Design & Implementation of Estimating Motion & Reconstruction Using Pixel Processing

Akhil Mahajan[1], Dr. Kamal Sharma[2], Deepak Kumar[3]

[1]Student, M. Tech, ESEAR, Ambala

[2]Professor, Dept. of ECE, E-Max group of Institutions, Ambala

[3]Assistant Professor, Dept. of ECE, E-Max group of Institutions, Ambala

Abstract— **Estimating the motion recognized in a set of images or an image sequence is aimportant problem in both image and video handling and computer vision. For instance, assessing the specious motion in a video sequence provides necessary information for many applications including navigation, industrial control, 3-D shape reconstruction, object recognition, object tracing, and automatic image sequence analysis. In image and video processing, the estimation of motion plays a energetic role in video compression as well as multi-frame image enhancement. In this, a new technique called Motion estimation and reconstruction using Pixel Processing is proposed. This technique is based on pattern search algorithm, in which size is dynamically determined for each pixel, based on the mean of two motion vectors of the neighbouring macro-blocks. The proposed technique attempts to decrease the computation time between the current macro-block and the macro-blocks in the reference frame.**

Keywords— **Motion detection, motion estimation, pixel processing, Pel approach.**

I. Introduction

Comparing digitally stored video sequences and those "stored" on celluloid and considering the fact that data-storage or data-transmission capacity still is restricted in computer technology, this comparison shows necessity of compressing the video-data: Using same approach to store videos digitally as they are in the classic way on celluloid would require at least 25 still images per second. A high-quality 90 minute movie with a resolution of 720*576 pixels and a 24-bit color depth per pixel could require of over 156 GB of storage capacity. Transmitting this amount of data over Internet is unreasonable, particularly when real-time performance is needed. This uncompressed video needs a transmission bandwidth of over 237 MBit/s. Similar problems occur when storing the data to disc – only very few memory devices have necessary capacity [1].

In analogous with planetary applications, digital image processing techniques initiated in late 1960s and early 1970s to be used in medical imaging, remote Earth resources explanations, and astronomy. The creation in early 1970s of computerized axial tomography (CAT), also named computerized tomography (CT) for short, is one of most important actions in application of image processing in medical analysis. Computerized axial tomography is a procedure in which a ring of detectors encloses an object (or patient) and an X-ray source, concentric withsensor ring, rotates about the object. The X-rays pass through object and are collected at opposite end by the corresponding detectors in ring. As the source revolves, this procedure is repeated. Tomography contains procedures that use sensed data to construct an image that represents a "slice" through object [2].

Motion of object in a direction perpendicular to ring of detectors creates a set of such slices, which constitute a three-dimensional (3-D) rendition of inside of article. Tomography was developed independently by Sir Godfrey N. Hounsfield and Allan Cormack, who shared 1979 Nobel Prize in Medicine for their invention. It is exciting to note that X-rays were discovered in 1895 by Wilhelm Conrad Roentgen, for which he acknowledged 1901 Nobel Prize for Physics. These two establishments, nearly 100 years apart, led to some of most active application areas of image processing today [3].

In video sequences, motion is a key source of information. Motion arises due to moving objects in 3D scene, as well as camera motion. Apparent motion, also known as optical flow, captures the resulting spatial-temporal variations of pixel intensities in successive images of a sequence. The purpose of motion estimation techniques is to recover this information by analyzing image content. Efficient and accurate motion estimation is an essential component in domains of image sequence analysis, computer vision and video communication [4].

As far as video coding is concerned, compression is attained by exploiting data redundancies in both spatial and temporal dimensions. Spatial redundancies reduction is largely attained by transform-coding, e.g. usingDiscrete Cosine Transform (DCT) or the Discrete Wavelet Transform (DWT), which efficiently compacts signal energy into a few significant coefficients. In turn, temporal redundancies are reduced by means of predictive coding. Observing that sequential correlation is maximized along motion trajectories, motion compensated prediction is used for this purpose. In this context, main objective of motion estimation is no longer to find 'true' motion in the scene, but rather to maximize compression efficiency. In additional words, motion vectors should deliver a precise prediction of signal. Moreover, the motion info should enable a compact representation, as it has to be conveyed as overhead in compressed code stream. Efficient motion estimation is a key to achieve large compression in video coding applications such as TV broadcasting, Internet video streaming, digital cinema etc.

