Efficient Hybrid Search for Motion Estimation

Deepak J. Jayaswal, *member IEEE* and Mukesh A. Zaveri

***Abstract:*--Motion Estimation is a process of estimating the motion of macro blocks (MB) of some predefined size from the reference frame to regenerate the current frame. This process exploits the temporal redundancy that is seen in videos due to the large number of frames present in a video sequence per second. Based on real world image sequence's characteristic of centre biased motion vector distribution, a new hybrid search (NHS) algorithm for fast block motion estimation is proposed in this paper. This paper aims at exploiting the centre-biased motion by considering the successive elimination algorithm (SEA) at centre of frame and an efficient diamond search (EDS) at the corner of the frame. This makes the algorithm considerably faster than the non-sub optimal block-matching algorithms (BMA). This paper also analyses the above mentioned algorithm in which both the time for reconstruction and quality in terms of the accuracy of motion compensation, using the PSNR ratio, MSE and average number of search points per frame are compared over various test video sequences.**

***Index Terms*—Macroblock, Motion estimation, Block matching algorithm, Successive elimination algorithm, Diamond search**

I. INTRODUCTION

MOTION estimation examines the movement of objects in an image sequence and obtains vectors representing the estimated motion. Each frame in a typical video sequence contains a lot of temporal redundancy and this can be exploited in terms of prediction of the current frame from the previous frame or a reference frame, except at scenes where the current frame is unrelated to the previous one i.e. the appearance of a new object in the current frame. Motion estimation and the exploitation of the strong correlation among the successive frames allows us to encode and transmit the motion vectors along with the error frame obtained using the regenerated current frame and the actual current frame and hence, reduces the number of bits used to convey the information. To achieve this bit reduction, various approaches and algorithms have been proposed.

The most accurate BMA is the exhaustive full-search (FS) method, which exhaustively evaluates all possible macro blocks (p x p) over a predetermined search window of size 2p+1 x 2p+1 to find the best match. The estimated motion vector is the best match achieved for a predefined block distortion measure (BDM). The only disadvantage of this method and perhaps the biggest flaw is the high computational cost associated with it. Other algorithms with reduced number of computations, for example, the successive elimination algorithm (SEA) [1], three-step search (TSS) [2], new three step search [3] , a novel four-step search (4SS) [4], efficient four step search [5], unrestricted center biased diamond search (UCBDS) [6], cross search [7], fast full search motion estimation [8], complexity bounded motion estimation [9] etc. have been proposed. Among these algorithms, the SEA is similar to the full search method except, the first one eliminates certain search points based on the Minkowiski's formula. Further reduction in number of search points has been achieved in TSS algorithm which starts with a step having nine uniformly spaced search points which get closer after every step until the step size reduces to 1. The best candidate search point in the previous step becomes the center of the current step. The main drawback of TSS is the relatively large search pattern in the first step having a distance of 4, which renders it inefficient for finding blocks with small motions. In order to exploit the characteristics of the center-biased motion vector distribution, a four-step search (4SS) algorithm has been proposed to speed up the search mechanism. It utilizes a nine-point search pattern on a 5 X 5 grid in the first step. 4SS requires only four search steps as it starts with a smaller search grid pattern. The total number of candidate search points in 4SS actually ranges from 17 (the best case) to 27 points (the worst case).

In this paper, an algorithm is proposed, namely, New Hybrid Search (NHS) which is trade off between above algorithms and the algorithms like a new fast algorithm for estimation of block motion vector [10], a dynamic search window adjustment for block matching [11], Displacement measurement and its application in interframe image coding [12], predictive coding based on efficient motion estimation [13], A new efficient block-matching algorithm for motion estimation [14] and Fast multiresolution motion estimation algorithms for wavelet-based scalable video coding [15]. It reduces the number of search points required for the motion estimation to the great extent by exploiting the best out of two search algorithms, one from sub-optimal and another from non-sub optimal category. It is computationally very efficient without any degradation in quality of the video which is very much clear from the exhaustive simulations on various video clips and results. The paper is organized as follows: Section II presents the proposed NHS algorithm in detail. The simulation and results is depicted in Section III. Our proposed algorithm, NHS, is also compared with other algorithms, namely, SEA, TSS and 4SS.

