ISSN No. 0976-5697

Volume 2, No. 4, July-August 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Novel Method to Improve Video Compression for Video Surveillance Applications Using SPIHT Algorithm

C. Lakshmi Devasena*
Department of Software Systems,
Karpagam University,
Coimbatore ,India
devaradhe2007@gmail.com

M. Hemalatha
Department of Software Systems,
Karpagam University,
Coimbatore,India

Abstract: This paper describes an improved video compression method for video surveillance applications using SPHIT algorithm. In the proposed system SPHIT algorithm is used to compress the video to track the moving object. This proposed motion detection algorithm consists of motion detection of successive frames and background frame subtraction. The motion detection of sequential frames is used to detect the continuous moving objects between two successive frames. After the compression of the video, background subtraction of frame is used to detect moving objects in the current frame. The proposed method is developed and tested using some sample videos. The proposed system can be used in applications like Video Surveillance Applications, Motion pixel Estimation Applications, Video Surveillance, Weapon Storage Area and Atomic Research Area etc.

Keywords: Background Subtraction, Data Compression, Discrete Wavelet Transform, SPIHT Algorithm.

I. INTRODUCTION

There are abrupt requirements for automated surveillance systems in commercial applications, Law enforcement and military applications. Installing video cameras is inexpensive, but finding accessible human resources to observe the output is costly. Although surveillance cameras are customary in banks, parking lots, and stores video data currently is used only as a forensic tool, thus losing its primary advantage as a dynamic, real-time medium. What is needed is round the clock monitoring of surveillance video to alert security officers to a theft in progress, or to an apprehensive individual loitering in the parking lot, while there is still time to prevent the offense. Apart from security applications, video surveillance can be proposed to measure traffic flow, monitor pedestrian congestion in public spaces, detect accidents on highways, compile consumer demographics in shopping malls, etc.

The main aim of this research is to implement and verifies the SPIHT algorithm on video files, to compress the videos and to show how beneficial this will be while handling large video files. The key of this project lies in the initialization and update of the background image. The effectiveness of both will affect the accuracy of test results. Therefore, this paper uses an effective method to initialize the background, and update the background in real time. Sometimes problem may occur while applying the encoding, decoding, DWT, Background subtraction methods to the large size videos even though the video, that is the process may take much time for producing the output for the selected large sized video.

To avoid this time consumption a method called compression is being carried out in this project. Compression is a method which is used to reduce the total size of the video which has been selected. The two main use of compression technique are

A. Reduced memory
B. Use of Efficient bandwidth

Sometimes problem may occur while applying the compression, DWT, Background subtraction methods to the large size videos even though the videos are being compressed, that is the process may take much time for producing the output for the selected large sized video. While splitting the video into frames if the video size is large, more number of frames will be produced. While considering all separate frames for encoding and decoding process it will consume large amount of time.

If the bandwidth is very low, great attention must be paid to cut down the bit rate of compressed video bit stream. The Frame Subtraction is generally difficult to obtain a complete outline of moving object, liable to appear the empty phenomenon; as a result the detection of moving object is not accurate. In Optical Flow method, a large quantity of calculation, sensitivity to noise, poor anti-noise performance, makes it not suitable for real-time demanding occasions.

The proposed research combines

A. SPIHT algorithm for compressing videos in video surveillance with static camera
B. Motion Detection technique
C. Finding the compression ratio between the videos which is before compression and after compression.
D. Selection method based on the background subtraction, and update background on the basis of accurate detection of object, this method is effective to enhance the effect of moving object detection.

The other existing methods regarding this project are;

A. Frame subtraction method is through the difference between two consecutive images to determine the presence of moving objects.
B. Optical flow method is to calculate the image optical flow field, and do clustering processing according to the optical flow distribution characteristics of image.
C. Statistical method.

