Continuous Monotonic Decomposition (CMD) is a special type of Ascending Subgraph Decomposition [1]. Since the size of the graph is \( n(n+1)/2 \), CMD is closely related to the theory of triangular numbers. The discussion in this paper leads to the identification of a collection of sequence of natural numbers that are not triangular. They are the following:

1. \( 3m^2+6mk+2k^2 \) is not a triangular number \( \forall m \in \mathbb{N} \) when \( k \) is not a multiple of 3.
2. \( 3m^2+6mk+2k^2 \) is not a triangular number \( \forall m \in \mathbb{N} \) when \( k \) is not a multiple of 3.
3. \( 3m^2+6mk+2k^2 \) is a triangular number if and only if \( m=3n(n+1)/2 \) \( \forall n \in \mathbb{N} \).
4. \( (m^2+4m+4)n+4m+4 \) is a triangular number if and only if \( n=(m^2+4m+4)/2 \in \mathbb{N} \).
5. \( (m^2+4m+4)n+4m+4 \) is not a triangular number if \( m=2 \) or \( n \in \mathbb{N} \) with \( m \neq c \).

Hence, \( K_{n+1} \) does not accept CMD.

Example 8

Fig 8 \( K_{n+1} \) does not accept CMD.

V. CONCLUSION

Continuous Monotonic Decomposition (CMD) is a special type of Ascending Subgraph Decomposition [1]. Since the size of the graph is \( n(n+1)/2 \), CMD is closely related to the theory of triangular numbers. The discussion in this paper leads to the identification of a collection of sequence of natural numbers that are not triangular. They are the following:

1. \( 3m^2+6mk+2k^2 \) is not a triangular number \( \forall m \in \mathbb{N} \) when \( k \) is not a multiple of 3.
2. \( 3m^2+6mk+2k^2 \) is not a triangular number \( \forall m \in \mathbb{N} \) when \( k \) is not a multiple of 3.
3. \( 3m^2+6mk+2k^2 \) is a triangular number if and only if \( m=3n(n+1)/2 \) \( \forall n \in \mathbb{N} \).
4. \( (m^2+4m+4)n+4m+4 \) is a triangular number if and only if \( n=(m^2+4m+4)/2 \in \mathbb{N} \).
5. \( (m^2+4m+4)n+4m+4 \) is not a triangular number if \( m=2 \) or \( n \in \mathbb{N} \) with \( m \neq c \).

Hence, \( K_{n+1} \) does not accept CMD.

Theorem 7: A complete tripartite graph \( K_{n+1} \) accepts CMD, \( \forall n \in \mathbb{N} \).

Proof: We have, \( q(K_{n+1}) = \frac{(m^2+4m+4)n+4m+4}{2} = 6m+5, m \in \mathbb{N} \) \( \vdots \) (19)

We know that \( G \) accepts CMD \( [H_1,H_2,H_3] \) if and only if \( q(G)=n(n+1)/2 \) \( \forall n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

\( \vdots \) (19) should be of the form \( n(n+1)/2 \) for some \( n \in \mathbb{N} \).

The main aim of steganography is to hide information in the other wrap media so that other persons will not observe the existence of the information. This is a major distinction between this method and the other methods of secret exchange of information because, for example, in cryptography, the individuals perceive the information by considering the implied information but they will not be able to realize the information [1]. However, in steganography, the existence of the information in the sources will not be noticed at all. Most steganography jobs have been carried out on video, video clips, texts, music and sounds. For video stream usually being accessible in compressed form, steganography algorithms that are not applicable in compressed bit-stream would require complete or at least partial decompression [2]. This is an unnecessary saddle best avoided. If the requirement of strict compressed domain steganography is to be met, the steganography needs to be embedded in the compressed domain. Nowadays, there are large amount of video watermarking algorithms have been proposed. Some of them are applied for compressed video.

