Effective Video Analysis Preprocessing Algorithm Based on Rough Sets in Compressed Domain

Department of Computer Engineering, Lanzhou Polytechnic College, Lanzhou 730050, China
College of Electrical Engineering and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China
College of Mathematics and Information Science, Northwest Normal University, Lanzhou 730070, China

Abstract: This study is aimed to overcome the characteristic of redundant data for video analysis, the study proposes a video analysis preprocessing algorithm based on Rough Sets (RS). Firstly, the representative data of video sequences is extracted. Secondly, the Information Table is constructed by extracted representative data. Finally, the Core of Information Table is achieved by making use of the attributes reduction theory of RS. As present experimental results indicate, the algorithm can get effective and scientific data to complete video analysis such as key frame extraction and shot segment. Compared to existing techniques, the proposed algorithm enjoys following advantages. Only a subset of frames need to be considered during video analysis. The limitations of requirements for a huge amounts of memory and CPU resource is overcome.

Key words: Attributes reduction, core, information system

INTRODUCTION

Advances in computer, telecommunications and internet technologies have brought huge amounts of multimedia information to a rapidly growing. More and more digital videos are made available over the internet or other storage device. Since of the characteristics of redundancy information in video data, many video analysis technologies become less available, especially in real-time system. So, video data preprocessing become an important and imperative work before video analysis.

There have been many considerable work reported on video preprocessing. All in all, the existing typical techniques can be categorised into following types: Video shot segmentation based technologies (Cooper et al., 2007; Lee et al., 2006), key frame extraction based technologies (Shaping and Xinggang, 2005; Kin-Wai Sze et al., 2005) and video summarization based technologies (Wen-Nung and Chun-Ming, 2004; Yu-Fei and Hong-Jiang, 2002). A challenging and difficult problem of above technologies is there is no efficient and scientific algorithm to processing redundant data either in uncompressed domain or compressed domain.

Rough Set (RS) is a novel and powerful tool for data analysis. It has successfully been used in many application domains, such as machine learning, expert system and pattern classification (Wojciech et al., 2008; Neil and Quang, 2008). The main advantage of rough sets is that it does not need any preliminary or additional information about data, like probability in statistics, or basic probability assignment in Dempster-Shafer theory and grade of membership or the value of possibility in fuzzy sets. Especially, the attributes reduction theory has widely used in many application domain.
(Changzhong et al., 2008; Yiyu and Yan, 2008). In this study, by using the advantage of RS theory, a RS based video analysis preprocessing algorithm is developed.

ROUGH SETS THEORY

Basic Concepts of Rough Sets

An Information System is a 4-tuple \( S = (U, A, V, F) \), where \( U \) non-empty finite set of objects, \( A \) is a non-empty finite set of attributes, \( V \) is the union of attributes domains.

Let \( U \neq \emptyset \) be a universe of discourse and \( X \) be a subset of \( U \). An equivalence relation, \( R \), classifies \( U \) into a set of subsets \( U / R = \{X_1, X_2, \ldots, X_n\} \) in which the following conditions are satisfied:

- \( X_i \subseteq U, X_i \neq \emptyset \) for any \( i \)
- \( X_i \cap X_j \neq \emptyset \) for any \( i, j \)
- \( \bigcup_{i=1, 2, \ldots, n} X_i = U \)

Any subset \( X_i \), which called a category, class or granule, represents an equivalence class of \( R \). A category in \( R \) containing an object \( x \in U \) is denoted by \([x]_R\). For a family of equivalence relations \( P \subset R \), an indiscernibility relation over \( P \) is denoted by \( \text{IND}(P) \) and is defined by Eq. 1:

\[
\text{IND}(P) = \bigcap_{R \in P} \text{IND}(R)
\]  

(1)

The set \( X \) can be divided according to the basic sets of \( R \), namely a lower approximation set and upper approximation set. Approximation is used to represent the roughness of the knowledge. Suppose a set \( X \subset U \) represents a vague concept, then the \( R \)-lower and \( R \)-upper approximations of \( X \) are defined by Eq. 2 and 3.

\[
\overline{R}X = \{x \in U : [x]_R \subseteq X\}
\]  

(2)

Equation 3 is the subset of \( X \) such that \( X \) belongs to \( X \) in \( R \), is the lower approximation of \( X \).

\[
\overline{\overline{R}}X = \{x \in U : [x]_R \cap X \neq \emptyset\}
\]  

(3)

Equation 3 is the subsets of all \( X \) that possibly belong to \( X \) in \( R \), thereby meaning that \( X \) may or may not belong to \( X \) in \( R \) and the upper approximation \( \overline{R} \) contains sets that are possibly included in \( X \). \( R \)-positive, \( R \)-negative and \( R \)-boundary regions of \( X \) are defined respectively by Eq. 4, 5 and 6.

\[
\text{POS}_R(X) = RX
\]  

(4)

\[
\text{NEG}_R(X) = U - RX
\]  

(5)

\[
\text{BNR}(X) = \overline{RX} - RX
\]  

(6)

The positive region includes all objects in \( U \) which can be classified into the equivalence classes determined by indiscernibility relation.