II. LITERATURE REVIEW

In 1997, the Defense Advanced Research Projects Agency (DARPA) Information Systems Office began a three-year program to develop Video Surveillance and Monitoring (VSAM) technology. The objective of the VSAM project was to develop automated video understanding technology for use in future urban and battlefield surveillance applications. Technology advances developed under this project enable a single human operator to monitor activities over a broad area using a distributed network of active video sensors. The sensor platforms are mainly autonomous, notifying the operator only of salient information as it occurs, and engaging the operator minimally to alter platform operations. A team composed of Carnegie Mellon University Robotics Institute and the Sarnoff Corporation were chosen to lead the technical efforts by developing an end-to-end test bed system demonstrating a wide range of advanced surveillance techniques: real-time moving object detection and tracking from stationary and moving camera platforms, recognition of generic object classes (e.g. human, sedan, truck) and specific object types (e.g. campus police car, FedEx van), active camera control and multi-camera cooperative tracking, object pose evaluation with respect to a geospatial site model, real-time data dissemination, recognition of simple multi-agent activities, data logging and dynamic scene visualization. [17].

A. *Works regarding Background Separation:*

Due to its pervasiveness in various contexts, background subtraction has been afforded by several researchers, and plenty of literature has been published [3, 4 & 5]. The usual method to find moving object detection is through background subtractions which require to maintain an up-to-date model of the background to detect moving objects by finding the deviation from such model [1]. Compared to other approaches, such as optical flow, this approach is computationally affordable for real-time applications. The main problem is its sensitivity to dynamic scene changes, and the consequent need for the background model adaptation via background maintenance. Such problem is known to be significant and difficult. Some of the well-known issues in background maintenance, that will be specifically addressed in the sequel, include:

***a. light Changes*:**
the background model should adapt to gradual illumination changes;

b. Moving Background:
the background model should include changing background that is not of interest for visual surveillance, such as waving trees;

***c. Cast Shadows*:**
the background model should include the shadow cast by moving objects that apparently behaves itself moving, in order to have a more accurate detection of the moving objects shape;

***d. Bootstrapping*:**
the background model should be properly set up even in the nonexistence of a complete and inert training set at the beginning of the sequence;

***e. Camouflage*:**
moving objects should be detected even if their chromatic features are similar to those of the background model.

In our proposed method moving object detection is based on the background model generated automatically by a self organizing method with no prior knowledge about the involved patterns [2]. The idea consists in adopting biologically inspired methods for moving object detection, where visual attention mechanisms are used to help detecting objects that keep the user attention in accordance with a set of predefined features, such as gray level, motion, and shape features.

It will be shown, by qualitative and quantitative results that our adaptive model can cope with all the above mentioned issues for background maintenance and achieves robust detection for different types of videos taken with stationary cameras. There is no unique classification of proposed methods.

III. METHODOLOGY

A. *Proposed System*

In this research first work is to design the form which consists of the buttons used for browsing the video which is being used for whole project, then for encoding and decoding the video, in this part video which is used will be compressed in size. The size of the original video and the compressed video is displayed. Mainly compression is used for two main purposes:

a) To reduce the memory
b) Efficient use of bandwidth

Then the compression ratio CR is found by using the formula:

CR = Fo / Fc

where Fo is the original File Size and Fc is the compressed File Size

After compressing the video using the encoding and decoding methods and showing the compression ration the reconstructed output is shown. This looks as like the original video but the size of the video is reduced.

Motion detection using Threshold method includes three parts as below;

a) Motion detection using consecutive frames
b) Background frame subtraction
c) Linear background frame update

In terms of functionality, motion detection between consecutive frames can be used for later linear background frame update. If there is change between two consecutive frames at the same location, the pixel at this location in the background should be updated; if not, the pixel at this location in the background is left unchanged. The background frame subtraction plays the key role in our motion detection algorithm, which directly separates moving objects from background. To meet the real-time demand, a linear background frame update is adopted.

Diff =current frame (for ground) - previous frame (back ground) for doing all these, Threshold point should be fixed in the video. That threshold point should be fixed without changing throughout the process In the proposed algorithm, the moving detection is Let In and In-1 denote the current frame and the frame before current frame, respectively. Th is defined as a threshold for motion detection and Diff(i, j) as a motion identifier for pixel (i, j), which can be calculated as:

$$Diff(i,j) = \begin{cases} 1 & |I_n(i,j) - I_{n-1}(i,j)| > Th \\ 0 & \text{else} \end{cases}$$

Final output can be viewed at this stage. The background fully viewed in black color and the foreground image can be viewed in white color. The threshold point of this moving image would not get changed through out the process. The final output can be viewed in several frames which are equal to the number of frames which is specified in the encoding and decoding stages. This can be changed manually. But if the number of frames which are selected is high the time which is taken for the whole process will be also very high.

