Interframe compression using Adaptive Discrete Cosine Transform (ADCT)

Ebtesam Fadhel Kanger,
University of Baghdad /College of Science / Astronomy and Space department
E.mail: shahab_alfrhan@yahoo.com

ABSTRACT

Image and video image compression techniques are paramount in the development of digital image and video systems. It is essential to develop compression methods which can both produce high compression ratios and preserve reconstructed quality in order for the creation of high quality, affordable image and video products. In this rsearch was produce New algorithm to optimize the compression performance by DCT classification scheme for the image blocks is adopted, a bit allocation matrices them, designed to assign a given number of bits for each retained transformed coefficients. Also was introduce New Three Step Search (NTSS) for motion estimation technique was produced.

الخلاصة

تقنيات ضغط الصور والصور الفديوية تعتبر الاهم في تطوير منضومه الصور الرقميه والفديويه. الاساسي في تطوير طرق الضغط هو الحصول على نسبة ضغط مع المحافظه على جودة الصوره المسترجعه . في هذا البحث قدمت خوارزميه لانجاز عملية الضغط تم التركيز فيها على مناقشة تقنية ترميز تشذيب الزمر وتحويل الجيب تمام المنفصل (DCT) .

تم اعتماد أسلوب جديد للحصول على أمثل قيمة للأداء بواسطة تصنيف الـ DCT لزمر الصور. تم تصميم مصفوفات – Bit لترميز عدد معين من الـ Bits لكل معامل تحويل. وكذلك قدمت طريقه جديده لتخمين الحركه الثلاثية الخطوه.

1- Introduction

In order to develop an efficient coding method for video images, the characteristics of the video images and the wide range of the image data must be taken into consideration [1]. Consequently, the inter–pixels correlation among adjacent pixels should be considered figure (1) illustrated the video image compression process flow. Therefore, in our research an Adaptive Discrete Cosine Transform (ADCT) technique is introduced to code video images. In this coding method, the variation consists of different regions which involve different degree of details, therefore, coding all image regions with the same number of bits is not an efficient coding method because it will produce high distortion in regions that are highly detailed. To overcome this problem many adaptive transform coding techniques have been introduced [2].

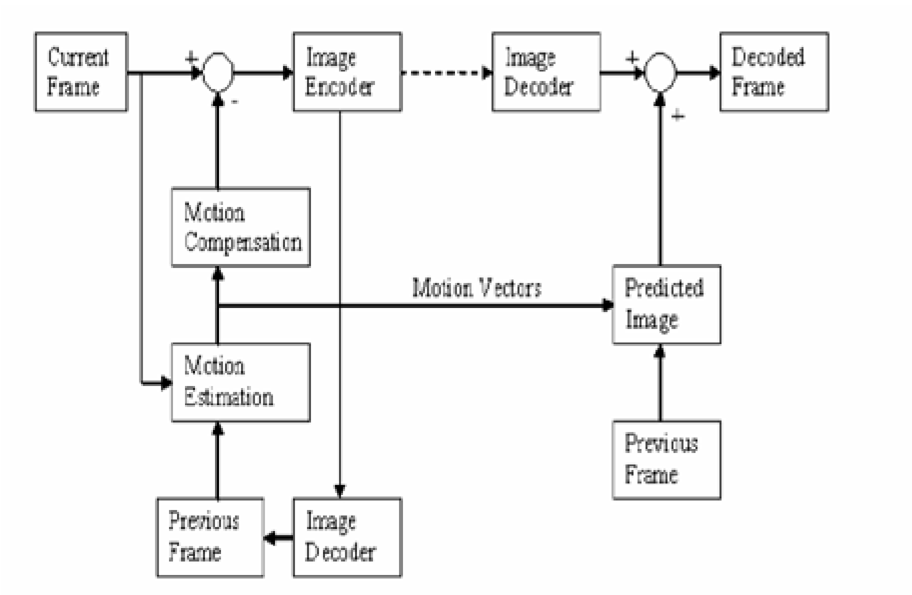


Figure .1 Video compression process flows [5].

This coding method is adopted in our work with many adaptations. Coded images, by this compression technique are divided into small non-overlapping block (e.g. 4×4, 8×8, or 16×16). Each block, is then divided again into smaller sub-block. Then, each block transformed using DCT transformation method, and classified into one of several classes, depending on their local activities. The total image bit – rate can then determined by computing the number of bits assigned for each image's block. Number of bits assigned for each block class is presented in the form of an array, usually, referred as the bit-assignment-matrix. The retained transformed coefficients, should first be classified according to the energy of the transferred block, then quantized differently by the Bit – Assignment – Matrix method [3].

