Frame Selection Key to Improve Video Compression

Talal Ali Hammouri[1] and Muzhir Shaban Al-Ani[2]

[1]Amman Arab University, Department of Computer Science,
Amman-Jordan, 11953
talalah@yahoo.com

[2]Amman Arab University, Department of Computer Science,
Amman-Jordan, 11953
muzhir@gmail.com

Abstract
The huge usage of digital multimedia via communications media, wireless communications, Internet, Intranet and cellular mobile leads to incurable growth of data flow through these media. The researchers go deep in developing efficient techniques in these fields such as data compression image compression and video compression. Recently, video compression techniques and their applications in many areas (educational, agriculture, medical …) cause this field to be one of the most intersect fields.
Wavelet transform is an efficient method that can be used to perform an efficient compression technique. The proposed approach deal with the developing of efficient video compression approach based on frame selection key that concentrated on the calculation of frame near distance (or difference between frames). The selection of the meaningful frame depends on many factors such as the compression performance, frame details and frame size and near distance between frames. In this paper, many videos are used and tested to insure the efficiency of the proposed technique, in addition a good performance results has been obtained.

Key Words: Discrete Wavelet Transform, Video Compression, Frame Selection Key, Compression Performance and Near Distance Calculation.

1. Introduction

There are many methods used for video compression, and the most famous and apply technique is Moving Picture Experts Group (MPEG), which is an ISO/ITU standard for compressing digital video. MPEG uses lossy compression within each frame similar to JPEG, which means pixels from the original images are permanently discarded. The most famous methods are concentrated of Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT).
The grey scale images gives 256 levels of possible intensity for each pixel, so these images refer to 8 bits per pixel (bpp). The typical RGB color images, with 8 bits for Red, 8 bits for Green, and 8 bits for Blue, then the intensity I is defined by (I=R+G+B). The human eye is most sensitive to variations in intensity, so the most difficult part of compressing a color image lies in the compressing of the intensity.
Digital video consists of a stream of images captured at regular time intervals. The images are represented as digitized samples containing visual (color and intensity)

information at each spatial and temporal location. Visual information at each sample point may be represented by the values of the three basic color components RGB color space. A video signal can be sampled in either frames (progressive) or fields (interlaced). In progressive video, a complete frame is sampled at each time instant. While an interlaced video only a half of the frame is captured (either odd or even rows of samples) at a particular time instant which are called fields. In this paper we deal with a sequence of complete frames in which we calculate the frame differencing between the consequences frames according to a specific threshold.

2. Literature Reviews

Many papers and researches are published related to this subject and some of these are listed below:

Chen Guanghua et al. proposed an AVS real time high definition video encoder for its high memory bandwidth and large calculation complexity caused by the new coding features of variable block size and 4-tap filter. The proposed design is implemented on FPGA with operating frequency of 150 MHz and can support 1080p (1920*1080)/30fps AVS real time encoder [1].

Xue Bai et al. introduced a new color model, Dynamic Color Flow, which incorporates motion estimation into color modeling in a probabilistic framework, and adaptively changes model parameters to match the local properties of the motion. The proposed model accurately and reliable describes changes in the science's appearance caused by motion across frames [2].

Esa Rahtu et al. implemented a new salient object segmentation method based on combining a saliency measure with a conditional random field. The proposed saliency measure is formatted using a statistical framework and local feature contrast in illumination, color, and motion information [3].

Mohammed et al. presented an improved DC prediction mode based on the distance between the predicted and reference pixels. The proposed system adaptively selects the number of prediction mode for each 4*4 or 8*8 blocks, which reduces the number of overhead bits and computational cost of the encoder [4].

Eugeniy Belyaev et al. proposed a new spatial scalable and low complexity video compression algorithm based on multiplication free three dimensional discrete pseudo cosine transform. This paper shows an efficient results compared with H.264/SVC as well as it can be used for robust video transmission over wireless channels [5].