Finally output has been produced to the selected video by applying the SPHIT Compression algorithm.

B. Methodology

a. Frame Separation:

i. Encoding and Decoding:

An encoder simply compares the block being coded with displaced blocks in the reference frame (a previous or future frame). The comparison can use mean squared error or some other metric of differences between images. The encoder selects the displaced block ith the smallest mean squared error difference! At no point has the encoder recognized an object in the image. The encoding process is called Motion Estimation. This finds the motion vector (or vectors) for each block. The decoding process is known as Motion Compensation. Motion Compensation accomplishes greater compression than Frame Differencing.

ii. Video Coding:

A typical system is shown in Figure 1:

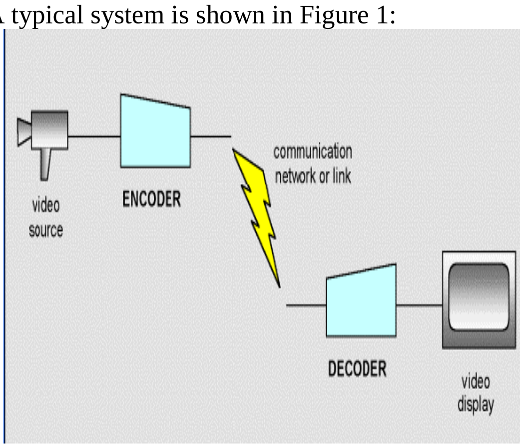


Figure 1. Basic Diagram for video compression.

Frames of the source video information are catched and are compressed by a video encoder. The compressed "stream" is transmitted across a network or telecommunications link and decoded (decompressed) by a video decoder and then can be displayed.

b. Compression:

Compression is a conversion of data to formats that requires less bits and performed to store and transmit the data more efficiently and it can be reversed back to its original data afterwards. The data size in compressed form (C) relative to the original size (O) is known as the compression ratio (R=C/O). If the inverse process, decompression, produces an accurate copy of the original data then it is lossless compression. Lossy compression used to image data does not allow duplication of an accurate imitation of the original image, but has a high compression ratio. Thus lossy compression permits only a rough calculation of the original to be generated. For image compression, the reliability of the approximation decreases when the compression ratio increases. Compression is similar to folding a letter before putting it in a small envelope to transport it more easily and cheaply. Compressed data is not easily readable and must first be decompressed to restore. Referred from papers [6-12] and [19].

i. Naturally Lossy Compression

Sufficiently low quality lossy compression will introduce visual artifacts that are highly annoying to the human viewer. An example is the blocking artifacts visible in highly compressed MPEG video and other block Discrete Cosine Transform based image compression codecs. At some point the lossy compression introduces unnatural artifacts which are perceived as new objects in the scene or fake lines within the image.

Objects in the visual system are delineated by edges. Anything like a codec algorithm that obliterates or creates an edge in an image is noticed [18].

ii. Unnatural Lossy Compression

Widely used video compression algorithms are lossy. Even they provide high compression, most of them will have problems with the edges in the image. Vector quantization, wavelet based image and video compression and block Discrete Cosine Transform naturally do not represent the intuitive notion of an edge or line.

iii. Image Compression Overview:

The Figure 2 shows the over view of Image Compression.

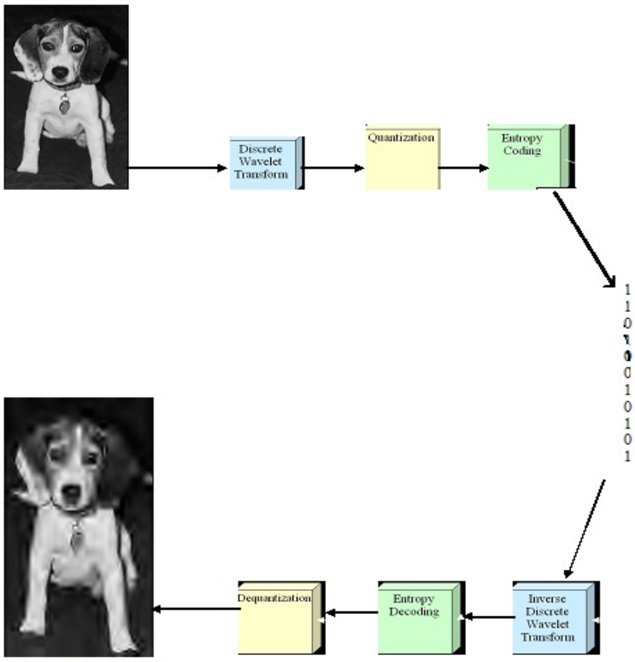


Figure 2. Image Compression Overview.

