this method. Liu et al. made use of only information in P and B frames to detect shot boundary changes.

The above related works show that different researchers took different features as input for shot boundary detection. Based on these features, few of them took feature selection as a module of their video retrieval system. They did not pay much attention to finding and selecting important and acoustic features in develop shot boundary detection system.

Video frequency data has the property of non-constitutive and the complexity, including information about text, picture, and sound. At the meantime, the features extracted from the video frequency data are not precise, whereas we use Rough Set in our abrupt shot detection. Rough Set theory was proposed by Z. Pawlak in 1982 as a powerful mathematical analysis tool to process incomplete data and inaccurate knowledge [7]. Without any mathematical description of attributes and features of detected objects in advance, it determines the knowledge reduction and educes decision rules via indiscernibility relations and classes directly from given knowledge classification. So far, Combined Rough Set theory with content based video retrieval, we have achieved a series of research results: proposed a partition algorithm for huge data set based on Rough Set [9]; developed a motion

information based video pre classification retrieval system using Rough classification [10]-[11]; combined Rough Set with video key-frame extraction and got an acceptable result[12]. In the aspect of shot-boundary detection, we achieved an algorithm referenced to [15] and improved its performance [10] which could effectively detect abrupt shot, but it was still a bit subject to light and object movements. In this paper, an approach to shot boundary detection using Rough Set is introduced which comprehensively utilizes the motion features of P frame in MPEG video sequence, the difference of pixel and the difference of histogram.

Since there are uncertainty and inconsistency in the classification of shot-boundary detection, the rough set method is adopted to construct a better shot type. In our former work, we used Rough Set as an approach for shot-boundary detection .We found that experiment result was not stable when Rough Set was taken as feature selection and classifier. Usually, a higher recognition rate can be achieved when Rough set is taken as feature selection and SVM as classifier. In this paper, we use the method of Rough Set and SVM for shot-boundary detection from video sequence images separately. That is, Rough Set is taken as a module of feature selection and SVM is taken as classifier in shot-boundary detection. Through the experiment we find some useful video features for shot-boundary detection.

After the overall introduction, the rest of this paper is organized as follows. In section 2, some related theory and work are introduced. In section 3, the framework of a shot-boundary detection system (CBVRSYSTEM) is proposed. In section 4, Simulation experiments are done by means of CBVRSYSTEM. Finally, conclusion and future works are discussed in section 5.

2. Introduction to Rough Set Theory and MPEG -2 Standards

2.1. MPEG-2 Standards

Many storage media are organized with MPEG-2 format. Two coding techniques, prediction and interpolation, are adopted in the MPEG-2 video compression algorithm in order to satisfy high compression ratio and random access. There are three specific types of frames in MPEG-2, I frame (intra-coded frames), P frames (predicatively coded frames) and B frames (bidirectional predicatively coded frames). I frames don't need reference frames when coded and their compression ratio is the lowest. I frames are random access points and reference frames for other types of frames. P frames are predicted by use of only one previous I or P

frames. Meanwhile they are reference frames for subsequent frames. Because P frames utilize the temporal redundancy of video, their compression ratio is high. B frames are predicted referring to at most two frames (I or P frames) and can't be used as reference frames. Their compression ratio is the highest because of their bidirectional prediction with motion compensation.

In the MPEG-2 standard, there is a data structure called GOP (group of pictures), which is shown in Fig. 1. The count of frames in GOP is a constant, and the first frame in GOP must be the I frame. A frame consists of certain slices, which are composed of arbitrary numbers of macro blocks. Macro blocks appear in raster scanning order in the video stream. A macro block, consisting luminance block or chroma block, is the unit of prediction with motion compensation. The block is the unit of DCT (discrete cosine transforms). Frames are predicted with motion compensation by use of the previous decoded frames (I or P frames), then they are reconstructed via combining the prediction with the coefficient data from the IDCT (inverse discrete cosine transform) outputs.
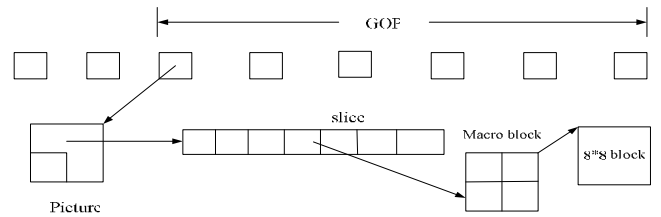

Fig. 1: Data Hierarchy of MPEG Video.