Linh Tran Ho et al., introduced a robust motion estimation scheme using a direction integrated genetic algorithm to speed up the encoding process of H.264/AVC video compression as well as to keep low bits to code frames. The obtained results show that the proposed motion estimation obtains the expressive results for both the number of bits to code a frame and the time cost to code a frame [6].

Evgency Kaminsky et al., proposed an effective DCT-domain video encoder architecture that decreases the computational complexity of conventional hybrid video encoders by reducing the number of transform operations between the pixel and DCT domain. The proposed system is based on the conventional hybrid coder and on a set of fast integer composition DCT transform. The proposed architecture may be used for the future Internet and 4G applications [7].

Cong Dao Han et al., implemented a novel search algorithm which utilizes an adaptive hexagon and small diamond search to enhance search speed. Simulation results showed that the proposed approach can speed up the search process with little effect on distortion performance compared with other adaptive approaches [8].

Jerome Gorin et al., produced an abstract decoder models (ADM) of MPEG decoders as programs described in a dataflow language, which naturally expresses potential parallelism between tasks of an application which makes an ADM description suitable for implementation to a wide variety of platforms for uni-processor systems to FPGAs [9].
Enrico Baccaglini et al., presented a performance compression between multiple description coding and unequal loss protection as tools to deliver multimedia data to mobile user. The compression is performed using as a case study JPEG 2000 coded images and H.264/AVC video sequences transmitted over lossy packet networks. The simulation results show that both schemes allocate the same amount of redundancy for any given encoding output rate to protect the transmitted information [10].
V. Vijayalakshmi et al., proposed a new video encryption scheme for sensitive applications. The objective of the proposed system is to analyze a secure and computational feasible video encryption algorithm for MPEG video to improve the security of existing algorithm by combining encryption in Intra and Inter frames and to test the algorithm against the common attacks [11].
Mohamed Haj Taieb et al., introduced a novel distributed video coding scheme with adaptive puncturing that sends more parity bits when it is not the case. This scheme is based on the argument that redirecting the party bits where they are the most effective, will improve the compression results. Simulation results demonstrate that lower bit rates are achieved [12].
Bibhas Chandra Dhara et al., presented a fast motion estimation algorithm and in residual frame coding proposed a fast method based on block truncation coding with pattern fitting concept is employed. The proposed video coding method is afast one with a good quality at the responsible bit rate [13].
Changfeng Niu et al., proposed a novel method for moving object segmentation in the H.264 compressed domain. The proposed method utilizes some characters of H.264 besides motion information with no more decoding required. Experimental results of several H.264 compressed video sequences demonstrate the good segmentation quality of the proposed approach [14].
Boxin Shi et al., introduced a new coding plus color correction scheme for multi-view video by exploring the color redundancy. Some advanced features in H.264 codec are utilized as automatic color annotation, and then an optimization based colorization is performed to render the color picture [15].

3. MPEG Types

MPEG is an asymmetrical system. It takes longer to compress the video than it does to decompress it in the DVD player, PC, set-top box or digital TV set. There are many types of MPEG such as [9,11]:

MPEG-1 (Video CDs)

Although MPEG-1 supports higher resolutions, it is typically coded at 352x240 x 30fps (NTSC) or 352x288 x 25fps (PAL/SECAM). Full 704x480 and 704x576 frames (BT.601) were scaled down for encoding and scaled up for playback. MPEG-1 uses the YCbCr color space with 4:2:0 sampling, but did not provide a standard way of handling interlaced video. Data rates were limited to 1.8 Mbps, but often exceeded.

MPEG-2 (DVD, Digital TV)

MPEG-2 provides broadcast quality video with resolutions up to 1920x1080. It supports a variety of audio/video formats, including legacy TV, HDTV and five channel surround sound. MPEG-2 uses the YCbCr color space with 4:2:0, 4:2:2 and 4:4:4 sampling and supports interlaced video. Data rates are from 1.5 to 60 Mbps.