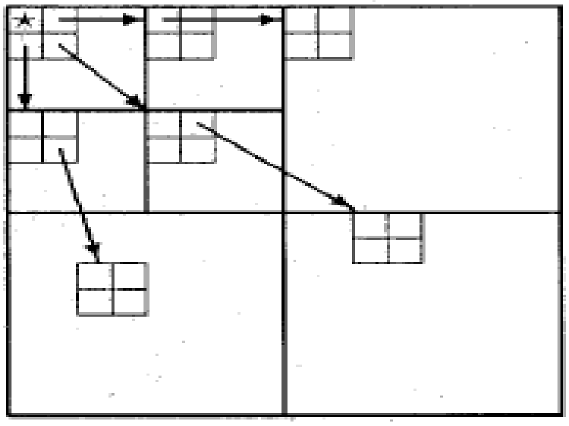
iv. SPHIT Algorithm description:

SPIHT algorithm is based on 3 concepts :

a) Ordered bit plane progressive transmission.
b) Set partitioning sorting algorithm.
c) Spatial orientation trees.

[a] Spatial Orientation Trees – (1)

Tree root is 2 by 2



[b] Spatial Orientation Trees – (2)

The sets of coordinates used to present the new coding method here are :

a) O (i, j) : set of coordinates of all offspring of node (i, j) ;
b) D (i, j) : set of coordinates of all descendants of the node (i, j) ;
c) H : set of coordinates of all spatial orientation tree roots ;
d) L (i, j) = D (i, j) – O (i, j) ;

[c] Coding Algorithm

In a practical implementation of SPIHT the significance information is stored in three ordered list :

a) LIS : list of insignificant sets
b) LIP : list of insignificant pixels
c) LSP : list of significant pixels

[d] A Simple Example Encoder: First Phase

26	6	13	10
-7	7	6	4
4	-4	4	-3
2	-2	-2	0

[e] Initialization :

Chose T0 = 2 └log226┘ = 16; **(n = 4)**
LIP = { (1,1), (1,2), (2,1), (2,2) }   LSP = { }
LIS = { D(1,2), D(2,1), D(2,2) }

[i] Process LIP

Sn(1,1) ◊ 26 ≥ T0, we transmit 1; 26 > 0, we transmit1;
Then move (1,1) to LSP.
Sn(1,2), Sn(2,1), Sn(2,2) are insignificant, transmit three 0;

[ii] Process LIS

Sn(D(1,2)) ◊ 13, 10, 6, 4 < T0, we transmit 0;
Sn(D(2,1)), Sn(D(2,2)) are insignificant, transmit two 0;

[iii] Needn't to process LSP (because LSP = NULL)

[iv] Update LSP

The transmitted bit stream : 11000000 (8 bits)
LIP = { (1,2), (2,1), (2,2) }
LIS = { D(1,2), D(2,1), D(2,2) }
LSP = { (1,1) }

v. Discrete Wavelength Transformation: (DWT)

Overview of Wavelet

Wavelets are mathematical functions that dice data into various frequency components, and each component is studied with a resolution matched to its scale. The basis functions of baby wavelets are obtained from a single prototype wavelet called mother wavelet, by dilations (scaling) and translations (shifts).

[a] Wavelet Transform

Wavelet Transform is a signal representation that can provide the frequency content at a particular instant of time. Wavelet analysis has advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes.[13-15]

[b] A DWT Example

When the Discrete Wavelet Transform is applied on an image, the result will be divided as shown in Figure 3.