The degree of data reduction achieved by a compression process or algorithm is called compression ratio, given by:

$$Compression \quad ratio = \frac{Uncompressed \quad File \quad Size}{Compressed \quad File \quad Size} \quad ..... \quad (1)$$

Therefore, measuring compression ratio as a function of the reproduced image quality define the efficiency characteristics of the adopted lossy compression technique, rather than a compression ratio in the decompressed pictures [1].The equivalent definition to the mean square $(SNR)_{ms}$ is the Peak Signal-to-Noise Ratio (PSNR) defined as [2,3] The PSNR can be represented in decidable (dB) unit as:

$$\text{PSNR} = 10\log_{10}\left[\frac{gray \quad scale \quad of \quad image}{MSE}\right]^2 \quad [1,3] \quad ..... \quad (3)$$

2- Middle Ac- Energy and Image Blocks Classification

The image blocks classification process is followed the DCT operation, in which transformed blocks are classified into n-classes (*in our presence work number of adopted classified blocks was n=4*). However, the classification process is based on the local activity or on the middle AC–energy (*MACE*) of each transformed block, i.e. if the size of the transformed blocks was P×P, and the size of the transformed sub-blocks was q×q, then the MACE, can be given by : -

$$MACE = \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |F(u,v)|^2 - \sum_{u=1}^{q-1} \sum_{v=1}^{q-1} |F(u,v)|^2 \quad \cdots\cdots(1)$$

The maximum and the minimum values of the computed MACE values should then be determined. However, in our present work, the histogram of MACE values is used to differential between block's classes. The classified blocks are then presented in a matrix, called a Bit-Classification-Map (BCM) which takes a size equal to (image size/ block size). This matrix must be coded and transmitted or stored in the overhead book information library.

3- Bit- Assignment – Matrix (BAM)

After generating the BCM, the average bit–rate for the coded image can be specified, automatically, by assigning a given number of bits for each classified block. A number of n-matrices must be simulated each represent the bit–rate distribution with each classified block. The size of each bit–assignment matrix is the same as the size of transformed blocks. However, number of bits assigned for each transformed coefficient, within each classified block is specified by adopting the ensemble average of the variance of each coefficient within each class type, it has been calculated according to the following formula : [5]where mu is the mean value of coefficients for each class.

The bit-assignment matrix (*BAM*) can now be generated by assigning of fixed number of bits (i.e.8.bits) for each DC–coefficients and for all blocks classes. The AC–coefficients (i.e. F(u,v)) assigned number as bits (NB (u,v)) determined of follows : - [5]

$$\sigma_L^2(u,v) = \frac{1}{K}\sum_{L=1}^{K} \left| F(u,v)^2 \right| - mu \qquad \ldots\ldots \quad (2)$$

$$NB(u,v) = \begin{cases} 0 & DS \ge \sigma^2(u,v) \\ \frac{1}{2}\ Log_2\ \frac{\sigma^2(u,v)}{DS} & DS < \sigma^2(u,v) \end{cases} \qquad \cdots\cdots(3)$$

In fact number of bits assigned for each Ac–coefficients is considered so as to be not greater than 7; i.e. NB (u,v) ≤ 7.

Finally, The distortion factor DS is a positive real number proportional to the discarding amount of the coefficients. The exact value of DS is specified by suggesting an initial value and, then iteratively, changing this value till we reach the desired bit–rate or compression ratio for coded image.

4-Quantization

Practically, the rounding of to the nearest integer operation could be represented as the simplest quantization approach. Generally, the quantization operation is performed by the following formula: -

$$Q_C(u,v) = Round\ off \left[\frac{(C(u,v) - MN)}{(MX - MN)} \times \left(G(u,v)\ -1\right) \right] \quad \cdots\cdots(4)$$

Where QC(u,v) is the output quantized transformed coefficients, while, C(u,v) is the input coefficients. MN is the minimum value of the input coefficients, MX is the maximum value of input coefficients, and G(u,v) is the number of quantization levels, suggested in the BAM [1].