Attributes Reduction and Core

In RS theory, an Information Table is used for describing the object of universe, it consists of two dimension, each row is an object, and each column is an attribute. RS classifies the attributes into two
types according to their roles for Information Table: Core attributes and redundant attributes. Here, the minimum condition attribute set can be received, which is called reduction. One Information Table might have several different reductions simultaneously. The intersection of the reductions is the Core of the Information Table and the Core attribute are the important attribute that influences attribute classification.

A subset B of a set of attributes C is a reduction of C with respect to R if and only if:

- \( \text{POS}_B(R) = \text{POS}_C(R) \)
- \( \text{POS}_{B \setminus a}(R) \neq \text{POS}_C(R) \), for any \( a \in B \)

And the Core can be defined by Eq. 7:

\[
\text{CORE}_C(R) = \{ c \in C \mid \forall c \in C, \text{POS}_{C \setminus \{c\}}(R) \neq \text{POS}_C(R) \} \tag{7}
\]

**THE PROCEDURE OF DATA PREPROCESSING**

**Acquisition of Raw Data**

Raw video data come from two aspects, the first is come from uncompressed domain, i.e., the frame in original video sequences. In this case, each video frame can be represented by the three primary colors of RGB. The second is come from compressed domain. In this case, the data is come from I, P and B frame. To simplify processing model, we only consider the I frame, i.e., DCT coefficients and DC coefficients. The DCT coefficients and DC coefficients can be acquired easily by Eq. 8 and 9.

\[
\text{vr}(P(x,y)) \xrightarrow{\text{mirror}} \text{DCT coefficients} \tag{8}
\]

\[
\text{DCT coefficients} \xrightarrow{\text{mirror}} \text{DCT coefficients} \tag{9}
\]

Equation 8 shows the DCT coefficients can be extracted from I frame of each video sequences. And Eq. 9 shows that by eliminating the AC coefficients of each block, the DC coefficients can be extracted.

**The Construction of Preprocessing Model**

According to MPEG video compressed standard, DC coefficients represent the major visual information of video sequences. As such, DC coefficients can satisfy the requirements of video analysis sufficiently. We have got the DC coefficients of each frame, so we can construct an Information System with it. Each row is a DC coefficient, and each column is the frame. This process can be described as following Eq. 10.

\[
\text{DC} \xrightarrow{\text{mirror}} \text{Information Table } S = \{ U, A, V, f \} \tag{10}
\]

where, \( U \) is sets, denotes all the object of Information System, \( A \) is also a sets, denotes all attributes in Information System, \( V \) is the sets of attributes value, \( f \) is a function denotes the relations between objects and attributes.

By using above process, we can get Information System table as Table 1. In Table 1, each row represent the DC coefficients of block in various video frame, consequently, this table contained the first and most important information of video sequences.
Table 1: Information system table constructed using DC coefficients

<table>
<thead>
<tr>
<th>DC coefficient</th>
<th>Frame 1</th>
<th>Frame 2</th>
<th>Frame 3</th>
<th>Frame 4</th>
<th>Frame 5</th>
<th>Frame 6</th>
<th>Frame 7</th>
<th>Frame 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC1</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.27771</td>
<td>0.35355</td>
<td>0.41572</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC2</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.41573</td>
<td>0.35354</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC3</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35356</td>
<td>0.19134</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC4</td>
<td>0.35356</td>
<td>0.35356</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35351</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC5</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35356</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC6</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.35356</td>
<td>0.19134</td>
<td>0.35345</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC7</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC8</td>
<td>0.23761</td>
<td>0.23761</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
</tbody>
</table>

Table 2: The reduced information system

<table>
<thead>
<tr>
<th>DC coefficient</th>
<th>Frame 1</th>
<th>Frame 2</th>
<th>Frame 3</th>
<th>Frame 4</th>
<th>Frame 5</th>
<th>Frame 6</th>
<th>Frame 7</th>
<th>Frame 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC1</td>
<td>0.35355</td>
<td>0.27771</td>
<td>0.35355</td>
<td>0.41572</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC2</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.41573</td>
<td>0.35354</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC3</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35356</td>
<td>0.19134</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC4</td>
<td>0.35356</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35351</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC5</td>
<td>0.35355</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35345</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC6</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35353</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC7</td>
<td>0.27779</td>
<td>0.27779</td>
<td>0.35355</td>
<td>0.19134</td>
<td>0.35353</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
<tr>
<td>DC8</td>
<td>0.23761</td>
<td>0.23779</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
<td>0.35355</td>
</tr>
</tbody>
</table>

Reduced Sets and Core

The attributes in the Information Table can be divided into two types by their roles: Core attributes and redundant attributes. From Table 1, we can show that the frame 2 and frame 8 can be reduced, the reduced Information System is showed as Table 2. If we introduce a threshold, more attributes can be reduced.