D.Jayaswal is a M-Tech. (Communication), pursuing Ph.D from SVNIT Surat, Gujrat, India(e-mail djjayaswal_vcet@yahoo.com)

Mukesh. Zaveri is the Head, Computer Engineering Department, SVNIT, Surat, Gujarat, India (e-mail mazaveri@gmail.com)

II. NEW HYBRID SEARCH ALGORITHM

NHS algorithm is devised with the sole aim of exploiting the center biased motion of successive frames, thus the SEA algorithm from the non-sub optimal category is selected to calculate the motion vectors at the center of the frame. The reason behind why SEA is preferred is that the reconstruction quality of SEA is much better than that of any sub optimal algorithms and it is also faster in terms of time required for reconstruction, among other non-sub optimal algorithms. Generally, videos are captured with very high frame rate, so the amount of motion observed at the corners of the successive frames is very small. To exploit this feature, the EDS algorithm is used with SEA in our proposed NHS algorithm. The EDS algorithm is used at the corners, as it is an extremely fast algorithm.

The SEA algorithm is used at the center of the frame and the EDS is used at the corners. The structure of SEA and EDS search pattern is depicted in the Fig 1. It explains to what extent the SEA algorithm is used and from which point the EDS algorithm begins. The following section describes EDS algorithm in detail.

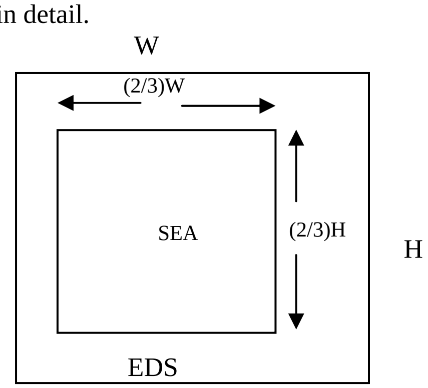
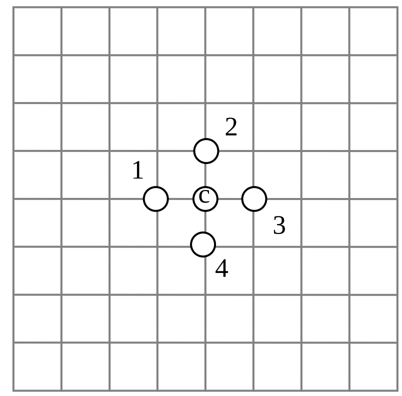


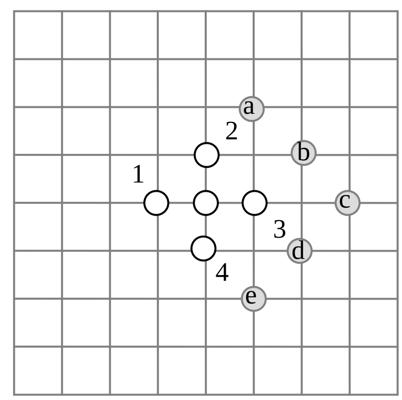
Fig. 1 Search pattern for NHS algorithm

A. Efficient Diamond Search (EDS)

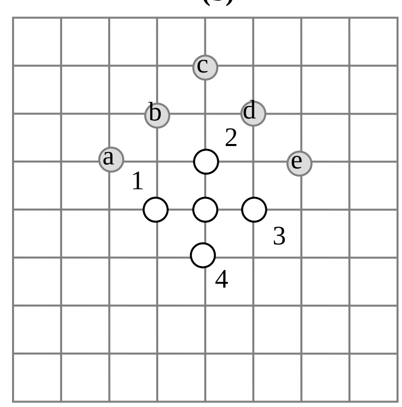
Generally, it is observed that the motion in successive frames follows a cross pattern with the motion usually taking place in the center. Block-based motion estimation algorithm divides each frame into blocks of a particular size in the form of an N X N grid. For our algorithm we consider a 15 X 15 search area, thus requiring 225 possible candidate search points per block when the FS is used. To reduce computational complexity we have to choose a suitable subset of these 225 points for a sub optimal version of the search algorithm. Fig 2(a) depicts a basic diamond search-point configuration used in the EDS algorithm. It consists of four candidate search points around the center. This pattern shows that the mandatory minimum search is to consider the 5 points.