The paper is ordered as follows. In section II, we discuss correlated work with recognition of fingerprint images. In Section III, It definesbasic motion estimation system, its technique and block matching technique.In Section IV, it describesproposed technique of motion estimation. After this, it designates the main results of this system.Finally, conclusion is explained in Section VI.

II. RELATED WORK

In literature, authors proposed a search technique based on conjugate directions, and another modest technique called one-at-a-time search based onrelationship of two methods, the latter technique is accepted on thebasis for further research. The accepted technique is compared with brute force search, current 2-d logarithmic search, and a modified version of it, for motion compensated estimate. To estimate the motion on a block-by-block basis by brute force requires extensive computations. These motion estimation techniques are applied to video sequences, and their larger performance compared to the existing techniques is illustrated based on quantitative measures of the prediction errors [5].

Some Authors proposed that three-step search (TSS) algorithm has been extensively used as motion estimation technique in low bit-rate video compression applications, owing to its easiness and effectiveness. However, TSS uses a steadily allocated checking point pattern in its first step, which becomes ineffective for estimation of small motions. A new three-step search (NTSS) algorithm is suggested in this paper. The features of NTSS are that it pays a centre-biased point pattern in its first step, which is derived by making search adaptive to motion vector distribution, and a halfway-stop technique to decreasecomputation cost. Simulation results show that, as related to TSS, NTSS is much more robust, produces smaller motion compensation errors, and has a very well-matched computational complexity [6].

Some proposed a new four-step search (4SS) procedure with centre-biased checking point design for fast block motion estimation. Halfway-stop technique is active in new algorithm with searching steps of 2 to 4 and total number of checking points is specked from 17 to 27. Simulation results show that suggested 4SS performs better than well-known three-step search and has similar performance to new three-step search (N3SS) in terms of motion compensation faults. In addition, 4SS also decreasesworst-case computational requirement from 33 to 27 search points and regular computational necessity from 21 to 19 search points as related with N3SS [7].

Authors proposed a block-based grade descent search (BBGDS) algorithm to perform block motion estimation in video coding. The BBGDS assessesvalues of a given objective function starting from a small centralized checking block. The least within checking block is found, and gradient descent direction where minimum is expected to lie is used to determinesearch direction and position of new checking block. The BBGDS is compared with full search (FS), three-step search, one-at-a-time search (OTS), and new three-step search. Experimental results show that proposed technique delivers competitive performance with reduced computational complexity[9].

III. MOTION ESTIMATION

Motion estimation (ME) techniques have been successfully useful in motion compensated predictive coding for falling temporal redundancies. They belong to class of nonlinear predictive coding techniques. An effective representation of motion is serious in order to reach high performance in video coding. Estimation techniques should on one hand provide good prediction, but on other hand should have low computational load [8].

A video sequence can be considered to be a discretized three-dimensional prediction of real four-dimensional continuous space-time. The objects in real world may move, rotate, or deform. The movements cannot be detected directly, but instead the light reflected from the object surfaces and estimated onto an image. The light source can be moving, and the reflected light varies depending on the angle between a surface and a light source. There may be objects occluding light rays and casting shadows. The objects may be see-through (so that several independent motions could be observed at the same location of an image) or there might be fog, rain or snow blurring the observed image. The discretization causes noise into the video sequence, from which video encoder makes its motion estimations. There may also be noises in image capture device (such as a video camera) or in electrical transmission lines. A perfect motion model would take all the factors into account and find motion that has the maximum likelihood from observed video sequence [10].

Changes between frames are mainly due to movement of objects. Using a model of motion of objects among frames, the encoder approximatesmotion that happened between reference frame and current frame. This process is called motion estimation (ME) [4]. The encoder then uses this motion and information to move contents of reference frame to deliver a better prediction of current frame. This process is known as motion compensation (MC), and prediction so produced is called motion-compensated prediction (MCP) or the displaced-frame (DF). In this case, coded prediction error signal is called displaced-frame difference (DFD).

