A Real-Time H.264/AVC Encoder&Decoder with Vertical Mode for Intra Frame And Three Step Search Algorithm for P-Frame

Dr. Mohammed H. Al-Jammas[1] and Mrs. Noor N. Hamdoon[2]

[1]Deputy Dean/College of Electronics Eng./University of Mosul/Mosul-Iraq
dr_mohammed_al_jammas@uomosul.edu.iq
[2]Department of Electrical Engineering /College of Engineering/Mosul–Iraq
Noornawaf81@yahoo.com

ABSTRACT

The video coding standards are being developed to satisfy the requirements of applications for various purposes, better picture quality, higher coding efficiency, and more error robustness. The new international video coding standard H.264 /AVC aims at having significant improvements in coding efficiency, and error robustness in comparison with the previous standards such as MPEG-2, H261, H263,and H264. Video stream needs to be processed from several steps in order to encode and decode the video such that it is compressed efficiently with available limited resources of hardware and software. All advantages and disadvantages of available algorithms should be known to implement a codec to accomplish final requirement. The purpose of this project is to implement all basic building blocks of H.264 video encoder and decoder. The significance of the project is the inclusion of all components required to encode and decode a video in MatLab .

KEYWORDS

H264/AVC , Intra frame (I-frame) , Inter frame (P-frame)

1. INTRODUCTION

A Digital video compression is an important techniques that enables efficient transmission bandwidth and storage space of multimedia. The H.264/AVC is a standard video coding that developed to achieve significant improvements, in the compression performance, over the existing standards. In fact, the high compression performance comes mainly from the prediction techniques that remove spatial and temporal redundancies. To remove spatial redundancy, H.264/AVC intra prediction supports many prediction modes to make better prediction. Inter prediction is enhanced by motion estimation (ME) to remove more temporal redundancy. However, the H.264/AVC coding performance comes at the price of computational complexity[1].

H.264/AVC intra encoding achieve higher compression ratio and picture quality compared with the latest still image coding standard JPEG2000].intra prediction is the first process of advanced video coding standard. It predicts a macro block by referring to its previous macro blocks to

reduce spatial redundancy. Intra prediction supports nine modes for 4x4 block and four modes for 16x16 blocks[2].

H.264 is an open, licensed standard that supports the most efficient video compression techniques available today. Without compromising image quality, an H.264 encoder can reduce the size of a digital video file by more than 80% compared with the Motion JPEG format and as much as 50% more than with the MPEG-4 Part 2 standard. This means that much less network bandwidth and storage space are required for a video file. Or seen another way, much higher video quality can be achieved for a given bit rate[3].

And also entered as an adjunct of the kind of evolutionary in public services such as video storage on the Internet and telecommunications companies and surveillance cameras used in industrial plants , and due to accept this kind of decoder over a wide range of frames during the second (60/30/25 (fps)) has been expanding in applications control of highways, airports , Most of the controversy over the techniques used to process coded information is how fast and accurate images and video after the code process it possible to re-information fully if it is reduced to half?. Will be answered by this research, which includes the representation of the coded H264 and decode of the file with a high level using the MATLAB simulation software to achieve complete system of coded data and return it in the same efficient used for the original system.

2. THE ENCODING PROCESS

H.264 encoder works on the same principles as that of any other codec. Figures (eg, Figure 1) shows the basic building blocks of H.264 video codec.

The input to the encoder is generally an intermediate format stream, which goes through the prediction block; the prediction block will perform intra and inter prediction (motion estimation and compensation) and exploit the redundancies that exist within the frame and between successive frames. The output of the prediction block is then transformed and quantized. An integer approximate of the discrete cosine transform is used (DCT) for transformation. It uses 4x4 or 8x8 integer transform, and outputs a set of coefficients each of which is a weighted value for a standard basis pattern. The coefficients are then quantized i.e. each coefficient is divided by an integer value. Quantization reduces the precision of the transform coefficients according to the quantization parameter (QP). Typically, the result is a block in which most or all of the coefficients are zero, with a few non-zero coefficients. Next, the coefficients are encoded into a bit stream. The video coding process creates a number of parameters that must be encoded to form a compressed bit stream[4]. These values include:

- Quantized transform coefficients.
- Information to re-create prediction.
- Information about the structure of compressed data and the compression tools used under encoding.