2.2. Introduction to Rough Set Theory

2.2.1 Basic Concepts

The expression of knowledge in Rough Set is generally formed as an information system. It is defined as:

$$K = (U, A, V, P), \tag{1}$$

where U is a finite set of objects and A= $C \cup D$ is a finite set of attributes, C is the condition attribute set and D= {d} is the decision attribute set. With every attribute $a \in A$, set of values V_a is associated. Each attribute determines a function:

$$f_a : U \to V_a. \tag{2}$$

Actually, a relation database is an information system; each row expresses the object information by relation attribute expression. Usually, a decision attribute D is

Cmvage	ModuleNum	Rog	Weight	RateMov	RateHChg	RateLChg	Rate	CmvShot	historm	pixsel	Type
0.919	1	0.227	1.095	0.261	0.196	0.542	0.081	1.263	1.450	0.148	0
5.458	3	0.270	9.137	0.159	0.417	0.423	0.145	5.352	0.433	6.850	0
2.069	2	0.428	14.80	0.054	0.801	0.145	0.197	1.316	6.334	8.887	1
4.624	1	0.159	0.998	0.375	0.304	0.321	0.156	4.929	0.972	3.789	0
6.404	4	0.061	5.027	0.156	0.625	0.219	0.572	1.865	17.14	0.922	0
4.624	1	0.159	0.998	0.375	0.304	0.321	0.156	4.929	0.973	3.789	0
3.766	1	0.286	12.23	0.091	0.727	0.182	0.127	4.928	1.632	7.199	1
7.672	4	0.163	6.693	0.145	0.690	0.165	0.169	5.029	3.239	30.905	1

Table 1: An example of decision table.

selected though the whole set of attribute A. Others is designed as condition attributes, defined as C.

2.2.2 Decision Table

A decision table is a formal representation of a data set to be analyzed and is a two-dimensional decision information table, in which each row denotes one example of U, each column denotes one attribute of A, and each element denotes the value of the information function. Table 1 is an example of decision table.

2.2.3 Attribute Reduction Algorithm for Feature Selection

There are many algorithms to reduce values, such as the inductive reduction algorithm, the heuristic reduction algorithm, the reduction algorithm based on decision matrices and so on.

Traditional techniques make use of feature merits based on the information theories, statistics correlation between each feature and the class, or the significant values based on Rough Set theory. All these measures only consider the single feature's effect on the class distinguish ability. However, in general, one feature does not distinguish classes by itself. Therefore, it is desirable to obtain the feature's correlation to the class in the context of other features. The reduction algorithm based on decision matrix [8] is adopted in the paper which is described as follow.

If P and Q are equivalence classes in U, the positive region of Q to P $POSp(Q)$ is defined as:

$$P\ POSp\ (Q) = \bigcup_{x \in U/Q} P_-(X).$$

For information table RED after reduction of attributes, let $X_i^+ (i = 1, 2, ..., \gamma), X_j^- (j = 1, 2, ..., \rho)$ denote the equivalence classes of the relation

$$R * (RED), X_i^+ \subseteq POS_{RED}^\beta (Y), X_J^- \subseteq NEG_{RED}^\beta (Y),$$

then the decision matrix $M = (M_{ij})_{\gamma\beta}$ is defined as:

$$M_{ij} = \{(a, f(X_i^+, a)) : a \in RED, f(X_i^+, a) \neq f(X_j^-, a)\},$$

M_{ij} contains all attribute pairs with different values in the equivalence class X_i^+ and X_j^-. Given an equivalence class X_i^+, if each element of M_{ij} is taken as a Boolean expression, the decision rule set can be expressed as:

$$B_i = \wedge_j (\vee M_{ij}).$$

It's seen that basic implication of B_i is actually the maximal generalized rule of the equivalence class X_i^+, which belongs to the positive region $POS_{RED}^\beta (Y)$. As a result, by finding out all basic implication of the decision function $B_i (i = 1, 2, ..., \gamma)$, all maximal generalized rules of positive region $POS_{RED}^\beta (Y)$ are computed.