A. Block Matching Motion Estimation

Block-matching motion estimation is extensively used motion estimation method for video coding. Interest in this method was initiated by Jain and Jain and he suggested a block-matching algorithm (BMA) in 1981. The current frame is first divided into blocks of M × N pels. The process then undertakes that all pels within the block undergo same translational movement. Thus, same motion vector, d is allocated to all pels within the block. This motion vector is projected by searching for best match block in a larger search window pelscentred at same location in a reference frame [5].This procedure is based on a translational model of motion of objects between frames. It also assumes that all pels within a block undergo similar translational movement [11].

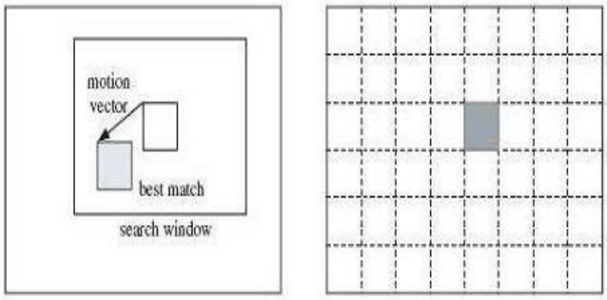


Figure 1: Block-Matching Motion Estimation [15]

B. *Matching Criteria for Motion Estimation*

Inter frame predictive coding is used to remove the hugeexpanse of temporal and spatial redundancy that happens in video sequences and helps in compressing them. In conventional predictive coding difference between current frame and the predicted frame is coded and transmitted. The better the prediction, smaller the error and hence the transmission bit rate when there is motion in a sequence, then a pel on same part of the moving object is a better prediction for the current pel [12].

1. *Block Size*

The important parameter of the BMA is block size. If the block size is smaller, it attains better prediction quality. This is due to a numeral of reasons. A smaller block size decreaseseffect of accuracy problem. In additional words, with a lesser block size, there is less possibility that block will comprise different objects moving in different directions. In addition, a smaller block size delivers a better piecewise translational approximation to non-translational motion. Since a smaller block size means that there are additional blocks (and consequently more motion vectors) per frame, this better prediction quality originates at the expense of a larger motion above. Most video coding standards use a block size of 16 × 16 as cooperation between prediction quality and motion overhead. A number of variable-block-size motion estimation methods have also been suggested in literature [14].

2. *Search Range*

The maximum allowed motion displacement dm, also identified as search range, has a direct influence on both computational complexity and prediction quality of BMA. A small dm results in poor compensation for fast-moving areas and therefore poor prediction quality. A large dm, on other hand, results in improved prediction quality but leads to an increase in the computational complexity. A larger dm can also result in longer motion vectors and consequently a slight increase in motion overhead [6]. In general, a maximum allowable displacement of dm = ±15 pels is enough for low-bit-rate applications. MPEG standard uses a maximum displacement of about ±15 pels, although this range can optionally be doubled with unrestricted motion vector mode [15].

3. *Search Accuracy*

Initially, the BMA was designed to estimate motion displacements with full-pel accuracy. Clearly, this limits performance of the algorithm, since in reality the motion of objects is entirely unrelated to the sampling grid. A number of workers in the field have proposed to extend BMA to sub-pel accuracy. For example, Ericsson demonstrated that a prediction gain of about 2 dB can be obtained by moving from full-pel to 1/8-pel accuracy. Girodoffered a sophisticated theoretical analysis of motion-compensating prediction with sub-pel accuracy. He calledresulting prediction gain the accuracy effect. He also presented that there is a “critical accuracy” beyond which possibility of further improving prediction is very small. He concluded that with block sizes of 16 × 16, quarter-pel accuracy is desirable for broadcast TV signals, whereas half-pel accuracy appears to be adequate for videophone signals [14].