These parameters are converted into binary codes using variable length coding or arithmetic coding. Each of these encoding methods produces an efficient, compact binary representation of the information. The encoded bit stream can now be transmitted or stored.
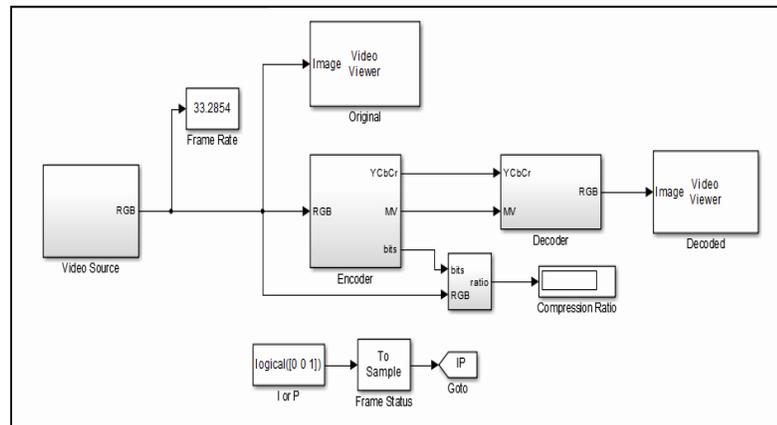
Figure 1. Encode and Decode circuit

3. VIDEO COMPRESSION WORKS

Video compression is about reducing and removing redundant video data so that a digital video file can be effectively sent and stored. The process involves applying an algorithm to the source video to create a compressed file that is ready for transmission or storage. To play the compressed file, an inverse algorithm is applied to produce a video that shows virtually the same content as the original source video. The time it takes to compress, send, decompress and display a file is called latency. The more advanced the compression algorithm, the higher the latency, given the same processing power. A pair of algorithms that works together is called a video codec (encoder/decoder).

Video codec's that implement different standards are normally not compatible with each other; that is, video content that is compressed using one standard cannot be decompressed with a different standard. For instance, an MPEG-4 Part 2 decoder will not work with an H.264 encoder. This is simply because one algorithm cannot correctly decode the output from another algorithm but it is possible to implement many different algorithms in the same software or hardware, which would then enable multiple formats to be compressed.

Different video compression standards utilize different methods of reducing data, and hence, results differ in bit rate, quality and latency. Results from encoders that use the same compression standard may also vary because the designer of an encoder can choose to implement different sets of tools defined by a standard. As long as the output of an encoder conforms to a standard's format and decoder, it is possible to make different implementations.

This is advantageous because different implementations have different goals and budget. Professional non-real-time software encoders for mastering optical media should have the option of being able to deliver better encoded video than a real-time hardware encoder for video conferencing that is integrated in a hand-held device. A given standard, therefore, cannot guarantee a given bit rate or quality. Furthermore, the performance of a standard cannot be properly compared with other standards, or even other implementations of the same standard, without first defining how it is implemented. A decoder, unlike an encoder, must implement all the required parts of a standard in order to decode a compliant bit stream. This is because a standard specifies exactly how a decompression algorithm should restore every bit of a compressed video [3]>

4. H.264 LEVELS

The Group focused joint development in determining work H.264 to find a solution is simple and flexible could include various applications through the use of a single standard, as is the case in video standards, etc., and is this flexibility in the provision of facilities for several profiles (represented groups of algorithms for the pressure data) and levels (level private suite of applications). The standard includes the following seven sets of capabilities, which are referred to as profiles, targeting specific classes of applications