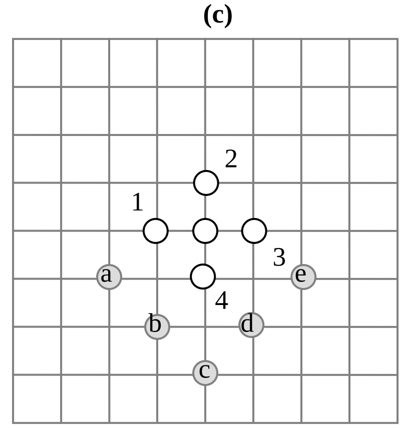
(a)



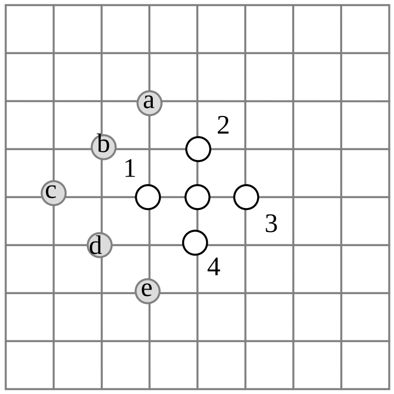
(b)



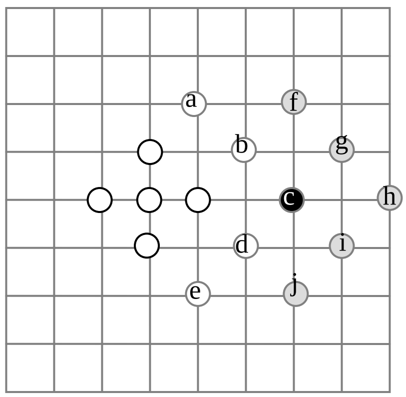
(c)



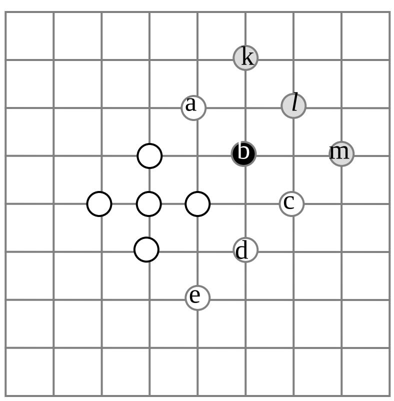
(d)



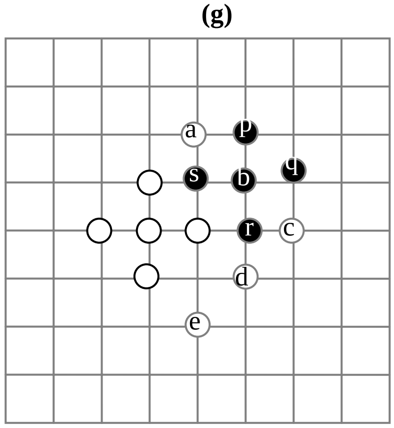
(e)



(f)



(g)



(h)

Fig 2. (a) Basic diamond search point, (b), (c), (d), (e) - Next step along a diamond's vertex and (f) Final step of NHS