In quantizing DC-coefficient, we fined that it is more suitable to normalize their values, by dividing them on a “NC” (normalized fa ctor NC >1),in dequantization operation the inverse process is performed by multiply each quantized DC-value by "NC" The AC-coefficients are quantized as follows; determine MN (*minimum*) and MX (*maximum*) value over all block. Extract the number of bits that were assigned to quantize these coefficients, given by :

$$G(u,v) = 2^{NB(u,v)} - 1 \quad \ldots\ldots(5)$$

Where NB(u,v) bits assignment matrix element, G(u,v) is the resultant levels. Apply the equation(4) to determine the quantize coefficients values. Finally, the dequantization of these coefficients is obtained, using the following formula : -

$$C\ (u, v)=(MX-MN) \times Qc\ (u, v) / G\ (u, v) + MN \quad \ldots\ldots(6)$$

5-Overhead Information

The overhead information can be considered as a header file for the coder and decoder. Without the information involve in this header the decoding process can not be achieved. Therefore, this information book must be transmitted first to make the decoding process possible. In this work, the information involved in the overhead information are : -

- The image size (X × X) (SI) .
- Images block size (P ×P) (BS); same method is followed

for saving this size information as that for image size, since larger block size that we used is 16 x 16, so we coded block size by 4 bits.

- Distortion parameter DS; which is presented by 8.bits.

- Bit Classification Map (BCM) matrix (MP); In this work only four classes were adopted, each of their element was coded with 2 bits consequently, number of bits which were preserved for these matrices are presented by :

$$MP = \frac{X \times X}{P \times P} \times 2 \qquad \ldots\ldots(7)$$

- Bit Assignment Matrices (BAMs) (BA); the range of coded value involved in the BAM is $0 \rightarrow 7$, therefore, only 3-bits were assigned to represent each of these elements. However; number of elements within each BAM matrix was (P×P–1), therefore, the total number of bits which were required to code the adopted 4- BAM is :

$$BA = 4\,[P \times P - 1] \times 3 \qquad (8)$$

Thus, the total number of bits which were preserved to represnt the whole information of the header is given by :

$$TB = SI + BS + DS + MP + BA \qquad \ldots\ldots\ \ldots\ldots(9)$$

$$AB = \frac{TB}{X \times X} \qquad \ldots\ldots\ (10)$$

Motion Estimation and Compensation

Successive video frames may contain the same objects (still or moving). Motion estimation examines the movement of objects in an image sequence to try to obtain vectors representing the estimated motion. Motion compensation uses the knowledge of object motion so obtained to achieve data compression. In interframe coding, motion estimation and compensation have become powerful techniques to eliminate the temporal redundancy due to high correlation between consecutive frames [4].

There are two mainstream techniques of motion estimation: pel-recursive algorithm (PRA) and block-matching algorithm (BMA) .in this search we was produce the(BMA) [6].