- Baseline profile (BP): It is primarily for applications with low cost resources with physical entity and this file is used widely in applications of mobile devices and applications to transfer data, such as video conversations.
- Main Profile (MP): This file is used for broadcast applications and storage, and the importance of this file vanished when placing a high level to include those applications.
- Extended Profile (XP): means the image streaming video, this file has the ability of a relatively high pressure, and some additional possibilities to avoid data loss.
- High Profile (HiP): The primary file used for broadcast applications and disk storage, particularly in the application of high-definition television and, for example, applications in the HD DVD and Blu-ray .
- High 10 Profile (Hi10P): This file adds support for the previous file in decryption process where they can processing 10 bits per sample of image resolution to decode the data.
- High 4:2:2 Profile (Hi422P): used in professional applications that use interlaced video, this file is used in the previous file basis (Hi10P) and added his coordination shorthand chromatography with 4:2:2 sample while it uses 10 bits per sample per unit decoding and gives the same quality.
- High 4:4:4 Predictive Profile (Hi444PP): used in the applications non are for the loss of data was added to the previous file format support reduction chrome quality of 4:4:4 and support for samples up to 14 bits per sample, and also supports encoding videos and pictures with tri-color separate.

5. TYPES OF FRAMES

H.264 consists of several different types of frames, such as (I-P-B), and can be used for encryption to get the required efficiency below illustrate the theoretical formula for each quality of frames.

- (I-intra frame) Is an autonomous framework which can encrypt and decrypt independently without need for another picture as a source of information retrieval, the first image of the video is for this type of frame, and the (I-frame) is the starting point for the video display as well as his importance in information retrieval synchronization if any damage in transport stream bit (bit stream), the flaw in this window that consumes the largest possible number of bits for encryption because it takes the window image full but on the other hand, the error rate is low. Encryption method for this type of window has two properties, depending on the method of dividing the cluster either ((16x16) or (4x4)) but in General is to convert the frame version (RGB)format (YCbCr) and separated from the other components of the final representation and is treated with a single image, so the representation of video format ((4:2:0) YCbCr) is to reduce the sensitivity of the eye where the eye responds to brightness by colors so the component (Y) represents the symbol of brightness luminance while (CbCr) represents the color (chrominance) taken the element Y with full size while the rest of the elements are reduced by deciding to half

the amount of action in the element size (Y) is (16x16) , the rest of the elements are the size of (8x8), this means that embedded type of encryption key encryption process. Encryption process as previously mentioned it is dependent on frame division, divides the frame into multiple blocks of size (16x16) and has (4) types of encryption as shown in Figures (eg, Figure 2).
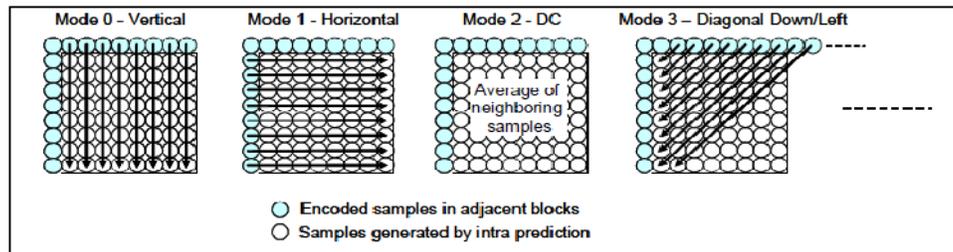

Figure 2. The types of patterns to divide block (16x16)

But the case of the split window to (4x4) , it has (9) types, as in Figures (eg, Figure 3)
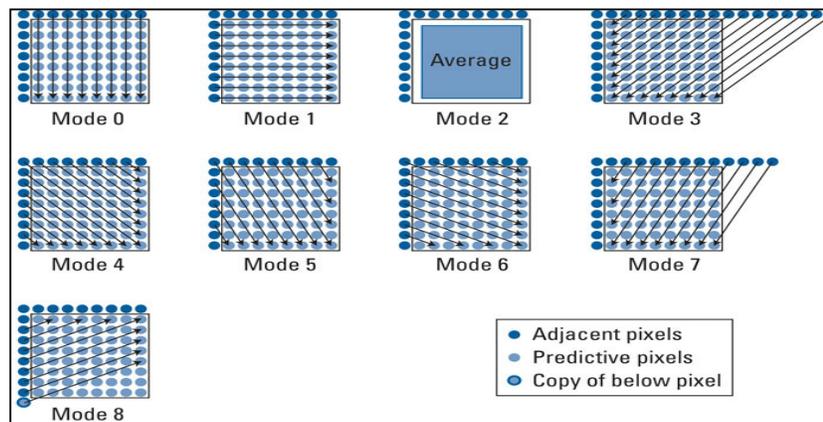
Figure 3. Types of Patterns to divide block (4x4)

Choose a style for the adoption application and competence required for the encryption process and the admissibility of the error rate, in most cases the amount of the error rate of the video is of a higher flexibility compared to the error rate in the case of a single image.