MPEG-4 (All Inclusive and Interactive)
MPEG-4 is an extremely comprehensive system for multimedia representation and distribution. Based on a variation of Apple's QuickTime file format, MPEG-4 offers a variety of compression options, including low-bandwidth formats for transmitting to wireless devices as well as high-bandwidth for studio processing.
A major feature of MPEG-4 is its ability to identify and deal with separate audio and video objects in the frame, which allows separate elements to be compressed more efficiently and dealt with independently. User-controlled interactive sequences that include audio, video, text, 2D and 3D objects and animations are all part of the MPEG-4 framework.
MPEG-7 (Meta-Data)
MPEG-7 is about describing multimedia objects and has nothing to do with compression. It provides a library of core description tools and an XML-based Description Definition Language (DDL) for extending the library with additional multimedia objects. Color, texture, shape and motion are examples of characteristics defined by MPEG-7.
MPEG-21 (Digital Rights Infrastructure)
MPEG-21 provides a comprehensive framework for storing, searching, accessing and protecting the copyrights of multimedia assets. It was designed to provide a standard for digital rights management as well as interoperability. MPEG-21 uses the "Digital Item" as a descriptor for all multimedia objects. Like MPEG-7, it does not deal with compression methods.
The Missing Numbers
MPEG-3 was abandoned after initial development because MPEG-2 was considered sufficient. Because MPEG-7 does not deal with compression, it was felt a higher number was needed to distance it from MPEG-4. MPEG-21 was coined for the 21st century. See MP3, M-JPEG, MPEG LA, MPEGIF, Pro-MPEG Forum, JPEG and interframe coding.

4. Frame Difference Measure

Color histograms are frequently used to compare images because they are simple to compute, and tend to be robust regarding small changes in camera view point. An image histogram H ( ) refers to the probability mass function of image intensities. Computationally, the color histogram is formed by counting the number of pixels belonging to each color. Usually a color quantization phase is performed on the original image in order to reduce the number of colors to consider in computing the histogram and thus the size of the histogram itself. There are a number of ways to compare color histograms. One simple method is the absolute difference between two color histograms as mentioned in the following equation.

$$D_{ifference\ \mathbf{rgb}}(I_i,I_j)=\sum_{K=1}^{n}\left(\left|H^r_i(k)-H^r_j(k)\right|+\left|H^g_i(k)-H^g_j(k)\right|+\left|H^b_i(k)-H^b_j(k)\right|\right)$$

Where, $H^r_i(j)$ is the histograms for color r in frame i.

5. **Two Dimensional DWT**

Two dimensional DWT is obtained via the implementation of low pass and high pass filters on rows and columns of image respectively. A low pass filter and a high pass filter are chosen, such that they exactly halve the frequency range between themselves. This filter pair is called the Analysis Filter pair. First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the lpf is a half band filter, the output data contains frequencies only in the first half of the original frequency range. the high pass filter is applied for the same row of data, and similarly the high pass components are separated, and placed by the side of the low pass components. This procedure is done for all rows. As mentioned above, the LL band at the highest level can be classified as most important, and the other 'detail' bands can be classified as of lesser importance, with the degree of importance decreasing from the top of the pyramid to the bands at the bottom. DWT is a multispectral technique used for converting signal or image into four different bands such as low-low (LL), low-high (LH), high-low (HL) and high-high (HH) as demonstrated in figure (1) [16,17].