3　Shot-Boundary Detection System

Shot boundary detection pre-classification model is shown in Fig.2, which is composed of such modules as feature extraction, feature selection, classification model.

Our system is developed based on MPEG-2 encoder [19]. So, video sequences based on MPEG-2 are taken into consideration as input.
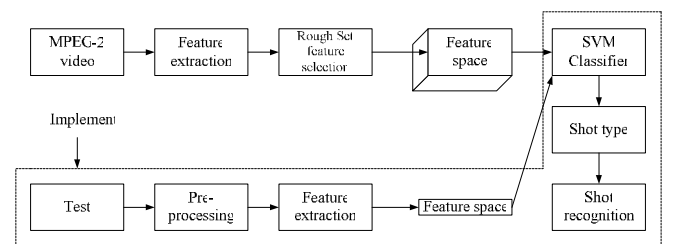

Fig. 2: Shot-Boundary Detection pre-classification model.

3.1 Feature Extraction

3.1.1 Motion Information from Compressed MPEG Stream

As for MPEG, each video sequence is composed of a series

of Group of Picture (GOP). While P can best represent movements of the whole GOP. Here, we will use the P frame information in compression region.

The macro block types are denoted as the field *macroblock_type* in MPEG stream. However, the field is just used in video decoding, but hardly in video retrieval. Furthermore, they can be divided into three types, which are low change (L), middle change (M) and high change (H). So the change ranges of frames are shown in the field *macroblock_type* in a certain degree. Via the field *macroblock_type*, ratios of macro block types are defined as follows:

$$RateH = Hcount / Tount , \qquad (3)$$
$$RateM = MCount / TCount , \qquad (4)$$
$$RateL = LCount / TCount , \qquad (5)$$

where *Tcount* is the whole count of a P frame, and the other *Xcount* are the count of the macro block type X.

Motion activity is a concept related with the activity of a video sequence or a movement pace. In this paper, we refer to literature [13] to define the motion activity as follow:

$$R_{xy}(i, j) = \sqrt{x(i, j)^2 + y(i, j)^2} . \qquad (6)$$

In the coordinate system for a P frame, $x(i, j)$ and $y(i, j)$ are the coordinates of every macro block in x and y directions respectively. We define: the energy of the macro block. $x(i, j)$ is horizontal motion vector of macro block and $y(i, j)$ is the vertical one. The average motion activity for every frame calculated as follows:

$$C_{mvavg}\left(R_{xy}(i, j)\right) = \frac{1}{MN}\sum_{i=0}^{M}\sum_{j=0}^{N} R_{xy}(i, j). \qquad (7)$$

We still need to consider the condition of motion activity of the shot with P frame belongs to. We define:

$$Rate_{mov} = \frac{C_{mvavg}(i)}{\sum_{i=1}^{n} C_{mvavg}(i)} , \qquad (8)$$

$$C_{mvShot} = \frac{\sum_{i=1}^{n} C_{mvavg}(i)}{n} . \qquad (9)$$

$Rate_{mov}$ denotes the proportion of the motion activity in the shot total activity in the P frame, while C_{mvShot} denotes the average motion activity of the shot where contains the P frame, and *n* in two formula is the total number of P frames in the shot.

These three attributes related with motion activity indicate the total activities of a video fragment and a few P frame activity in each shot. In the approach of which based on motion activity intensity, the shot-boundary detection must be selected depend on these attributes.

Motion spatial distribution must be paid attention to

which describes the distribution of the activity in an image. Two measurements, *Rog* and *Weight*, are introduced to describe the special distribution of each image. Referring to the literature [14], we define:

$$m_{pq} = \sum_{i=-I/2}^{I/2} \sum_{j=-J/2}^{J/2} i^p j^q f(i, j) , f(i, j) = \begin{cases} 0 & E_{ij} = 0 \\ 1 & E_{ij} \neq 0 \end{cases} , \quad (10)$$

$$Weigh_x = \frac{m_{10}}{m_{00}}, Weigh_y = \frac{m_{01}}{m_{00}}, Weigh = \sqrt{Weigh_x^2 + Weigh_y^2} , \quad (11)$$

$$Rog_x = \sqrt{\frac{m_{10}}{m_{00}}}, Rog_y = \sqrt{\frac{m_{01}}{m_{00}}}, Rog = \sqrt{Rog_x^2 + Rog_y^2} . \quad (12)$$