- (P-Inter Frame) Predictive Inter Frame: is derived from the current frame to the video sequence frame by reducing the time between frames increase unlike previous quality work only within the space of pixels, the principle of its work essentially compare the block of
- the current window with the block of the previous frame and the centre of block is search for match, this called (matching block), all theories have one and is the best possible match and this is called motion estimation (ME), after finding the best match, we put the block of the original block and the remaining known as compensation (motion compensation), the link between the location of the current block with original block is the transmission (motion vector (MV)) shown in Figures (eg, Figure 4).
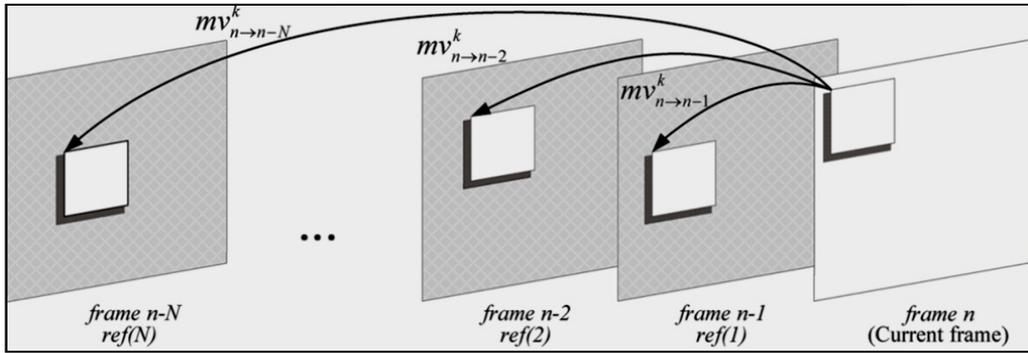
Figure 4. The basic idea to represent the predictive inter frame (P-frame)

- B-frames (Bi-predictive inter frame), this type of frame be intermediate between (I, and B frames) used at high levels for perfect efficiency but complex where the highest of qualities as follows based on the comparison between more than one source for block, meaning most forecasts from original source and source is expected as in Figures (eg, Figure 5).

When retrieving the information in the decryption process is (I-frame), the former to decrypt followed by (B-P frames) if used, the decryption depend one upon the other in the information retrieval of the original frame. H264 has several ways in the encryption that uses encryption (I-frame) or use (I&P ,(I&B&P)) and each method has its qualities, if you use the first method, the quantity of bits encoded be high compared with other cases but the error rate low because all the encrypted individually without relying on the previous window, this method is used in some applications that need high resolution cameras also in prisons and banks to get clearer picture during the up seizing process as in Figures (eg, Figure 6) .
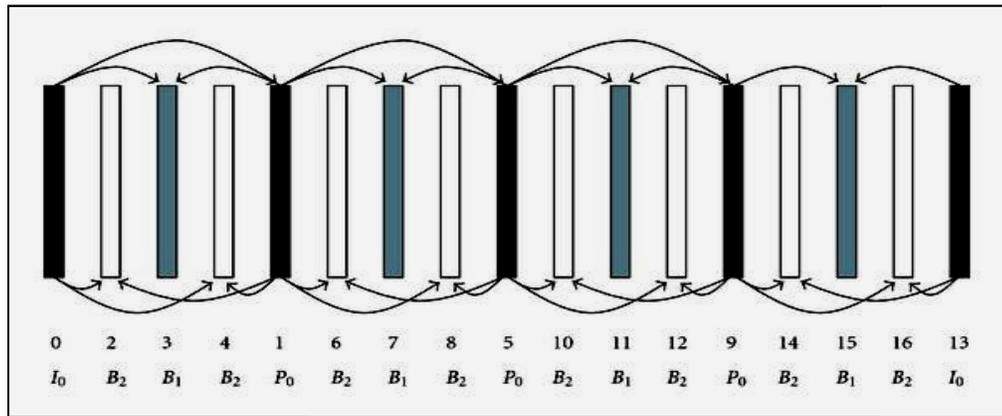