The minimum BDM is calculated between the reference and the current frame. We have used the SAD which is given by,

$$\mathrm{SAD}_{(p,q)}(m_x,m_y) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} |f_t(p+i,q+j) - f_{t-\triangle t}(p+i+m_x,q+j+m_y)|$$

where $-W \leq m_x, m_y \leq +W$

Here, each block with its upper left corner at a position in the current frame is estimated / predicted from a block of the same size in the previous (reconstructed) frame by means of a motion vector. The original diamond pattern as shown in Fig.2(a) is placed at (0,0), the center of the search window. The BDM is evaluated for each of the *five* candidate search points. Depending on where the best match is found, the following cases are considered:

Case 1: CENTER
If the best match is found at the center of the small diamond, then the search stops for that particular block and the motion vectors for that point is sent across the channel. The center of diamond is labeled 'c' in the Fig 2(a).
Case 2: INITIAL CORNER
If the best match is found at any of the four corners of the initial diamond then the five points around that corner are considered. Then depending on where the best match is found, go to the cases as shown below. Four corners of the initial diamond are numbered '1', '2', '3' and '4' respectively in Fig 2 (a). and five points are shown with light gray color in the Fig. 2 (b), (c), (d) ,and (e), labeled as 'a', 'b', 'c', 'd', and 'e'.
Case 3: FURTHER CORNER
If the best match is found at the one of the corner point labeled 'c' in Fig 2(f), then the next 5 new points around that corner point are considered. The 5 points are shown with a light gray color in the Fig. 2(f), labeled as 'f', 'g', 'h', 'i', and 'j'.
Case 4: SIDE
If the best match is found at the side as shown in Fig. 2(g), labeled 'b', then the 3 new points shown in gray color in Fig. 2(g), and labeled as 'k', 'l', and 'm', are considered.
Case 5: ENDING
If the current point has the minimum BDM than a smaller diamond at a distance of one is considered keeping the current point, i.e., the minimum BDM point as the center. These points are shown in Fig. 2(f) using the black color nodes around the current point 'b' and the points that forms the small diamond are labeled with 'p', 'q', 'r', and 's'. These points are shown in Fig. 2(h).
Now comparing the BDM's of the smaller diamond points the minimum BDM is found and the motion vector of that point is sent for reconstruction. The above process is considered until the search reaches the end of the 15 X 15 search window as shown in Fig 3.

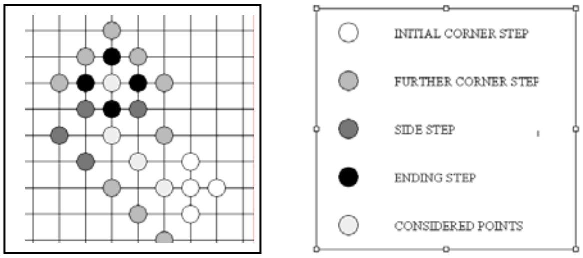


Fig 3. Efficient diamond search pattern

III. SIMULATION AND RESULTS

We have evaluated our algorithm using the large number of standard video clips. These video clips are of different nature. The simulation and results of some of clips are described here. First video clip used for experiment is that of ‘Miss America’. It contains a person presenting a report. This video only has subtle lateral movement of the person, keeping the background stationery. The NHS algorithm works very well for such sequences, as the best match at the corners will be found to be at the center itself, providing the best case search of only 5 points at the corners. Second video is the ‘Tennis Sequence’ which contains a person playing table tennis. In the clip the person serves by tossing the ball in the air off his racquet before serving. Second clip has more amount of motion compared to first one. Third one is ‘Flower Garden Sequence’. In the sequence, a camera is moving laterally, and recording a garden sequence and the amount of motion is extremely large. Moreover, a large number of new objects appear in the successive frames.

For the same set of sequences, we have experimented using different search algorithms, namely, FS, SEA, TSS, 4SS and UCBDS, and compared with the proposed NHS algorithm. The proposed algorithm is evaluated based on three parameters, namely, peak signal to noise ration (PSNR), mean square error (MSE) and average number of search points required per macro block (MB) for motion estimation. Second parameter provides how best the search points are selected through out the sequence. MSE also depicts the quality of the video reconstructed. Naturally the minimum MSE is expected for better quality of the video using the best algorithm. The third parameter provides the information about the computational complexity, i.e., few the number of search points less the computational complexity. The best algorithm is one which gives minimum MSE and very few search points. For all the clips 80 numbers of successive frames are considered for simulation.