IV. MOTION ESTIMATION USING PIXEL TECHNIQUE

A. *Proposed Algorithm*

The proposed steps for motion estimation using pixel approach are:

1. Interface WEBCAM with MATLAB
   a) Install image acquisition device.
   b) Retrieve hardware information.
   c) Create a video input object.
   d) Preview video stream (optional)
   e) Configure object properties.
   f) Acquire image data.
   g) Starting the video input object.
   h) Triggering the acquisition.
2. Image Reading
3. Image Enhancement
4. Image Conversion
5. Image Segmentation
6. Apply thresholding process
7. Feature Extraction
8. Detect the final motion
9. Reconstruction Original Image

Then results are compared with three step block based technique.

B. *Description of Proposed System*

1. *Webcam Interfacing and Video Processing*

It acquire the images from camera depends on lots of software. Creating the interface between computer and the camera can be done using lower level language like C or C++ will give lots of flexibility, but it will also invoke lots of work and background knowledge. MATLAB’s image acquisition toolbox has a variety of simple functions.

- Specifications of WEB CAM
  - ImageResolution: 640x480,352x288,320x240,176x144,160x120
  - Frame Rate : Up to 30 fps
  - Image Control : Brightness, contrast ,saturation
  - Environment : Indoor, Outdoor
  - I/O interface : USB 1.1,2.0

- Image Format : RGB_ 24

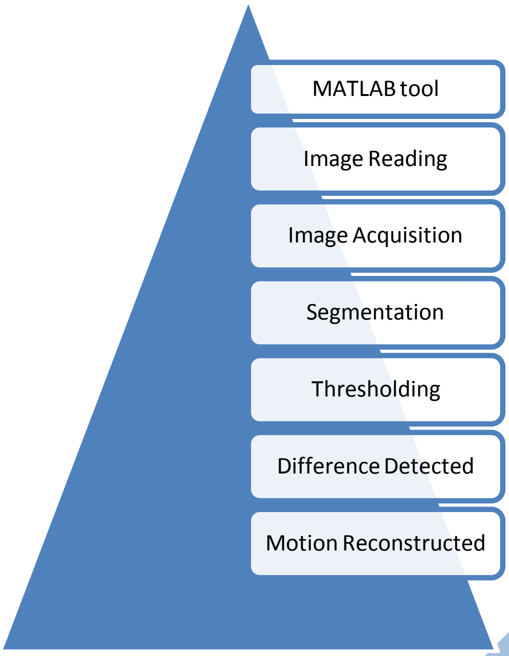


Figure 2: Block Representation of Proposed System

2. *Image Reading*

img= imread(strcat(a,num2str(i),b));

This takes the grey values of all pixels in grey scale image and puts them all into a matrix img. This matrix img is now a MATLAB variable, and so we can perform many matrix operations on it. In general the imread function reads pixel values from an image file, and returns a matrix of all pixel values.

3. *Image Segmentation*

Segmentation refers to process of partitioning an image into component parts, or into separate objects. Segmentation of an image entails the division or separation of the image into regions of similar attribute. The most basic attribute for segmentation is image luminance amplitude for a monochrome image and color components for a color image. Image edges and texture are also useful attributes for segmentation. There are a collection of ad hoc methods that have received some degree of popularity. Because the methods are ad hoc, it would be useful to have some means of assessing their performance. The objective of segment at ion is to classify each RGB pixel in a given image as having a colorin the specified range or not.

4. *Thresholding*

A grey scale image is converted into a binary (black and white) image by first choosing a grey level in original image, and then converting every pixel black or white according to whether its grey value is better than or less than T :

A pixel becomes: {white if its grey level is > T
{Black if its grey level is<T

Thresholding is a vital part of image segmentation, where we wish to isolate objects from background. It is also an significant component of robot vision. Thresholding can be done very simply in MATLAB.

5. *Displaying the Images*

Now to show image, the toolbox includes two image display functions: **imshow** and **imtool**. Imshow is toolbox's fundamental image presentation function. Imtool starts the Image Tool which presents an combined environment for displaying images and performing some common image processing tasks. The Image Tool delivers all image display capabilities of imshow but also offers access to several other tools for navigating and exploring images, such as scroll bars, Pixel Region tool, Image Information tool, and Contrast Adjustment tool. You can use either function to show an image. Here we use **imshow**.