Figure 5. The representation frame (B)

Figure 6. Video encoding using spatial locations for pictures

But if you use the second method as shown in Figures (eg, Figure 7) they have characteristics that they reduce the number of bits encoded and the error rate is acceptable and this method is used in video compression in general and cameras , In the third grade are more complex than both methods but with a reduced data encrypted and contains a higher delay method because search matches more than one source.
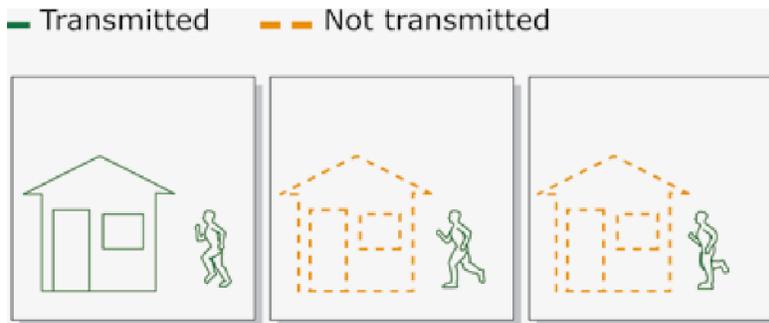

Figure 7. The first image is a frame (Frame-I) and encodes a single, the second and third picture only encrypts mobile part

6. SIMULATE ENCRYPTION AND DECRYPTION USING MATLAB

Previously we mentioned about the quality of the frames and how to process encoded in H.264 encoding process is divided into two ways:
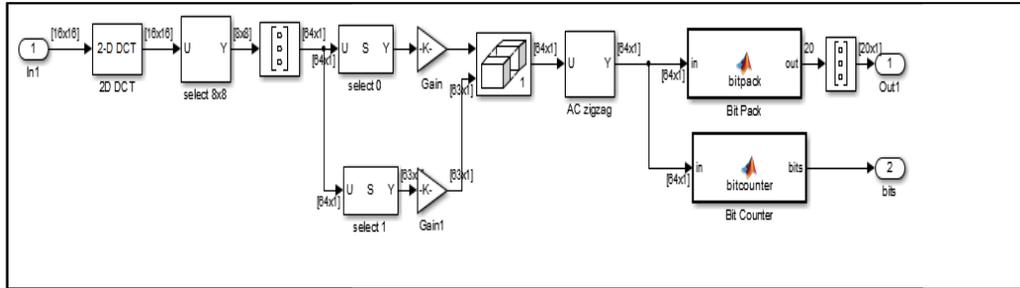
6.1 Encryption Process (I-frame)

I-frame is spatially encoded with a specific kind of styles in this research is the use of block size (16 x16), divided into (8x8), Tables (eg, Table 1) shows real-time to encrypt and decrypt the frame interconnection pattern within the first mode (Vertical mode) has been achieved through Figures (eg, Figure 8). Each block is handled independently of and separately from other 8x8 block are taken separately and applied the first style (vertical mode) and enters the cosine (DCT). This conversion generates a representation of each block of (8x8) in the frequency domain rather than the spatial domain. The resulting values of DCT process usually consists of a few large values and many small values represent the relative sizes of these transactions how important each block of information in the decryption stage of min after this operation involved transactions to bring it shown in Figures (eg, Figure 8).