frame Figure .2 illustrates a process of block-matching algorithm. I

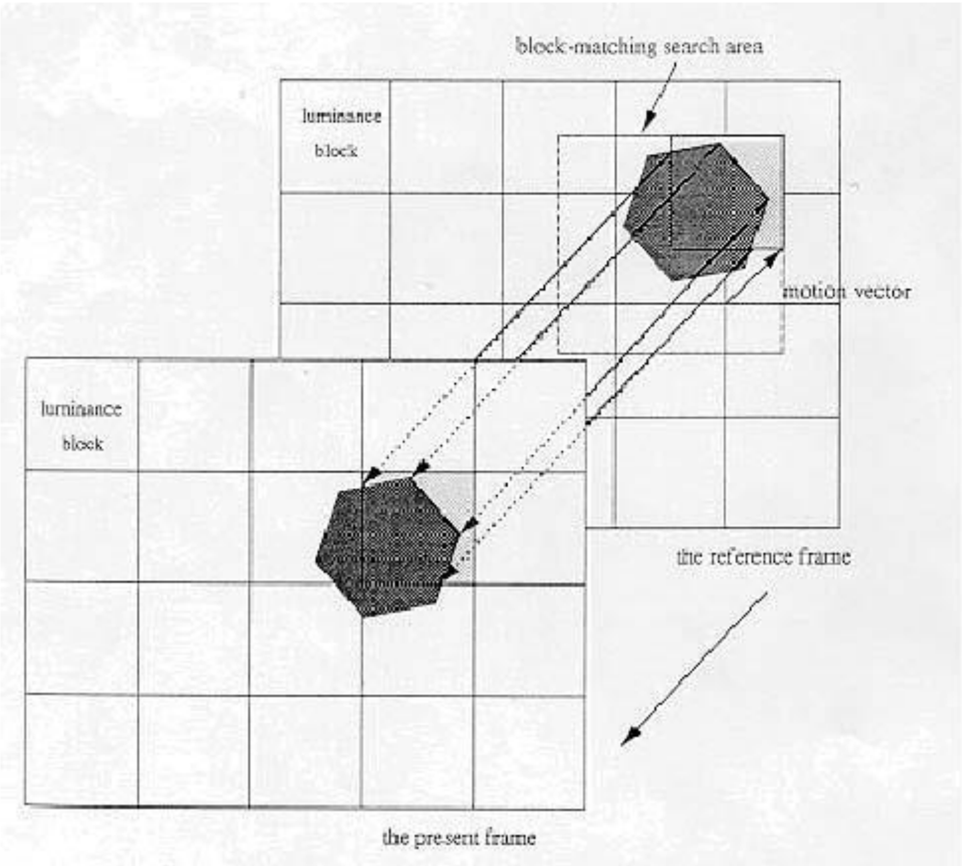


Figure. 2 illustrates a process of block-matching algorithm.

There are various cost functions, of which the most popular and less computationally expensive is mean Absolute Difference (MAD) given by equation (4). Another cost function is Mean Squared Error (MSE) given by equation (5).The Mean Absolute Difference (MAD), defined as :

$$MAD(dx, dy) = \frac{1}{mn} \sum_{i=-n/2}^{n/2} \sum_{j=-m/2}^{m/2} F(i, j) - G(i + dx, j + dy) \quad .... (4)$$

Where $F(i,j)$ represents an $(m \times n)$ macro block within the current frame, $G(i,j)$ represent the corresponding macro block within reference frame (past or future), *(dx,dy)* a vector representing the search location[3,7] .

The Mean-Squared Difference (MSD) cost function is defined as :

$$MSD(dx,dy)=\frac{1}{mn}\sum_{i=-n/2}^{n/2}\sum_{j=-m/2}^{m/2}[F(i,j)-G(i+dx,j+dy)]^2 \quad .... (5)$$

7. New Three Step Search (NTSS) for motion estimation:

NTSS [4] improves on TSS results by providing a centerbiased searching scheme and having provisions for half waystop to reduce computational cost. It was one of the firstwidely accepted fast algorithms and frequently used for implementing earlier standards like MPEG 1 and H.261The TSS uses a uniformly allocated checking pattern for motion detection and is prone to missing small motions. The NTSS process is illustrated graphically in Fig 4. In the first step 16 points are checked in addition to the search origin for lowest weight using a cost function [7]. Of these additional searchlocations, 8 are a distance of S = 4 away (similar to TSS) and the other 8 are at S = 1 away from the search origin. If the lowest cost is at the origin then the search is stopped right here and the motion vector is set as (0, 0). If the lowest weight is at any one of the 8 locations at S = 1, then we change the origin of the search to that point and check for weights adjacent to it. Depending on which point it is we might end up checking 5 points or 3 points.. The location that gives the lowest weight is the closest match and motion vector is set to that location. On the other hand if the lowest weight after the first step was one of the 8 locations at S = 4, then we followthe normal TSS procedure. Hence although this process might need a minimum of 17 points to check every macro block, it also has the worst-case scenario of 33 locations to check [8].