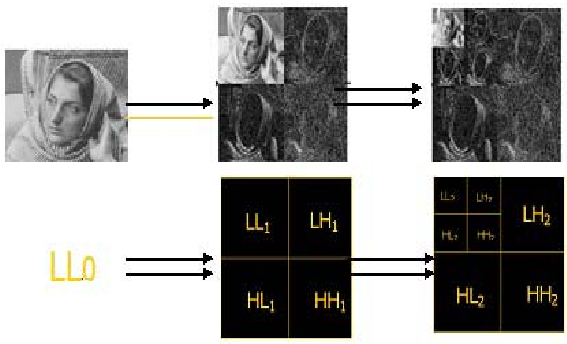


Figure 3. Example of a Discrete Wavelet Transform.

C. System Flow Diagram

The Overall system flow diagram and Encoding and Decoding diagrams are as shown in Figure 4, 5 and Figure 6.

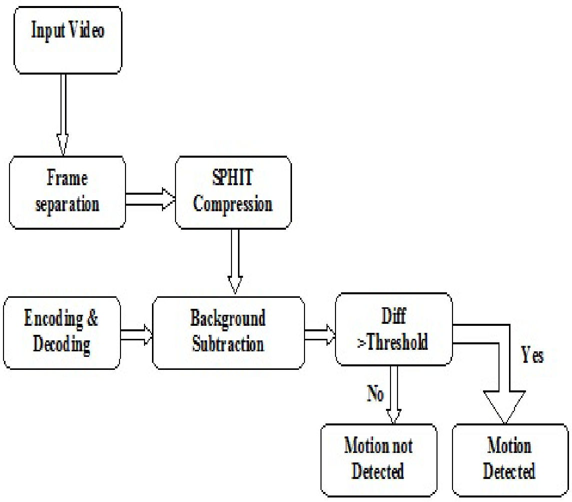


Figure 4. Overall System Flow Diagram.

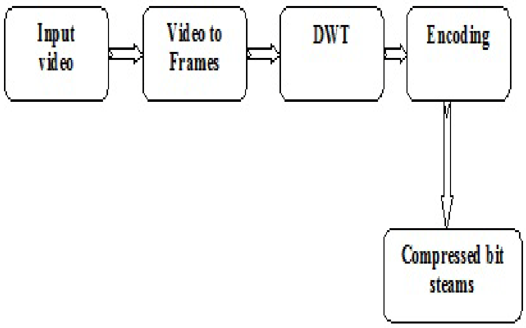


Figure 5. System Flow diagram for Encoding.

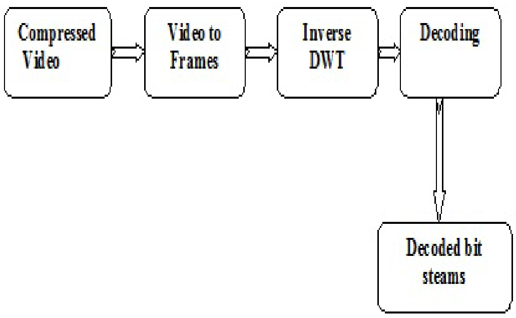


Figure 6. System Flow diagram for Decoding.

IV. IMPLEMENTATION AND RESULTS

This research work is implemented using Matlab. The Main Interface of this work is shown in Figure 7.

Figure 7. Main Interface.

When we press the 'Browse_Video' button, a browse window get opened, from that the needed video can be selected. After that the selected video can be viewed as shown in Figure 8.

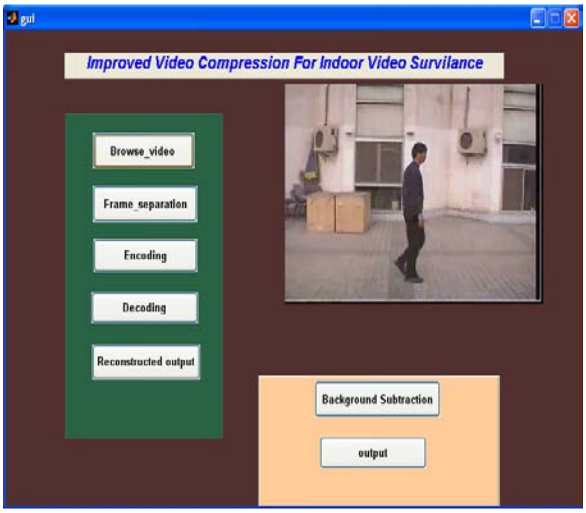


Figure 8. Input Video Selected.