3.1.2 Pixel and Color Histogram

Pixel and color histogram are two important clues for the description of video content. For a video frame, the color information summarizes the appearance of the embedded objects well, not considering the location of the pixel. So, it is possible the histograms of two frames are similar, but the contents are completely different. The pixel information summarizes the distribution of the embedded objects well. But, it is easily subjected to the voice and motion. In this paper, we combine two features merit. They can complement each other nicely.
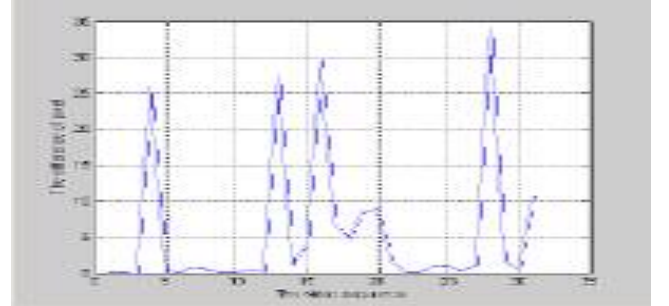

Fig. 3: Pixel difference distribution.

Fig. 4: A series of video sequence.

The difference of pixel is easy to get. $D(k, k + 1)$ is

defined as the difference between frame K and frame K+1:

$$D(k, k+1) = \frac{1}{XY} \sum_{x=1}^{X} \sum_{y=1}^{Y} \left| L_k(x, y) - L_{k+1}(x, y) \right|, \qquad (13)$$

where X, Y is the width and height of the picture; $L_k(x, y)$ and $L_{k+1}(x, y)$ is the value of the frame K and frame K+1 respectively. The shot-boundary will appear obviously changed based on pixel. Fig.3 and Fig.4 show the pixel difference of a video sequence.

As the three color components of the HSV color space are independent, we first calculate the histogram differences of consecutive video frames on each color component respectively. Fig.5 shows the distribution of the HSV histogram between shot-boundary. And then, we use the maximum histogram difference as the feature. The HSV histogram difference HD_i is calculated as follows:

$$HD_i = \max(H_i, S_i, V_i), \qquad (14)$$

$$H_i = \sum_{b=1}^{B} \left| h_i^V(b) - h_{i+k}^V(b) \right|, \qquad (15)$$

$$H_i = \sum_{b=1}^{B} \left| h_i^H(b) - h_{i+k}^H(b) \right|, \qquad (16)$$

$$H_i = \sum_{b=1}^{B} \left| h_i^S(b) - h_{i+k}^S(b) \right|. \qquad (17)$$

H_i, S_i and V_i are the histogram differences of H, S, and V component, respectively. h_i^H, h_i^S and h_i^S are the 64 bin normalized histograms of H, S and V, respectively.

Fig. 5: The distribution of HSV histogram.

3.2 Feature Selection

Feature selection module is the key part of our shot-boundary detection system. The purpose of this module is to select valuable features for shot-boundary detection from all features. In this module, features for shot-boundary detection are decreased. Thus, the complexity and cost for the following classification procedure are reduced. It could improve the efficiency of the whole system.

In our system, features selection module is based on the Rough Set theory. The attribute reduction algorithm adopted to select features is introduced in section 2.

3.3 Classification Module

SVM are taken as the classifier in our shot-boundary detection system. There are two cases of classification problems: linear separable and linear non-separable. In practice it is necessary to employ the non-separable approach. In statistical learning theory [17] [18], we bound the difference between the expected risk $R(\alpha)$, and the empirical risk $R_{emp}(\alpha)$, when both training and tests sets are assumed to be generated from the same underlying probability distribution $P(x, y)$. The empirical risk is calculated by:

$$R_{emp}(\alpha) = \frac{1}{2l} \sum_{i=1}^{l} \left| y_i - f(x_i, \alpha) \right|, \qquad (18)$$

where l is the size of the training set, α is the model parameter and $f(x_i, \alpha)$ is the classifier output for a training vector x having a corresponding label $y_i \in \{-1, 1\}$. The risk for an unseen test vector x is

$$R(\alpha) = \int \frac{1}{2} |y - f(x, \alpha) dP(x, y)|. \qquad (19)$$

4 Experiments and Analysis

In our experiments, the RIDAS system is adopted as a data mining platform, which is developed by Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications in China [16]. The system integrates almost 30 classical algorithms regarding rough set theory. The experiment data includes 938 frames which are randomly divided by 5 and chosen 376 as training data.