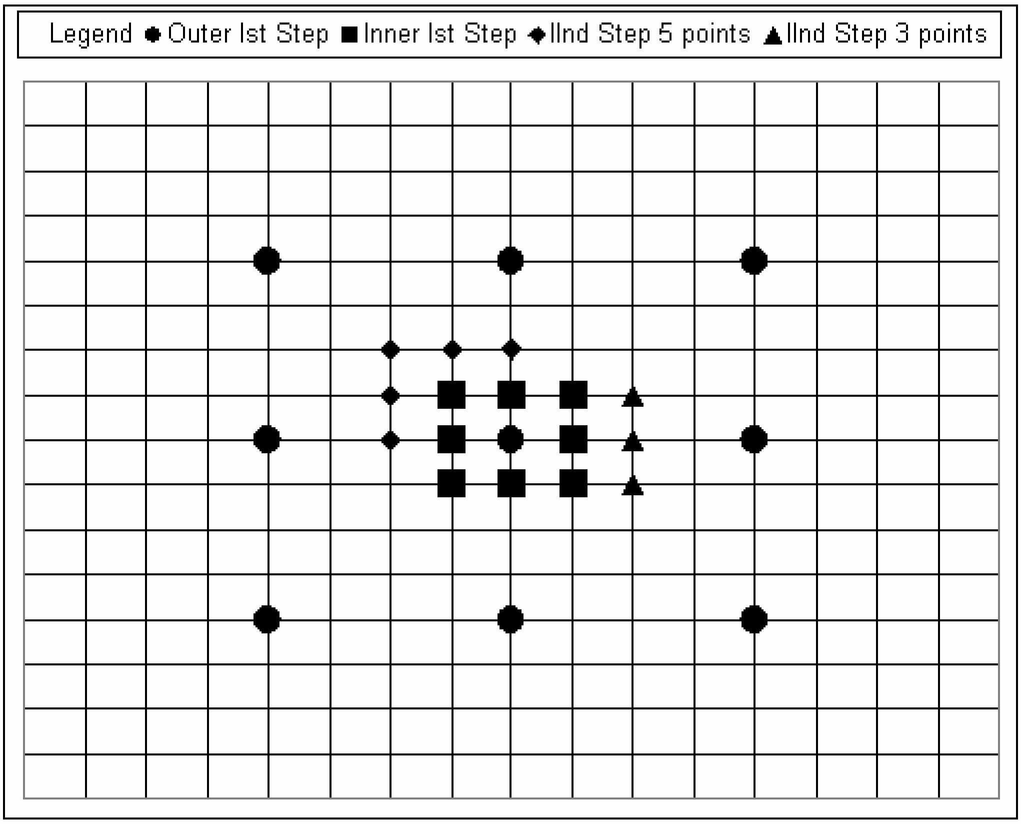


Figure.3 New Three Step Search block matching: Big circles are checking points in the first step of TSS and the squares are the extra 8 points added in the first step of NTSS. Triangles and diamonds are second step of NTSS showing 3 points and 5 points being checked when least weight in first step is at one of the 8 neighbors of window center.

8-Interpolation (IBBP):

Interpolation is a simple yet efficient and important method in image and video compression . In image compression, we may only transmit, say, every row. We then try to interpolate these missing rows from the other half of the transmitted rows in the receiver . In this way, compress the data to half . since the interpolation is carried out within a frame, it is referred to as spatial interpolation . In video compression, for instance. In the receiver we may try to interpolate the dropped frames from the transmitted frames . This strategy immediately drops the transmitted data to one third [9]. I or basic concepts of zero – order interpolation, bilinear interpolation, and polynomial interpolation, readers are referred to signal processing texts.

In temporal interpolation, the zero – order interpolation means creation of a frame by copying its nearest frame along the time dimension.

The IBBP frame model is a new feature of video coding which consists of three pictures (one p-picture and two BB-pictures). The P-pictures are predicted from previous decoded (I) [9] .

The BB-pictures are bidirectional predicted both from the previous decoded (I) and P-picture . This model has been used in this work. Figure (4) represent this model

$$f(x,y,t) = \frac{I_2}{I_1 + I_2} f(x,y,t_1) + \frac{I_1}{I_1 + I_2} f(x,y,t_2) \qquad \text{……….(11)}$$