When we press 'Frame_Seperation' button, the selected video is divided into several frames and that frames are stored in the target folder and the message 'Frame separation is completed' is shown (Figure 9).

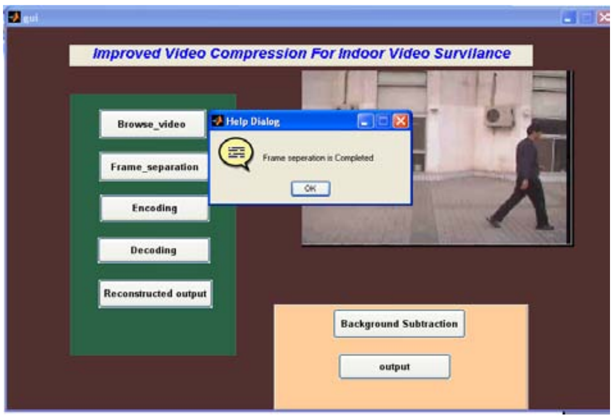


Figure 9. Frame Separation Completed.

After frame separation encoding process has to be carried out. In this process the entire frame is divided into four equal parts by using Discrete Wavelet Transformation method, each frames are named as LL, LH, HL, LL. Video can be viewed in all frames but the clear video can be viewed only in the LL part (Figure 10).

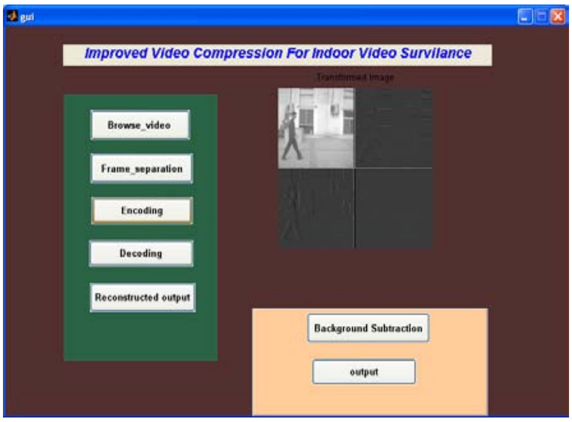


Figure 10. Encoding is in Progress.

After completing the encoding process, the message 'Encoding process is completed' will be shown (Figure 11).

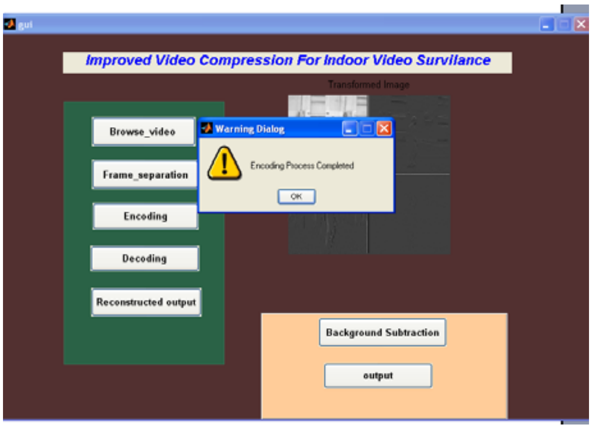


Figure 11. Encoding is Over.

Next to that Decoding process is carried out by applying IDWT- Inverse Discrete Wavelet Transformation and the message 'Decoding Process Completed' will be shown (Figure 12).

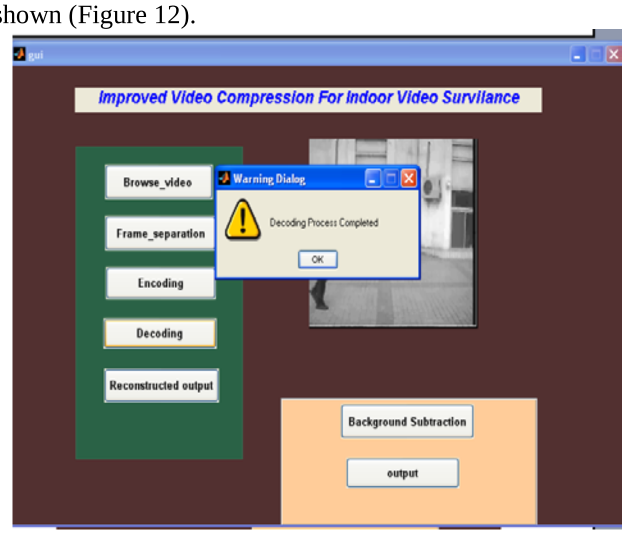


Figure 12. Decoding is Done.