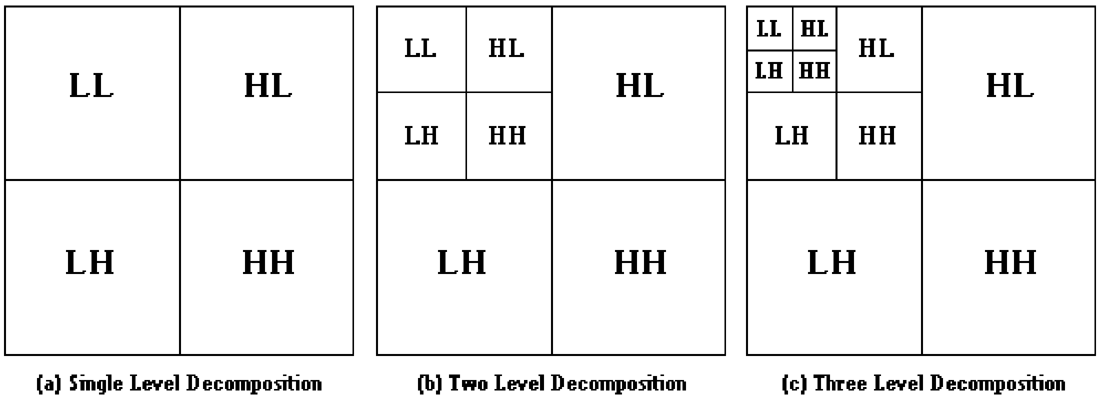


Figure (1) Decomposition of image applying DWT

6. **The Proposed System**

The proposed system concentrated on the adequate frame selection key in which used for the compression technique. This system is implemented via many stages as illustrated in figure (2) and these stages are listed below:

- Filtering stage that generates the de-nosing video.
- Frame extraction that generates frames for a certain time.
- Frame selection that generates effective frames.
- Frame reordering that prepared frames.
- Applying 2D DWT stage that generates the LL-band.
- Video construction that construct the final compressed video.

The main core of this system is the frame selection key, in which pass the similar frames and select the different frames depends on a certain specified threshold. In addition the selected frames are compressed via applying two dimensional discrete wavelet transform.

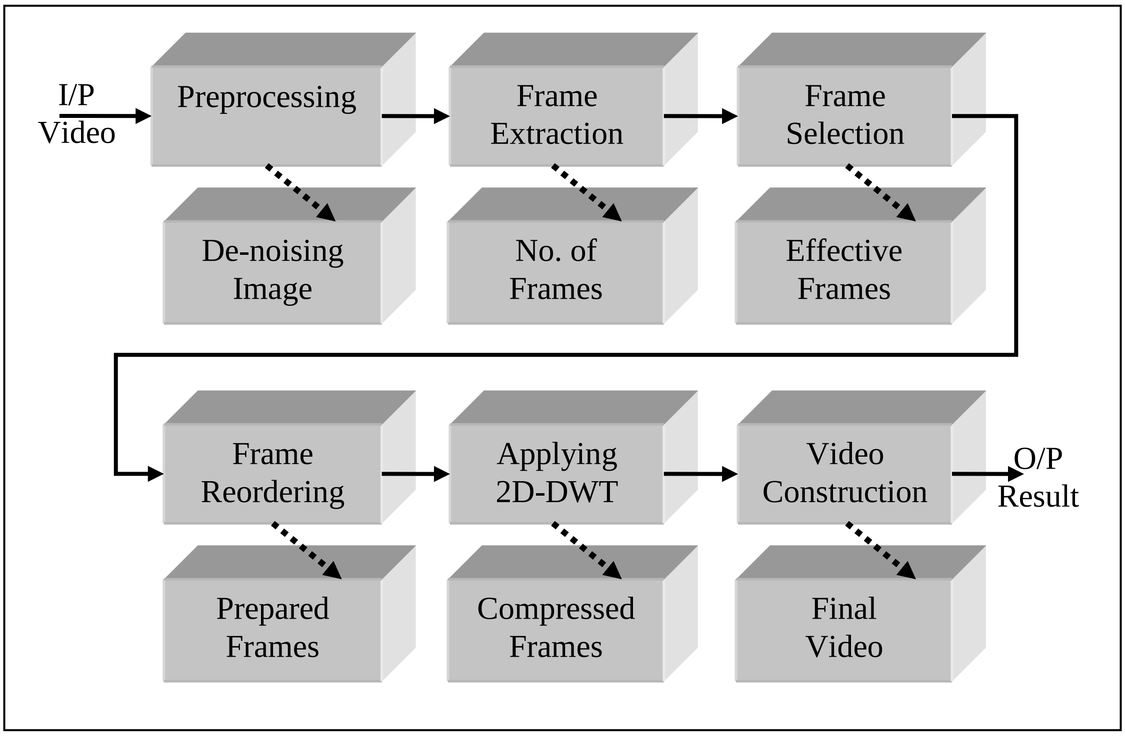
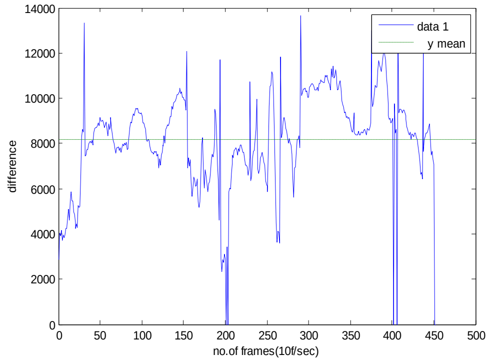