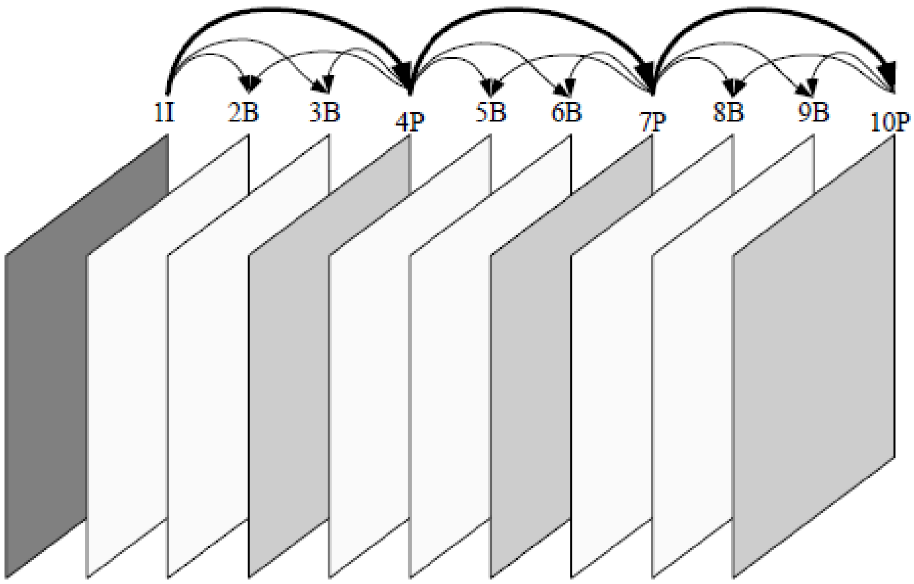


Figure.4 The BB-pictures are bidirectional predicted both from the previous decoded (I) and P-picture[3]

Experimental Result:

In this paper, an Adaptive Discrete Cosine Transform (ADCT) technique is introduced to code video images.. The algorithm is of low computational complexity.

The new efficient video image compression using (ADCT) with different block size is applied to gray Man the image size (128x128) and reference block size is (4x4)(8x8) and (16x16) pixel with different Distortion parameter DS(20, 40,60,80) we was found from the experimental result the best Distortion parameter DS is(60) . Also was applied adaptive search for motion Estimation technique and reference block size is(8x8) pixel .We will show the usefulness of the motion estimation in interframe compression. The compression and motion basic. In performance tests of the proposed algorithm apply block truncation and New Three Step Search (NTSS) for motion estimation.

. The figure (5) represent the origin (man) images. Figure (6) represent the reconstruction (I), (p) frame when we apply (ADCT) technique block size (4x4) Distortion parameter DS(60). Figure (7) represent reconstruction images (I) and (P) and block different when we apply (NTSS) used (MAD) criteria with thr.4, in figure shown the block different between (I),(P) frames . Figure (8) represent reconstruction images (I) and (P) when we apply used (MSD) criteria with thr.4, in figure shown the block different between (I),(P) frames . Figure (9) represented the frames (I), (B1),(B2) and (P) respectively Where BB-pictures are bidirectional predicted both from the previous decoded (I) estimation are performed using the visual

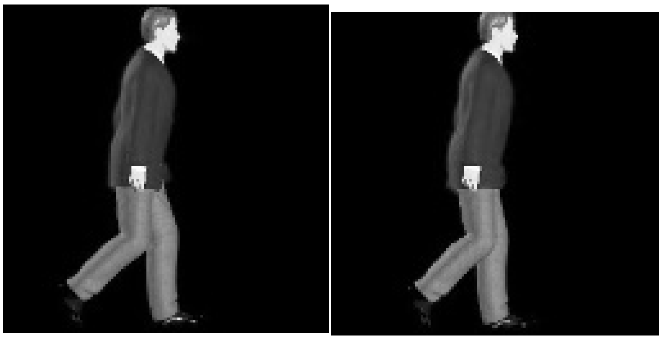
and P-picture . Table (1) numerical result when we apply (ADCT) technique block size (4x4). Table (2) numerical result when we apply (ADCT) technique block size (8x8). Table (3) numerical result when we apply block (ADCT) technique block size (12x12), Table (4) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size (4x4) DS(60). Table (5) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size (8x8) DS(60). Table (6) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size(16x16)DS(60)

Conclusions:

From the result we were found the best image reconstruction at block size (4x4) and DS (60) with good value of compression ratio. We have presented motion estimation technique, New Three Step Search (NTSS), The experimental results show that this new method can significantly reduce the computational complexity and low search time while achieving comparable performance for many video sequences . Also we have found the criteria (MAD) was gave higher number of block different than criteria(MSD) then low C.R when we used(MAD) than (MSD).Threshold (4) was gave highest number of block different than other thresholds.