To be useful, a steganographic technique should not be easily detectable. If the existence of secret message can be detected with a probability higher than random guessing, the corresponding steganographic technique is considered to be invalid [3]. Similar to cryptography and steganography may suffer from the attack method (steganalysis). Much of the research work in the field of steganalysis has been carried out on images. One approach is based solely on the first order statistics and is applicable only to idempotent embedding. Steganography Terms

Carrier File: A file which has hidden information inside of it.
Steganalysis: The process of detecting hidden information inside a file. Stego-Medium: The medium in which the information is hidden.
II.haar Wavelet Transform

Wavelets are mathematical functions that were developed by scientists working in several different fields for the purpose of sorting data by frequency. Translated data can then be sorted at a resolution which matches its scale [8]. Studying data at different levels allows for the development of a more complete picture. Both small features and large features are discernible because they are studied separately. Unlike the discrete cosine transform, the wavelet transform is not Fourier-based and therefore wavelets do a better job of handling discontinuities in the data [7].

The Haar wavelet operates on data by calculating the sums and differences of adjacent elements. The Haar wavelet operates first on adjacent horizontal elements and then on adjacent vertical elements. The Haar transform is computed using:

\[ \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \]

One nice feature of the Haar wavelet transform is that the transform is equal to its inverse. As each transform is computed the energy in the data in relocated to the top left corner.

III. Methodology

As discussed earlier Digital video comprises a series of orthogonal bitmap digital images displayed in rapid succession at a constant rate. These images are called frames. The frames in a video are nothing but normal images with some extra information such as the index and other Meta data related to the video [10]. Each frame can be extracted individually from the video and can be converted into an image. This can then be treated as the cover object and the data can be embedded in it using one of the usual techniques used for data hiding in images. The stego-image then obtained can then be converted back into frames and arranged in sequence to obtain the stego-video. This video contains the embedded data which can be obtained by extracting each frame and extracting the data from it.

Our technique spreads the data evenly over the entire video instead of concentrating it in into one single frame, thereby making the detection of the data even more impossible. For example if there are 4 bits of data to be embedded into a video with more than four frames then each bit will be embedded in one of the four frames in the video. This approach gives a huge advantage in the aspect of increasing the imperceptibility of the data embedded into the video. Our methodology is quite simple and unsophisticated and hence very fast when compared to other existing techniques.

IV. The Coding Algorithm

In the usual steganography algorithms, information is hidden in the sequential pixels. Therefore anyone with the knowledge of the coding algorithm can extract the hidden information from the image. In this paper a new approach has been studied for selecting pixels according to a password. This password would enable us to select the pixels in a random manner.

In the usual steganography algorithms, if the size of the information is small in comparison with the size of image, the attacker can find the pattern of altered pixels and extract the hidden information. But in this method, information is embedded into a random order pixels in each block, and extracting the hidden information is difficult.

On the other hand if the size of the information is large, the algorithm reaches the end of image. For solving this problem, it has to return to the beginning of the image and hide information in an empty pixel (an empty pixel is defined as a pixel of original image that has no hidden data). This process needs a large amount of memory to remember all empty pixels, but in the mobile phones we have a limited amount of memory. After all, finding an empty pixel needs a lot of time in coding or decoding phases.

V. Proposed Algorithms For Embedding Scheme

Let the Cover video consist of N number of frames. Each of these frames are extracted from the video for the purpose of embedding the data in them.

1. Read the data to be embedded and convert it into binary form containing B bits.
2. Considering the amount of data to be embedded and the capacity of video calculate the value of k (no. of LSBs to be modified in each pixel), using the following formula.

\[ K = \frac{\text{MAXDATA} \times B}{N \times \text{Height} \times \text{Width}} \]

where MAXDATA = N*Height*Width, K should not exceed 4 in order to maintain satisfactory video quality.

3. Divide the B into N number of blocks of data. Let these blocks be called BLi.

4. Divide each block into groups of K bits.

5. Consider one frame at a time and embed the K groups of bits in blocks BLi into the Pixels in the corresponding frame. For example block BLi is embedded into the ith frame. For embedding OPAP (Optimal Pixel Adjustment Process) is used.

6. The frames with embedded data are again combined together to obtain the Stego-video.

VI. Proposed Algorithm for Decoding Scheme

1. Read the Stego- video with .avi extension.
2. Let the Stego video consist of N number of frames. Each of these frames are extracted from the video for the purpose of retrieving the data from them.
3. Calculate the value of K using the following formula.

\[ K = \frac{\text{MAXDATA} \times \text{CHARBITS}}{8} \]

where CHARBITS = (number of characters that have been embedded in the video)* 8.
4. Calculate the size of block of data to be extracted from each frame of the video using the formula

\[ \text{S} = \frac{\text{MAXDATA} \times \text{CHARBITS}}{K} \]

This method for hiding information in images can be used for secure communication, copyright protection, preventing undesirable changes in digital documents, protecting from unauthorized copying and other applications.
5. Extract the BLi bits of data from the ith frame using the OPAP technique.
6. The Extracted data which is a stream of bits should be grouped into blocks of 8 bits each and converted back into the character format in order to retrieve the data embedded in the video.