Figure (2) The proposed frame selection key approach

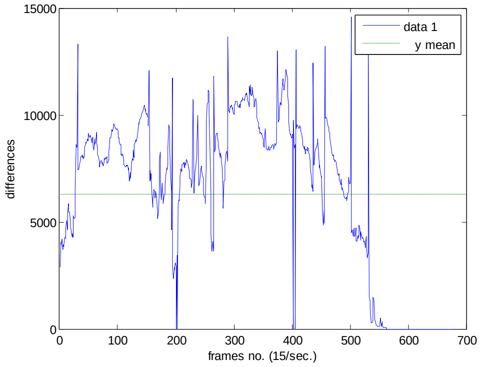
7. Results and Analysis

Many types of videos are used as a test video. These videos are extracted to forms different sets of frames. Different sizes of frames where examined. Then a mean averages are measured between each two nearest frames, and this procedure continue up to the end of frames. The frame selection key depends on the similarity between frames that are demonstrated in figure (4). Figure (4-a) indicates the results obtained in case of 10 frames per second, in which the mean value is concentrated on 8301. Figure (4-b) indicates the results obtained in case of 15 frames per second, in which the mean value is concentrated on 7612. Figure (4-c) indicates the results obtained in case of 20 frames per second, in which the mean value is concentrated on 6747. Figure (4-d) indicates the results obtained in case of 25 frames per second, in which the mean value is concentrated on 6185.
The threshold selection key depends on the averaging of all frames that are demonstrated in figure (5). Figure (5) a, b, c & d indicates the results obtained in case of 10, 15, 20 & 25 frames per second, in which data1 (blue), data2 (green) & data3 (red) depend on threshold selection. Thresholds data1, data2 & data3 depend on the difference between frames that are of values 3,4 & 5 respectively. It is clear from these results that data of threshold 5 have minimum values difference so it leads to better results of compression.