(I) **(P)**

Figure (5) represent the original images of man (I), (P) frame

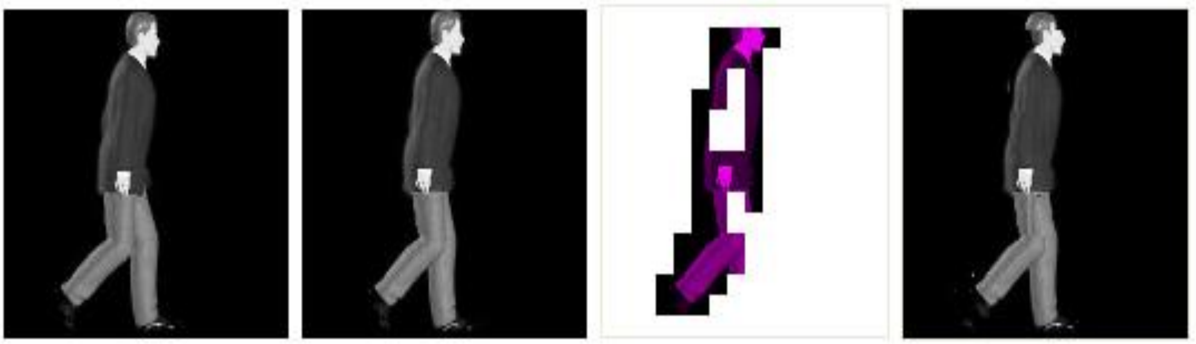


(I) **(P)**

Figure (6) represent the reconstruction images (I), (P) frames when apply(ADCT) technique**,** block size (4z4)**.**

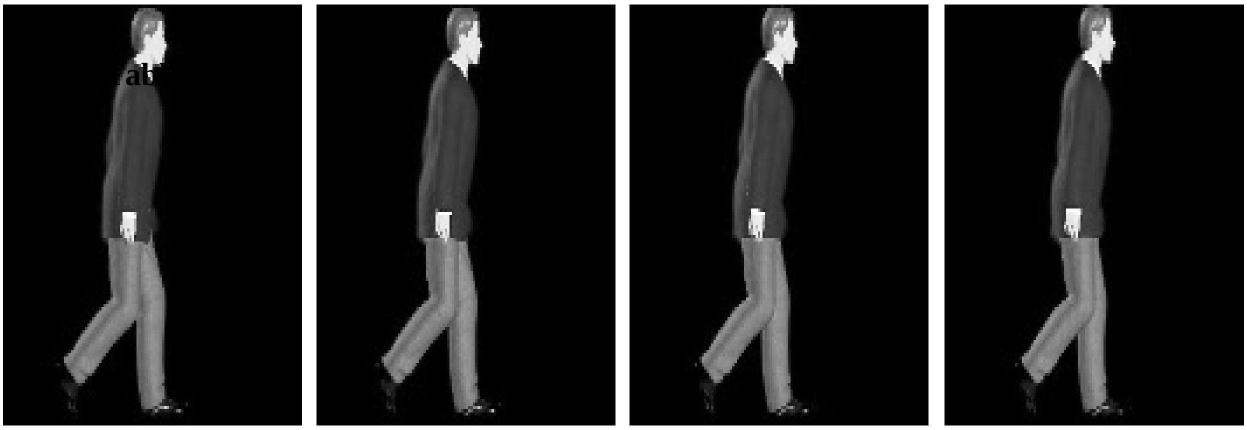
(I) **(P)** **Block different** **New (P)**

Figure (7) represent the (I), (P) frame when we apply (NTSS) for motion estemation used (MAD) ,with thr.4



(I) **(P)** **Block different** **New (P)**

Figure (8) represent the (I), (P) frame when we apply (NTSS) for motion estimation used (MSD) ,with thr.4



I B B P

1 2

Figure (9) represent the (I), (B1),(B2) and (P) frame when we apply (ASMET) used (MAD) ,with thr.4

Table (1) numerical result when we apply (ADCT) technique block size (4x4)

DS	PSNR	MSE	Comp. ratio
40	26.2	123.59	4.5
50	26	125.54	4.5
60	25.6	139.82	5.81
70	24.5	171.9	7.9
80	23.40	18.66	8.1

Table (2) numerical result when we apply (ADCT) technique block size (8x8)

DS	PSNR	MSE	Comp. ratio
40	25.7	169.41	7
50	25.5	173.16	7.24
60	23.8	202.08	9.57
70	23.78	206.13	9.88
80	22.93	211.02	9.92

Table (1) numerical result when we apply (ADCT) technique block size (16x16)

DS	PSNR	MSE	Comp. ratio
40	22.3	291.62	10.5
50	22.1	297.96	11.92
60	21.3	315.10	13.2
70	21.11	319.34	14.8
80	20.91	322.01	15.2