By attributes reduction, we can get a succinct and informative Information Table, this Table 2 contain major visual information of video sequences, so it is sufficient to utilize the table to video analysis.

The Description of Algorithm

algorithm: Generate_core(GC)
input: Original Information Table
output: Reduced attributes sets
processing:
(1)
   for i = 1 to n / n is the number of Information Table recorder
   for j = 1 to m / m is the number of Information Table attributes
   if pm - pm/k = 1,2,...n
      mn = ∅
   else
      mn = Ai[f(x,a) + f(x,j,a)] i,j = 1,...n /m is the sets of all attributes that can discem x and xj
(2)
   for i = 1 to n
      for j = 1 to m
         if and [xjA\xjA] = ∅
            output Ai
   }
EXPERIMENTAL RESULTS

Various MPEG video sequences are selected to examine the performance of the proposed algorithm. Figure 1 illustrates some key frames extracted from shots using the proposed algorithm. Figure 2 shows some shots extracted from experimental video sequences using our proposed algorithm. We can see obviously that the data can complete the video analysis efficiently. Table 3 and 4 show the comparison results.

Fig. 1: Some key frames extracted from shots using proposed algorithm

Fig. 2: Shots extracted from experimental video sequences using our proposed algorithm

<table>
<thead>
<tr>
<th>Video sequences type</th>
<th>Before preprocessing</th>
<th>After preprocessing</th>
<th>Error detected</th>
<th>Results evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gym</td>
<td>112 Frames 7 Shots</td>
<td>31 Frames 6 Shots</td>
<td>1</td>
<td>Excellent</td>
</tr>
<tr>
<td>Animation</td>
<td>174 Frames 13 Shots</td>
<td>53 Frames 14 Shots</td>
<td>1</td>
<td>Excellent</td>
</tr>
<tr>
<td>Scenery</td>
<td>317 Frames 23 Shots</td>
<td>106 Frames 23 Shots</td>
<td>0</td>
<td>Good</td>
</tr>
<tr>
<td>Story</td>
<td>126 Frames 13 Shots</td>
<td>97 Frames 13 Shots</td>
<td>0</td>
<td>Excellent</td>
</tr>
<tr>
<td>News</td>
<td>153 Frames 14 Shots</td>
<td>101 Frames 17 Shots</td>
<td>3</td>
<td>Good</td>
</tr>
</tbody>
</table>

55
Table 4: Comparison result of key frame extraction by various video sequences

<table>
<thead>
<tr>
<th>Video sequences type</th>
<th>Before preprocessing</th>
<th>After preprocessing</th>
<th>Results evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frames</td>
<td>Key frame</td>
<td>Frames</td>
</tr>
<tr>
<td>Gym</td>
<td>112</td>
<td>26</td>
<td>31</td>
</tr>
<tr>
<td>Animation</td>
<td>174</td>
<td>47</td>
<td>53</td>
</tr>
<tr>
<td>Scenery</td>
<td>317</td>
<td>69</td>
<td>106</td>
</tr>
<tr>
<td>Story</td>
<td>126</td>
<td>50</td>
<td>97</td>
</tr>
<tr>
<td>News</td>
<td>153</td>
<td>81</td>
<td>101</td>
</tr>
</tbody>
</table>

With above comparison and evaluation, we can see obviously that preprocessed data can satisfy the video analysis for key frame extraction and shot segmentation, while the amounts of the frames and the data of each frame are reduced dramatically. As such, the efficiency is improved.

CONCLUSION

As the characteristics of video stream data, the first and most difficult task in video analysis is how to process redundant data. In this study, we propose a effective and efficient video analysis preprocessing algorithm based on rough sets in compressed domain. The algorithm introduce concepts of rough sets theory, the amount of data is reduced dramatically and scientifically. The experimental results prove that proposed algorithm can process video such as key frame extraction or shots segment without considering full data.

ACKNOWLEDGMENT

The study is supported by Gansu Natural Science Foundation of China under Grant (No. 3ZS051-A25-047).

REFERENCES