The Table-I depicts average number of search points required for motion estimation using 8 x 8 MB size among the different search algorithms for various video clips. From Table-I it is observed that average number of search points required for motion estimation in our proposed algorithm is less in comparison with FS and SEA algorithms and more in comparison with other algorithms, TSS, 4SS and UCBDS. Table-II and Table-III describe MSE using different algorithms for various clips. MSE in Table-II and Table-III are derived using two different MB sizes 8 x 8 and 16 x 16 respectively. The MSE for our proposed algorithm is less compared to TSS, 4SS and UCBDS algorithms and more in comparison with FS and SEA algorithms.

These results are very much expected as it is mentioned earlier that the NHS algorithm is proposed to reduce the computational complexity without much degradation in the quality of a video. FS and SEA algorithms give better quality of video at the cost of computational complexity. TSS, 4SS and UCBDS algorithms reduce the computational complexity at the cost of quality of the video. It is true that the proposed algorithm is not optimal one as FS but at the same time its performance is much better than sub-optimal algorithms (TSS, 4SS and UCBDS). From these tables it is concluded that NHS algorithm reduces the computational complexity without much degradation in the quality of a video.

Another criteria used for comparison is PSNR. Fig. 4, Fig. 5 and Fig.6 depict the comparison in terms of PSNR using different search algorithms for various clips. From these figures it is clear that the proposed NHS algorithm performs as good as FS and SEA and much better than TSS.

TABLE-I
AVERAGE NO. OF SEARCH POINTS PER MOTION VECTOR FOR THE FIRST 80 FRAMES

Algorithm	**Miss America**	**Flower Garden**	**Tennis**
FS	274	274	274
SEA	105.31	152.53	175.62
NHS	**57.83**	**79.54**	**102.71**
TSS	26	26	26
4SS	27.22	26.06	27.34
UCBDS	24.74	26.72	23.97

TABLE-II
AVERAGE MSE OF THE FIRST 80 FRAME
(8 X 8 BLOCK)

Algorithm	**Miss America**	**Flower Garden**	**Tennis**
FS	18.37	2330.99	1331.30
SEA	18.37	2330.99	1331.30
NHS	**19.68**	**2686.49**	**1440.86**
TSS	24.90	2685.68	1439.86
4SS	28.08	2840.31	1489.99
UCBDS	28.30	2887.68	1498.06

TABLE –III
AVERAGE MSE OF THE FIRST 80 FRAME
(16 X 16 BLOCK)

Algorithm	Miss America	Flower Garden	Tennis
FS	18.99	2419.21	1286.91
SEA	18.99	2419.21	1286.91
NHS	**19.71**	**2978.73**	**1508.99**
TSS	35.25	3372.11	1618.63
4SS	36.43	3477.62	1660.13
UCBDS	36.05	3502.61	1660.63

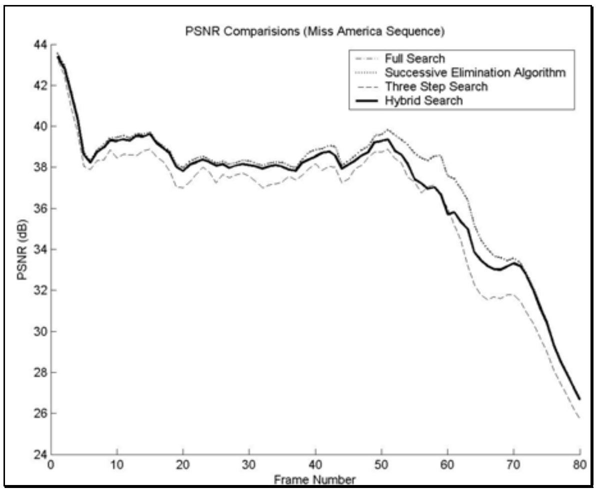


Fig 4. PSNR comparison of NHS with FS, SEA & TSS algorithms for the MISS AMERICA sequence.