Table 1. Time encode and decode for I-frame

Video / 30 fps	Time code /(sec/frame)	Time decode (sec/frame)
Forman(288x352)	0.0383	0.034
Vipmen(160x120)	0.00052	0.00104
Bride(720x1280)	0.468	0.159
Piano(352x480)	0.054	0.04

Figure 8. The encrypted representation service

6.2 Encryption Process (P-frame)

We mentioned already that this type of encryption is encryption And this means that it uses the correlation between the current frame and the frame (or frames) to achieve compression. Tables (eg, Table 2) Shows real-time to encrypt and decrypt for P-frame, Temporal encryption is achieved using vector motion (motion vectors.) the basic idea is to match each block in the current window size (16x16) pixel in the frame of reference, the match here can be calculated in many ways, but in H.264 uses a simple and more common is the sum of absolute differences (SAD) offset from the current transaction to the last movement is represented in reference to two values (horizontal, vertical vectors vector) are searched to find the best tankers in the selected area to search only when a less value of the total difference match between blocks this process is carried out under several algorithms one three-step algorithm to search for blocks to neighbourhoods Match blocks beginning with predictive search steps greater than half the search steps used in the previous methods and mechanism of action as follows:

- Compared with nine points in each step, three score and six points on the sides.
- Under search by one space after each step and the search stops when the size of the search space by one pixel.
- At each step a new search moving space research to find the best match for the Center block from the previous search, blue circles of number one as in Figures (eg, Figure 9) represent the first step of three steps when creating less differences between the current and previous bloc begin the second step of the green circles and also look on the differences is less up to the level of a single pixel of Orange and last circles represent the end of the search for this theory.
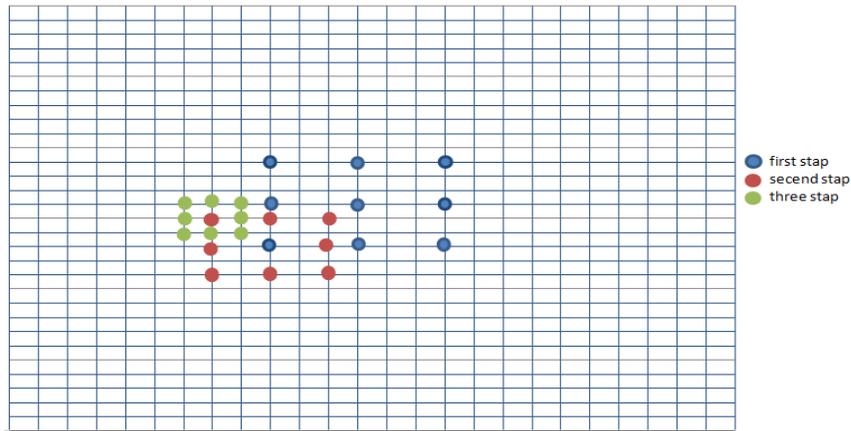
Figure 9. The theory of the three steps

Motion vector is a simple way to move a lot of information as shown in Figures (eg, Figure 10), but not always give an exact match to give the best quality, taking output subtraction between the original frame and block the cluster framework forecast output encrypted as shown in Figures (eg, Figure 11), the encryption process for residual image is similar to the encryption (I-frame), but the difference is in the process of rounding and through practical results as in Figures (eg, Figure 12) and found that the compression ratio is 70%of the original size.

Table 2. Time encode and decode for P-frame

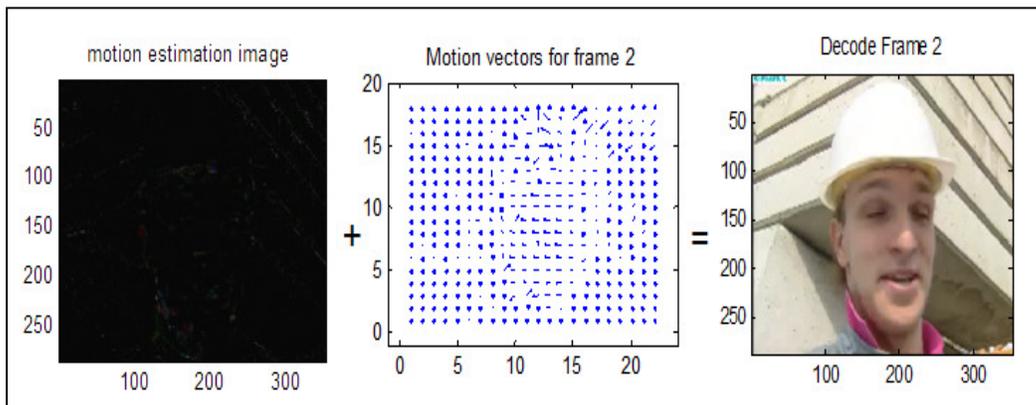| Video 30 fps	Time code (sec/frame)	Time decode (sec/frame)
Forman(288x352)	0.0013	0.0011
Vipmen(160x120)	0.00032	0.001
Bride(720x1280)	0.0161	0.0054
Piano(352x480)	0.00186	0.00173