6. *Detection of Motion*

1. Scan the entire area for finding out difference of two images.
2. Comparing that point row and column vise.
3. If there is difference in pixel value, then it takes the difference of images and hence motion is detected.
4. Then image is reconstructed using motion vectors.

7. *Motion Compensation*

Motion estimation creates a model of the current frame based on available data in one or more previously encoded frames ('reference frames'). These reference frames may be 'past' frames (i.e. earlier than the current frame in temporal order) or 'future' frames (i.e. later in temporal order). The design goals for a motion estimation algorithm are to model the current frame as accurately as possible. The encoder forms a model of the current frame based on the samples of a previously transmitted frame. The encoder attempts to 'cornpensate' for motion in a video sequence by translating (moving) or warping the samples of the previously transmitted 'reference' frame. The resulting motion-compensated predicted frame (the model of the current frame) is subtracted from the current frame to produce a residual 'error' frame.At the same time, the encoded residual is decoded and added to the model to reconstruct a decoded copy of the current Frame (which may not be identical to the original frame because of coding losses). This reconstructed frame is stored to be used as a reference frame for further predictions [13].

V. RESULTS

A.Implementation

MATLAB is one of a number of commercially accessible, sophisticated mathematical computation tools, which also comprise Maple, Mathematica, and MathCAD. Despite what followers may claim, no single one of these tools is "the best." Each has strengths and weaknesses. Each allows you to perform simple mathematical calculations. They differ in way they handle symbolic calculations and more complex mathematical procedures, such as matrix manipulation. For example, MATLAB (short for **Mat**rix **Lab**oratory) excels at computations relating matrices, whereas Maple excels at symbolic calculations. At a central level, you can think of these programs as sophisticated computer-based calculators. They can perform same functions as your scientific 'C' calculator—and **many more**. If you have a computer on your desk, you may find yourself using MATLAB instead of your calculator for even the simplest mathematical applications—for example, balancing your checkbook. In many engineering classes, the use of programs such as MATLAB to perform computations is replacing more traditional computer programming.

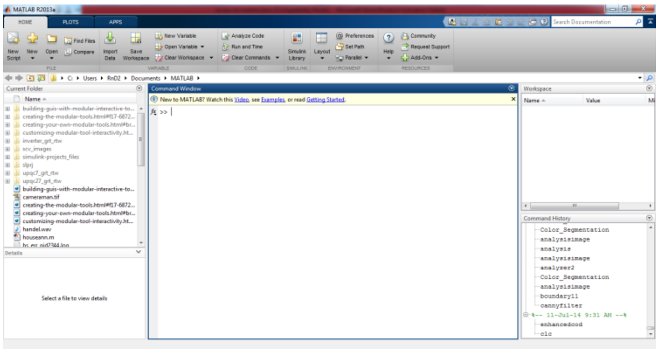


Figure 3: MATLAB Tool

B. *Proposed Results*

A video image is a projection of a 3-D scene onto a 2-D plane. A still image is a 'snapshot' of the 2D representation at a particular instant in time whereas a video sequence represents the scene over a period of time. Fig 4 shows the real time current image and its previous image.

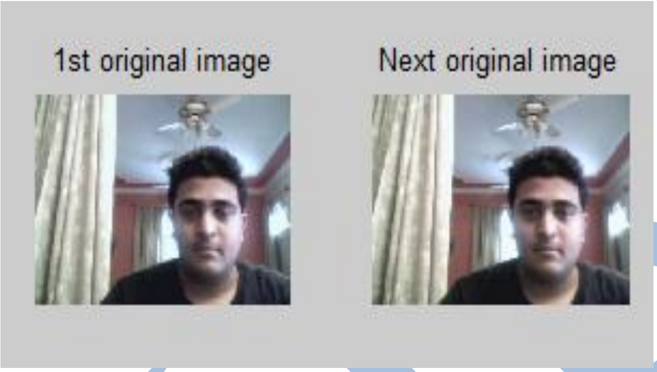


Figure 4: Input Original Images

After reading, it converts the grayscale image to a binary image. The output image substitutes all pixels in the input image with luminance better than level with value 1 (white) and replaces all other pixels with value 0 (black) as shown in fig 5.