As above introduction, we choose 11 features as the set of shot-boundary condition attributes:

C= $\{C_{mvavg}$、 $Rate$、 C_{mvShot}、 $ModuleNum$、 $RateMov$、 $RateHChg$、 $RateLChg$、 Rog、 $Weight$、 $Pixel$、 $Histom\}$.

From Table2, we can see that the results are not stable using Rough Set as classifier. Using these features, we form the decision table as shown in Table1. Let $A = C \cup D$ denotes attributes set of P frame, C is the set of total condition attributes of P frame as listed above, and D is the decision attribute of the P frame, D = { type }. Then with the

Fig. 6: Comparison of the proposed shot-boundary detection method (the 2nd row) versus the algorithm [15] (the 1st row).

reduction algorithm of [8], 5 features are selected:

C= $\{C_{mShot}$、$ModuleNum$、$RateMov$、$RateHChg$、$Historm\}$.

In these 5 features selected by Rough-Set-based attribute reduction algorithm, there are 4 compressed features and 1 non-compressed feature. These 5 features are taken as the input of the classifier of SVM. At the mean time, we define: if the frame is a hard cut frame, type amount 1; if the frame is a gradual transition frame, type amount 2; otherwise type amount 0. In this way, all the frames are classified into 3 types: hard cut frame, gradual transition frame and common frame.

Then, we use Rough Set as feature selection and SVM as classifier. The result is shown in table 3.

From Table2, Table3 and Table4, we can conclude that Rough Set can remove some unimportant condition attributes, while SVM has a good classification result. At the meanwhile, the video features we extract and select can effectively distinguish shot-boundary.

The proposed method is tested on three digital videos. The contents of the experimental data are advertisement, movie fragments and dace fragments. The results are shown in Table5. We calculate Precision and Recall by:

Precision= (Correct boundaries)/
(Correct boundaries + Error boundaries)
Recall= (Correct boundaries)/
(Correct boundaries + Missed boundaries)

A comparative experiment using color algorithm, pixel algorithm and the algorithm of [15] is done as is shown in Table6. In order to validate the advantage of our method based on Rough Set+SVM in abrupt shot detection, we go on to conduct a contrastive experiment The test video, a

fragment of movie lasts for 80 seconds and contains 2684 frames. The performance of our method is compared with the algorithm [15] and the comparison is shown in Fig.6, we just choose some represent shot-boundary frames.

The algorithm of [15] is based on dual (one big and one small) sliding windows. From the 1st row of Fig.6, we can see that the shot-boundary frames extracted by it are still subject to light and object movements. The two adjacent frames are considered as different shot-boundary. In fact, they belong to the same shot. On the contrary, we may see that our method behaves well in overcoming those defects.

	Recognition count	Percentage (%)
Right	311	82.7%
Wrong	49	13.0%
Unknown	16	4.7%

Table 2: Distribution of recognition ratio.

SVM					
	1	2	3	4	5
Total count	187	188	187	188	188
Recognition count	174	173	172	166	177
Recognition ratio	93.05%	92.02%	91.98%	88.30%	94.15%
Average ratio	91.90%				

Table 3: Experiment results using SVM.

SVM+Rough Set					
	1	2	3	4	5
Total count	187	187	188	188	188
Recognition count	175	177	174	172	171
Recognition ratio	93.58%	94.15%	92.55%	90.96%	91.49%
Average ratio	92.55%				

Table 4: Experiment results using SVM+Rough Set.

	Total shots	Frame	Abrupt	Error	Miss	Precision	Recall
Movie	39	443	38	3	3		
AD	26	269	24	1	1	94.03%	93.3%
Dance	67	1208	64	5	4		

Table 5: Experimental results by implementing proposed method.

Abrupt Shot Transition		
	Precision	Recall
Color Histogram	70.68%	82.79%
Pixel	71.80%	81.94%
Double Slide Window	93.24%	91.83%
Our Method	94.03%	93.33%

Table 6: Comparative results.