Table (4) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size (4x4) DS(60).
Comp. ratio (I)= 5.81 PSNR=25.6

threshold	Tim of search(sec)	No. of block change	Over all comp. ratio	PSNR (B1)	PSNR (B2)	PSNR (P)
4 (MAD)	0.411	90	19.5	24.7	24.4	25.5
8 (MAD)	0.232	85	22.5	24.7	24.4	25.5
12 (MAD)	0.161	55	29.2	24.7	24.4	25.5
4 (MSD)	0.411	75	21.9	24.1	23.5	24.7
8 (MSD)	0.232	54	34.8	24.1	23.5	24.7
12 (MSD)	0.161	42	38.7	24.1	23.5	24.7

Table (5) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size (8x8) DS(60).
Comp. ratio (I)=9.57 5.2 PSNR=23.8

threshold	Tim of search(sec)	No. of block change	Over all comp. ratio	PSNR (B1)	PSNR (B2)	PSNR (P)
4 (MAD)	0.411	90	19.5	22.7	22.4	23.2
8 (MAD)	0.232	85	22.5	22.7	22.4	23.2
12 (MAD)	0.161	62	29.2	22.7	22.4	23.2
4 (MSD)	0.411	55	21.9	22.1	22.1	22.6
8 (MSD)	0.232	54	34.8	22.1	22.1	22.6
12 (MSD)	0.161	42	38.7	22.1	22.1	22.6

Table (6) numerical result when we apply motion estimation (NTSS) with different criteria (MAD) and (MSD) with (ADCT) technique block size (8x8) DS(60).
Comp. ratio (I)=13.2 PSNR=21.3

threshold	Tim of search(sec)	No. of block change	Over all comp. ratio	PSNR (B1)	PSNR (B2)	PSNR (P)
4 (MAD)	0.411	90	19.5	20.7	20.4	21.2
8 (MAD)	0.232	85	22.5	20.7	20.4	21.2
12 (MAD)	0.161	62	29.2	20.7	20.4	21.2
4 (MSD)	0.411	55	21.9	20.3	20.1	20.9
8 (MSD)	0.232	54	34.8	20.3	20.1	20.9
12 (MSD)	0.161	42	38.7	20.3	20.1	20.9

Conclusions:

From the result we were found the best image reconstruction at block size (4x4) and DS (60) with good value of compression ratio. We have presented motion estimation technique, New Three Step Search (NTSS), The experimental results show that this new method can significantly reduce the computational complexity and low search time while achieving comparable performance for many video sequences . Also we have found the criteria (MAD) was gave higher number of block different than criteria(MSD) then low C.R when we used(MAD) than (MSD).Threshold (4) was gave highest number of block different than other thresholds.

Refrences:

1- R.C. Gonzalez and R.E. Woods, "Digital Image Processing",Prentice Hall, 2[3d] edition, 2008.

2- Zhen Li"EW METHODS FOR MOTION ESTIMATION WITH APPLICATIONS TO LOW COMPLEXITY VIDEO COMPRESSION"A Thesis Submitted to the Faculty of Purdue University In Partial Fulfillments of the Requirements for the Degree Of Doctor of Philosophy December 2005.

3-**1.**Eric J.Balster " Video compression and Rate control methods Based on the wavelet Transform " Dissertation Degree of Doctor .The Ohio state University .,2004.

4- Nishat Kanvel, and Elwin Chandra Monie"Adaptive lifting based image compression scheme fornarrow band transmission system" International Journal of Physical Sciences Vol. 4 (4), pp. 194-164,April, 2009

5- Aroh Barjatya, *Student Member, IEEE"* Block Matching Algorithms For Motion Estimation" ,2004

6- Toshiaki Kondo_, Waree Kongprawechnon **"** Robust motion estimation methods using gradient orientation information" doi: 10.2306/scienceasia1513-1874.2009.35.196*ScienceAsia* 35 (2009)**.**

7-Dr.V.V.Gohokar Professor Electonics Dr.V.N.Gohokar Professor

"Adaptive - Selection of Motion Estimation Block Size forRate-Distortion Optimization" International Journal of Computer Applications (0975 – 8887)Volume 17– No.4, March 2011

8- Svetislav Momcilovic, Nuno roma and Leonel Sousa "An ASIP Approach for Adaptive AVC Motion Estimation ",PRIME 2007 .

9-**.** Yun, Q "Image and Video Compression for multimedia engineering" ,New jersey institute of technology Newark, NJ,2000.