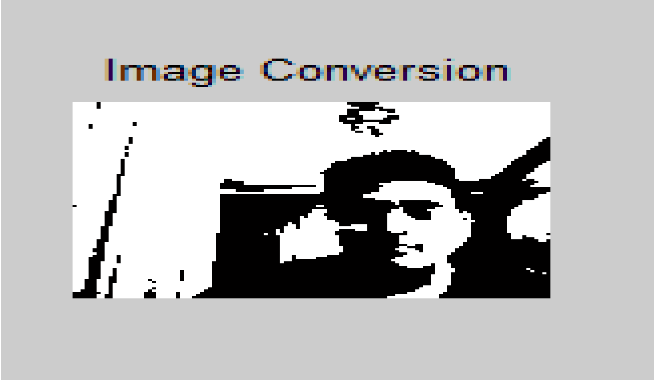


Figure 5: Image Conversion Output

In Figure 6 the motion estimation module creates a nodel by modifying one or more reference frames to match the current frame as closely ao possible (according to a *matching* criterion). And reconstructed image is shown in fig 7.

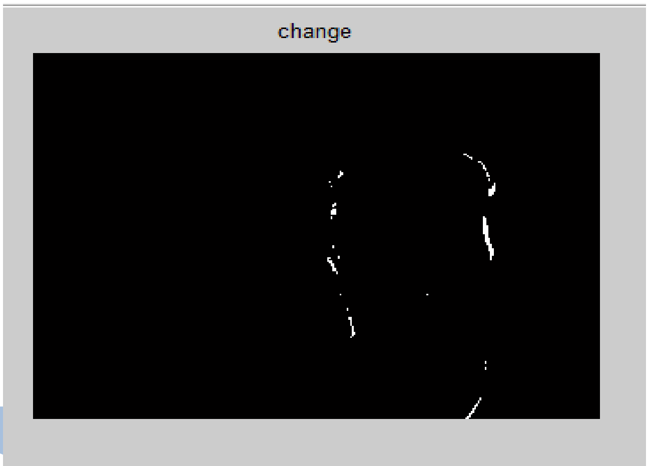


Figure 6: Motion Detection Output

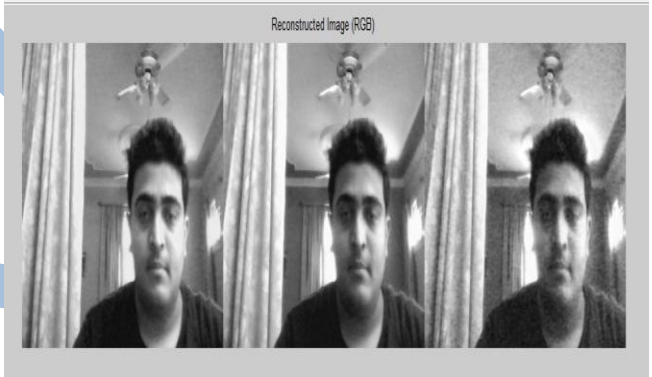
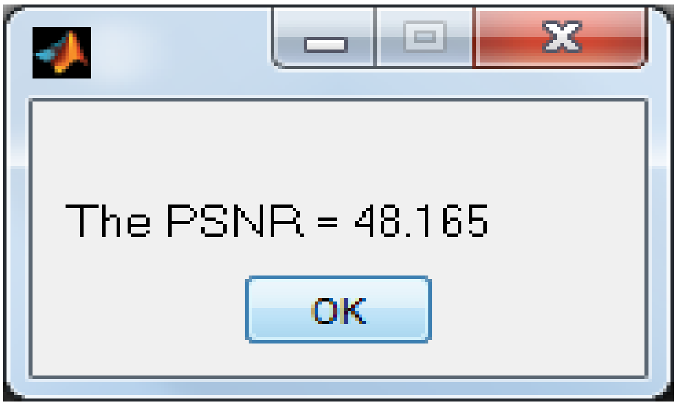


Figure 7: Reconstruction Image Output

The peak signal to ratio (PSNR) and MSE is an important parameter in image processing. Mean squared error provides a measure of the energy remaining in the difference block. Varying block sizes, or irregular- shaped regions, can be more efficient at matching true motion than fixed 16 x 16 blocks. The PSNR is shown in fig 8.