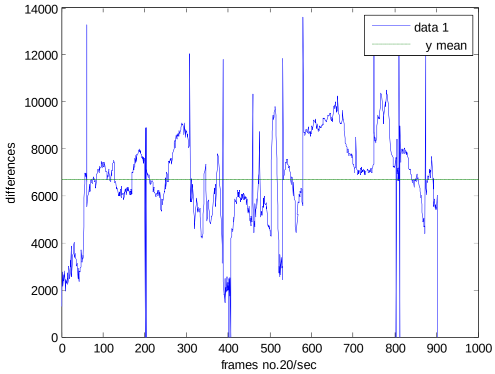

Figure (3) Different types of videos



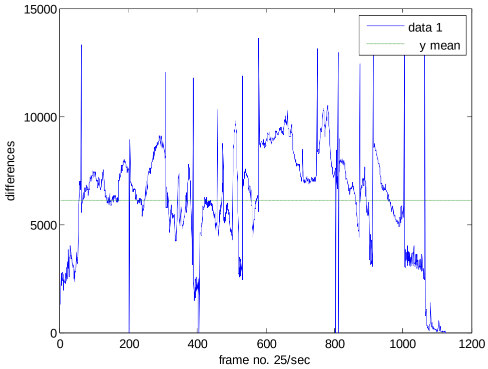
(a) 10 frames per second



(b) 15 frames per second

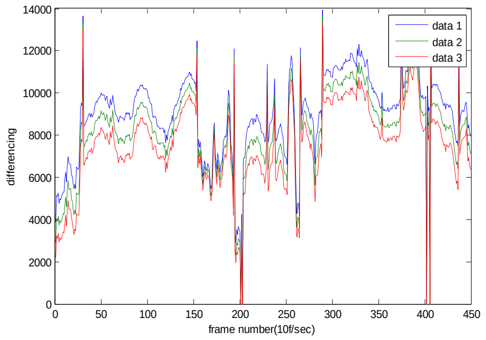


(c) 20 frames per second

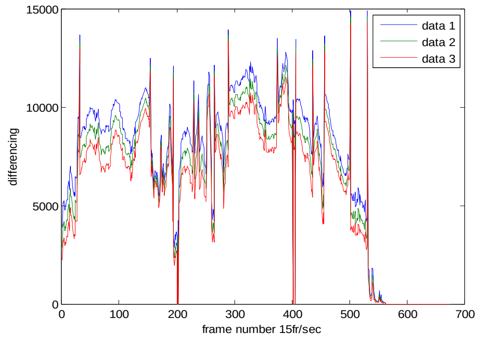


(d) 25 frames per second

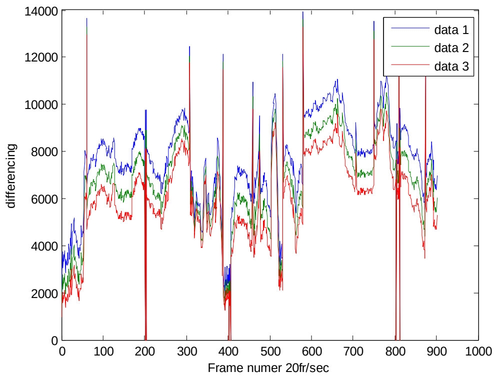
Figure (4) The results calculated for different number of frames



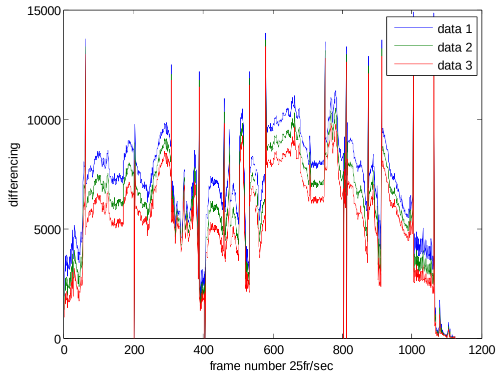
(a) 10 frames per second



(b) 15 frames per second



(c) 20 frames per second



(d) 25 frames per second

Figure (5) The results calculated for different number of threshold

8. Conclusion

An efficient video compression approach based on frame selection key is developed that concentrated on the calculation of frame near distance. Many factors are applied in the selection of meaningful frames, in which eliminate the similar frames. The proposed system passes into many steps; preprocessing, frame extraction, frame selection, frame reordering, 2D-DWT, then video construction. Different types of videos are introduced to test the system. The output compressed video is in a good quality and good performance as well as it has a specific compression ratio.