After the decoding process, the reconstructed video can be viewed. In that process same video is viewed but after compressing it. The decompression rate is considerably high and the size of the video file is considerably small. For the video taken for tests, the average compression rate is more than 50% and the compression speed is also high.

The result of experiment of the algorithm on one of such video taken is discussed. The input video is a scene of a walking man. This video has a dimension of 720 × 576 of 7 second duration. The novel algorithm converts this input video into 51 frames. It gives the output with 30 frames of duration 3 second with dimension of 468 × 378 with compressed size of 754 KB. The mean noise ratio was employed to evaluate the quality of reconstructed images objectively, although it is not always reliable as a gauge of subjective visual quality.

V. CONCLUSION AND FUTURE WORK

In the light of the results from the experiments, some conclusions may be given about the success of the current research, according to the goals that were established for the project. These objectives were the compression of video with improve quality, is satisfactorily achieved with good compression ratio. The SPHIT video compression algorithm is fast and effective. This algorithm is automatic and requires no parameter adjustment and no a priori knowledge of the acquisition conditions. This is because functions evaluate their parameters or use pre-adjusted defaults values. Inverse filtering gives good results but generally requires a priori knowledge on the environment. But in our preprocessing, filtering needs no parameters adjustment so it can be used systematically on underwater images. The proposed system can be used in applications like Video Surveillance Applications, Motion pixel Estimation Applications, Video Surveillance, Weapon Storage Area and Atomic Research Area etc.

VI. BIBLIOGRAPHY

[1] Lucia Maddalena, Alfredo Petrosino "A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection", Neural Computing and Applications Journal, Volume 19 Issue 2, March 2010 10.1007/s00521-009-0285-8.

[2] Maddalena, L. Petrosino, A. "A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications ", IEEE, ISSN: 1057-7149, Pgs 1168 – 1177.

[3] Massimo Piccardi, "Background subtraction techniques: a review" 2004.

[4] Volkan Cevher et.al "Compressive Sensing for Background Subtraction" 2008

[5] Sen-Ching S. Cheung and Chandrika Kamath "Robust techniques for background subtraction in urban traffic video" UCRL conference 2007.

[6] Khalid Sayood, "Introduction to Data Compression 2/E", M.K. (2000)

[7] David Salomon, "A Guide to Data Compression Methods", Springer (2002).

[8] Introduction to Data Compression , 2nd Ed., by K. Sayood, Academic Press, Morgan Kaufmann Publishers, 2000.

[9] Djohan, T. Q. Nguyen, andW. J. Tompkins, "ECG compression using discrete symmetric wavelet transform," in Proc. 17th Int. Conf. IEEE
[10] JPEG2000: Image Compression, Fundamentals, Standards, and Practice by D. S. Taubman and M. W. Marcellin, Kluwer Academic Publishers, 2002.
[11] Data Compression: The Complete Reference by D. Salomon, Springer Verlag, 2000.
[12] A Guide to Data Compression Methods (Paperback) by D. Salomon, Springer Verlag, 2002.
[13] www.tenthofmarch.com/.../how-to-compress-video-before-uploading-to-youtube/
[14] www.filmmakingwebinars.com/.../web-video-compression-techniques/
[15] Wavelet Image and Video Compression , P. N. Topiwala, Editor, Kluwer Academic Publishers, 1998.
[16] Wavelet Transforms: Introduction to Theory and Applications by R. M. Rao and A. S. Bopardikar, Addison Wesley Longman, 1998.
[17] http://www.cs.cmu.edu/~vsam/.
[18] http://www.jmcgowan.com/avialgo.html
[19] John McGowan "AVI Overview: Video Compression Technologies" 2000.