The PSNR = 48.165

Figure 8: PSNR Output of Pixel Approach

VI. CONCLUSION

Because of the Internet is more and more universal and the technology of multimedia has been progressed, the communication of the image data is a part in life. In order to employ effect in a limit transmission bandwidth, to convey the most, high quality user information .It is necessary to have more advanced compression method in image and data. Pel-recursive motion estimation is a well-established approach. The proposed algorithm can reduce the computational time as compared to block based technique.Motion Estimation (ME) and compensation techniques, which can exclude temporal redundancy between adjacent frames effectively, have been widely useful to popular video compression coding standards such as MPEG-4. The displacement of each picture element in each frame forms the displacement vector field (DVF) and its estimation can be done using at least two successive frames. The pixel based approaches depend upon intensity of image and its performance is affected by presence of noise. While block based techniques depends upon motion vectors and it has high computation time as compared to pixel approaches. The experiment results have shown that the proposed algorithm can reduce the computational time for the Pixel based technique required to determine matching macro block from the reference frame to the current macro-block as compared to three step search approach.

REFFERENCES

[1] R. Gonzalez, R. Woods and S. Eddins"Digital Image Processing Using Matlab", 2004, Prentice Hall.

[2] Gonzalez, R., Woods, R., "Digital Image Processing", 2nd Edition, Prentice-Hall, (2002).

[3] Acharya, T. and Ray, A.K, Image processing: Principles and Application, Wiley, and Hoboken, NJ, 2005.

[4] I. E. G. Richardson. H.264 and MPEG-4 video compression. Wiley , Chichester, England, 2003.

[5] Lai-Man Poet. , "Displacement measurement and its application in interframe image coding," IEEE Trans. Communications, vol.COM-29,PP. 1799–1807, Dec.1981.

[6] S.kappagantula and K.R.Rao, "Motion compensated interframe image Prediction", IEEE Trans. Commun 33,PP, 1011–l015,Sept, 1985.

[7] B.Liu and A.Zaccarin, New fast algorithms for the estimation of block motion vectors, IEEE Trans. Circuits Syst. Video technology, Vol.3, pp.440–445,Dec 1995.

[8] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 4, pp. 438442, Aug. 1994.

[9] Ram Srinivasan And K. R. Rao. Predictive Coding Based On Efficient Motion Estimation. Ieee Transactions On Communications, Vol. Com-33, No. 8, August 1985

[10] Renxiang Li, Bing Zeng, And Ming L. Liou. A New Three-Step Search Algorithm For Block Motion Estimation. Ieee Transactions On Circuits And Systems For Video Technology, Vol. 4, No. 4, August 1994.

[11] Lai-Man Po Wing-Chung Ma. A Novel Four-Step Search Algorithm For Fast Block Motion Estimation. Ieee Trans. Circuits Syst. Video Technol., Vol. 6, No.2, Pp. 313-317,Jun. 1996.

[12] Lurng-Kuo Li And Ephraim Feig. A Block-Based Gradient Descent Search Algorithm For Block Motion Estimation In Video Coding. Ieee Transactions On Circuits And Systems For Video Technology, Vol. 6, No. 3, August 1996.

[13] Jo Yew Tham, SurendraRanganath, MaitreyaRanganath, And Ashraf Ali Kassim. A Novel Unrestricted Center-Biased Diamond Search Algorithm For Block Motion Estimation. Ieee Transactions On Circuits And Systems For Video Technology, Vol. 8, No. 3, August 1998

[14] Ce Zhu, Xiao Lin, And Lap-Pui Chau. Hexagon-Based Search Pattern For Fast Block Motion Estimation. Ieee Transactions On Circuits And Systems For Video Technology, Vol. 11, No. 5, May 2002

[15] Chun-Ho Cheung, Member, Ieee, And Lai-Man Po, Member, Ieee. Novel Cross-Diamond-Hexagonal Search Algorithms For Fast Block Motion Estimation. Ieee Transactions On Multimedia, Vol. 6, No. 1, February 2005.