References

[1] Chen Guanghua et al., "VLSI Implementation of sub-pixel interpolator for AVS encoder", LSMS/ICSEE 2010, part II, LNCS 6329, pp 351-359, 2010. (fulltext2)

[2] Xue Bai et al. "Dynamic Color Flow: A Motion Adaptive Color Model for Object Segmentation in Video", ECCV 2010, part V, LNCS 6315, pp 617-630, 2010. (fulltext4)

[3] Esa Rahtu et al., "Segmenting Salient Objects from Images and Videos", ECCV 2010, part V, LNCS 6315, pp 366-379, 2010. (fulltext5)

[4] Mohammed et al., "Enhanced Intra Coding of H.264/AVC Advanced Video Coding Standard with Adaptive Number of Modes", ATM 2010, LNCS 6335, pp 361-372, 2010. (fulltext11)

[5] Eugeniy Belyaev et al., "Scalable Video Coding Based on Three Dimensional Discrete Pseudo Cosine Transform", ruSMART/NEW2AN 2010, LNCS 6294, pp 448-459, 2010. (fulltext12)

[6] Linh Tran Ho et al., "Direction Integrated Genetic Algorithm for Motion Estimation in H.264/AVC", ICIC 2010, LNAI 6216, pp 279-286, 2010. (fulltext18)

[7] Evgency Kaminsky et al., "DCT-domain coder for Digital Video Applications", Journal of Real Time Image Processing, 14 August 2010. (fulltext21)

[8] Cong Dao Han et al., "An Adaptive Fast Search Algorithm for Block Motion Estimation in H.264", Journal of Zhejiang University Science C (Computer & Electron), Vol. 11, No. 8, pp 637-644, 2010. (fulltext30)

[9] Jerome Gorin et al., "LLVM – Based and Scalable MPEG-RVC Decoder", Journal of Real Time Image Processing, 27 July 2010. (fulltext32)

[10] Enrico Baccaglini et al., "Image and Video Transmission: A Comparison Study of Using Unequal Loss Protection and Multiple Description Coding", Multimedia Tools Applications, 27 July 2010. (fulltext33)

[11] V. Vijayalakshmi et al., "Efficient of Intra and Inter Frames in MPEG Video", CNSA 2010, CCIS 89, pp 93-104, 2010. (fulltext34)

[12] Mohamed Haj Taieb et al., "Turbo Code Using Adaptive Puncturing for Pixel Domain Distributed Video Coding", ICISP 2010, LNCS 6134, pp 342-350, 2010. (fulltext42)

[13] Bibhas Chandra Dhara et al., "A video Coding Technique Using Octagonal Motion Search and BTC-PF Method for Fast Reconstruction", AST/UCMA/ISA/CAN 2010, LNCS 6059, 480-490, 2010. (fulltext44)

[14] Changfeng Niu et al., "Moving Object Segmentation in the H.264 Compressed Domain", ACCV 2009, part II, LNCS 5995, pp 45-54, 2010. (fulltext59)

[15] Boxin Shi et al., " Color Correction and Compression for Multi-View Video Using H.264 Features", ACCV 2009, part III, LNCS 5996, pp 43-52, 2010. (fulltext60)

[16] Chen Hai Lin et al., "Multiple Descriptions Based Wavelet Image Coding", Journal of Zhejiang University Science, vol.5, no.1, pp55-61, Jan. 2004.
[17] Don Hong et al, "Wavelet and Biorthogonal Wavelets for Image Compression", Robotic Welding and Automation, LNCIS 299, pp281-303, 2004.

[1]**Talal Hamori**

[2]**Muzhir Shaban Al-Ani** has received Ph. D. in Computer & Communication Engineering Technology, ETSII, Valladolid University, Spain, 1994. Assistant of Dean at Al-Anbar Technical Institute (1985). Head of Electrical Department at Al-Anbar Technical Institute, Iraq (1985-1988), Head of Computer and Software Engineering Department at Al-Mustansyria University, Iraq (1997-2001), Dean of Computer Science (CS) & Information System (IS) faculty at University of Technology, Iraq (2001-2003). He joined in 15 September 2003 Electrical and Computer Engineering Department, College of Engineering, Applied Science University, Amman, Jordan, as Associated Professor. He joined in 15 September 2005 Management Information System Department, Amman Arab University, Amman, Jordan, as Associated Professor, then he joined computer science department in 15 September 2008 at the same university.