5 Conclusions and future work

In this paper, based on Rough Set theory, an approach of shot-boundary detection is proposed. Based on Rough Set reduction, some important features for shot-boundary detection are discovered according to our simulation experiment results. Depending on these features, an average recognition rate of 92.55% is achieved using SVM + Rough Set. That is, the video features we select can effectively use to distinguish shot-boundary especially in abrupt shot-boundary detection. In the future, we are going to study further on the gradual transition detection, application to manage the video data, and explore its applications, such as the system of video retrieval. Effective reduction algorithm for video features selection for video retrieval system will also studied.

Acknowledgement

References

[1] H. Zhang, J.Y.A.Wang, and Y.Altunbasak, An integrated system for content-based video retrieval and browsing, *Pattern Recognition*, 30:634-648, 1997.

[2] K. Otsccji, Y. Tonomura and Y. Ohba, Video browsing using brightness data, *Proceedings of the SPIE on Visual Communications and Image Processing*, 1606: 980-989, 1991.

[3] A. Nagasaka and Y. Tanaka, An automatic video indexing and video search for object appearances, *Proceeding of the 2nd Working Conference on Visual Database Systems*, pp. 119-133, 1991.

[4] H. Zhang, A.Kankanhalli and S. Smoliar, Automatic partitioning of full motion video, *Multimedia System*, l1:10-28, 1993.

[5] I. K. Sethi and N. Patel, A Statistical Approach to Scene Change Detection, *Proceeding of the SPIE on Storage and Retrieval for Image and Video Databases*, 2420: 329-338, 1995.

[6] H. J. Zhang, et al, Video parsing, retrieval and browsing: An integrated and Content-Based Solution, *In: Proc. of ACM Multimedia 95 San Francisco*, pp. 15-24, 1995.

[7] Z. Pawlak, Rough Sets, *International Journal of Computer and Information Sciences*, 11:314-356, 1982.

[8] S.H.Nguyen, and A.Skowron, Quantization of Real Value Attributes-Rough Set and Boolean Reasoning Approach, *Proc of the second Joint Conference on Information Sciences*, pp. 34-37, 1995.

[9] Z.R. Qin, Y. Wu and G.Y. Wang, A Partition Algorithm for Huge Data Sets Based on Rough Set, *Mode Identification and Artificial Intelligence*, pp. 249-256, 2006.

[10] Z.Yuan, Y. Wu, G.Y. Wang and J.B. Li, Motion-Information-Based Video Retrieval System Using Rough Pre-classification, *Lecture Notes in Computer Science, Book Transactions on Rough Sets V*, 4100: 306-333, 2006.

[11] Z.Yuan, Y. Wu, G.Y. Wang and J.B. Li, A Global-Motion Analysis Method via Rough –Set-Based Video Pre-classification, *10th International Conference, RSFDGRC,* pp. 323-332, 2005.

[12] T. Wang, Y. Wu and L. Chen, An Approach to Video Key-frame Extraction Based on Rough Set, *Multimedia and Ubiquitous Engineering*, pp. 590-596, 2007.

[13] T. L. Yu and S. J. Zhang. Video Retrieval Based on the Global Motion Information, http://cs.cqupt.edu.cn/videoretrieval

[14] S. Jeannin, A. Divakaran., MPEG-7 visual motion descriptors, *IEEE Transactions on Circuits and Systems for Video Technology*, 11: 720-724, 2001.

[15] H. B. Lu, S. J. Zhang, An efficient algorithm for detecting abrupt scene change, *Journal of Image and Graphics*, 10:805-810, 1999,

[16] G. Y. Wang, Z. Zheng, Y. Zhang, RIDAS-A Rough Set Based Intelligent Data Analysis System, *Proceedings of the First Int. Conf. on Machine Learning and Cybernetics*, pp. 646-649, 2002.

[17] C.J.C.Burges, A tutorial on support vector machines for pattern recognition, *Data Mining, and Knowledge Discovery*, 2(2): 121-167, 1998.

[18] A.B.Hur, D.Hom, H.T.Seigelmann, V.Vapnik, A Support Vector Method for Clustering, *Advances in Neural Information Processing Systems*, 13: 367-373, 2001

[19] http://www.mpeg.org/MPEG/MSSG/#source.