The proposed approach used to embed any number of files in audio or video. Precious human life could be saved in e-dia:nosis. These approaches concentrate on achieving higher compression ratio without sacrificing the quality of the image. Since processing power required in the mobile handset is limited, a new approach is developed with energy efficient, computing efficient and adaptive image compression and communication techniques. Performance analysis is made in terms of accuracy and computational time, which is a positive scope of this paper. Most of the computational burden is reduced in Haar wavelet transform.

Video (Covering medium)

(a) Before encoding (b) After encoding

Decrypted images

Table I Experimental Evaluation

<table>
<thead>
<tr>
<th>Carrier File Types</th>
<th>Compression Size</th>
<th>Compression Ratio</th>
<th>Embedding and Compression Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video</td>
<td>1.72MB</td>
<td>98.7%</td>
<td>5.12</td>
</tr>
<tr>
<td>Video</td>
<td>5.23MB</td>
<td>98.7%</td>
<td>9.26</td>
</tr>
<tr>
<td>Video</td>
<td>4.31MB</td>
<td>98.7%</td>
<td>9.15</td>
</tr>
<tr>
<td>Audio</td>
<td>2.65MB</td>
<td>99.3%</td>
<td>7.24</td>
</tr>
<tr>
<td>Audio</td>
<td>2.84MB</td>
<td>99.3%</td>
<td>7.29</td>
</tr>
<tr>
<td>Audio</td>
<td>7.24MB</td>
<td>99.3%</td>
<td>12.00</td>
</tr>
<tr>
<td>Text</td>
<td>3.76MB</td>
<td>99.3%</td>
<td>8.21</td>
</tr>
</tbody>
</table>

The probability that one can detect a steganano image is relatively low, due to the high volume of images exchanged between mobile phones and computers. The password is not stored in the steganano image; therefore it is difficult to detect the password. The decoding program fast enough to retrieve the images with few kilobytes of memory. This approach is used on mobile phones with no limitation for selecting the password.

VII. CONCLUSION

The proposed approach used to embed any number of files in audio or video. Precious human life could be saved in e-dia:nosis. These approaches concentrate on achieving higher compression ratio without sacrificing the quality of the image. Since processing power required in the mobile handset is limited, a new approach is developed with energy efficient, computing efficient and adaptive image compression and communication techniques. Performance analysis is made in terms of accuracy and computational time, which is a positive scope of this paper. Most of the computational burden is reduced in Haar wavelet transform.

REFERENCES


Abstract - Clustering of related or similar objects has long been regarded as a potentially useful technique for helping users to navigate an information space such as a document collection. But, the major challenge in document clustering is high dimensionality. Data mining and statistical techniques have been applied with some success to large set of documents to automatically produce meaningful subsets. Many clustering algorithms and techniques have been developed and implemented since the earliest days of computational information retrieval but as the sizes of document collections have grown, these techniques have not been scaled to large collections because of their computational overhead. Traditional document clustering is usually considered as an unsupervised learning. It cannot effectively group documents under user's need. To solve this problem, the proposed system concentrates on an interactive text clustering methodology, topic oriented probability based and semi supervised document clustering. It suggests interactive approach for document clustering, to facilitate human refinement of clustering outputs. The proposed system evaluates system efficiency by implementing and testing the clustering results with Dbscan and K-means clustering algorithms. Experiment shows that the proposed document clustering algorithm performs with an average efficiency of 94.4% for various document categories.

Keywords : Document Clustering, Text Documents, Word Frequency, Probability, Tokenization, Structural Filtering

I. INTRODUCTION

With the rapid development of Information technology, the number of electronic documents and digital content of documents exceed the capacity of manual control and management. People are increasingly required to handle wide ranges of information from multiple sources. As a result, document clustering techniques are implemented by organizations to manage their information and knowledge more effectively. Document clustering can be defined as