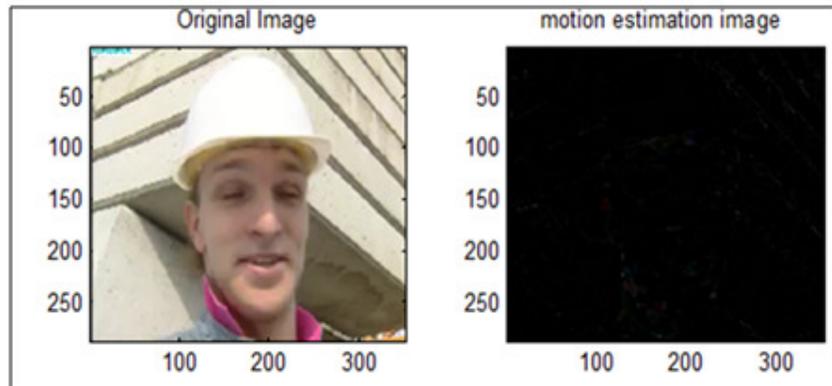Figure 10. The motion vectors for Residual Image
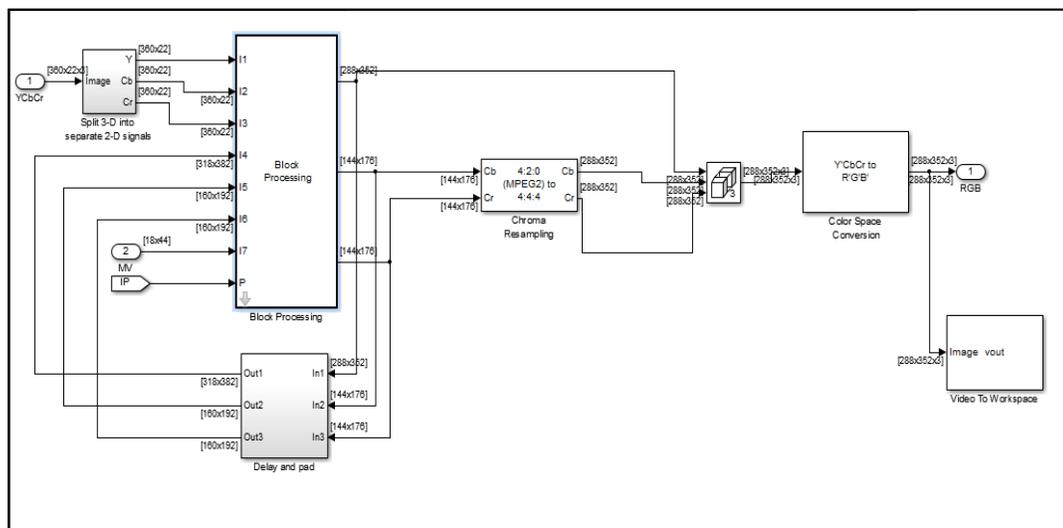
Figure 11. The Residual Image

Figure 12. The Decode Circuit

In Figures (eg, Figure 13) shows the encryption and decryption of video sequences Forman and note that in the case of image encryption and decryption have same properties.
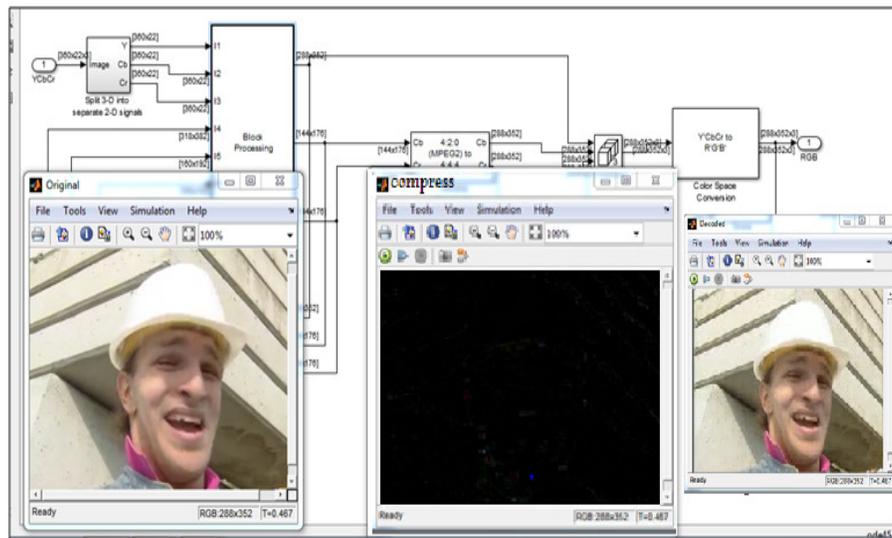
Figure 13. The Original Foreman Sequence and Return Within Three-Step algorithm

7. CONCLUSIONS

DisplaysH.264 A major step forward in the field of video compression technology. and provides techniques which enable better compression efficiency, due to more accurate forecasting capabilities, as well as improving the ability to minimize errors. It provides new possibilities for creating video encoders that managed to get high quality video and high frame rates at per second and higher resolution at bitrates (compared to the preceding criteria), and through the practical results of the MATLAB simulation was found that data compression during real time up to(70%) of the video size Original part time implementation 261.89 s and time clk (4ns) and highest value can be obtained for a reference to the noise up to (45 db) also shown in Figures (eg, Figure 14).

Figure 14. Simulink profile for H264/AVC

H.264 It is expected to replace other compression standards and methods used today, and form became H.264 more widely available species in network cameras, video encoders and video management software, designers of systems at the present time, the network video products support both H.264 and Motion JPEG is perfect for maximum flexibility and possibilities of integration.

REFERENCE

[1] A. Ben Atitallah, H. Loukil , and N. Masmoudi, FPGA DESIGN FOR H.264/AVC ENCODER, International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.1, No.5, October 2011, pp 119-138.

[2] Manjanaik.N, Dr.Manjunatha.R, Development of Efficient Intra Frame Coding in Advanced Video Standard Using Horizontal Prediction Mode, International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 2, February 2013, pp 192-196.

[3] H.264 video compression standard, New possibilities within video surveillance, Axis Communications, White paper, 2008.

[4] Amruta Kiran Kulkarni, Implementation of Fast Inter-Prediction Mode Decision in H.264/AVC Video Encoder, Master Thesis, May 2012.

Authors

Mohammed H. AL-Jammas (Jun'02) born in 1966 in Mosul-Iraq. He awarded BSc in Electronic and Communication Engineering from the University of Mosul, Mosul-Iraq in 1988. Next, he awarded the MSc in Communication from the University of Mosul, Mosul-Iraq in 1994, and PhD in Computer Engineering from the University of Technology, Baghdad-Iraq in 2007. From 2002-2006, Dr. Mohammed worked with the University of Technology in Baghdad. From 2007, he acts as an Assistance dean of the College of Electronics Engineering at the University of Mosul.

Through his academic life he published over 7 papers in field of computer engineering, and information security.

Noor N. AL-Sawaf (August'08) born in 1988 in Mosul-Iraq. she awarded BSc in Electronics Engineering from the University of Mosul, Mosul-Iraq in 2010. Next, she awarded the MSc in Electrical from the University of Mosul, Mosul-Iraq in 2014.

Through her academic life she published 1 papers in field of image and video compress .