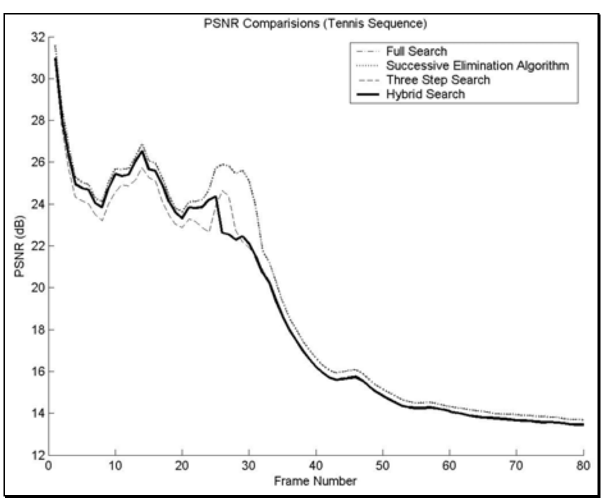


Fig 5. PSNR comparison of NHS with FS, SEA & TSS algorithms for the Tennis sequence.

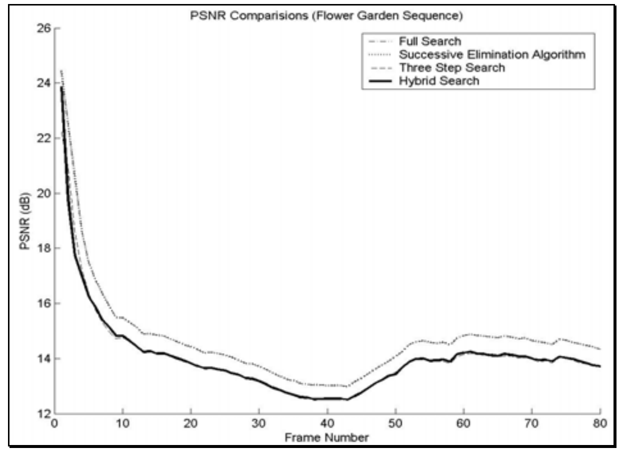


Fig 6. PSNR comparison of NHS with FS, SEA & TSS algorithms for the Flower Garden sequence.

IV. CONCLUSION

In this paper, we have proposed the New Hybrid Search algorithm (NHS), which exploits the best out of two algorithms, one sub optimal algorithm, namely, the Efficient Diamond Search (EDS) and another non-sub optimal algorithm, namely, the Successive Elimination Algorithm (SEA). The proposed algorithm takes the advantage of the central biased motion present in most videos. From simulation results it is concluded that the proposed algorithm NHS performance is very much comparable with that of FS and SEA and much better than TSS, 4SS and UCBDS. The NHS algorithm reduces the computational complexity without any degradation in the quality of a video.

V. REFERENCES

[1] W. Li and E. Salari, “Successive Elimination Algorithm for Motion Estimation”, *IEEE Transactions on Image Processing*, Vol. 4, No. 1, January 1995.

[2] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, “Motion compensated interframe coding for video conferencing,” in *Proc. National Telecommunications Conf.*, Nov. 29–Dec. 3, 1981, pp. G.5.3.1–G.5.3.5.

[3] R. X. Li, B. Zeng, and M. Liou, “A new three step search algorithm for block motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, pp. 438–442, Aug. 1994.

[4] Lai-Man Po and Wing-Chung Ma, “A Novel Four-Step Search Algorithm for Fast Block Motion Estimation”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 3, June 1996.

[5] Kuan-Tsang wang , “ Efficient four-step search’’, *IEEE international symposium on Circuits and Systems*, 1998, ISCAS apos,’98 Vol. 4, 31-May – 3 Jun 1998.

[6] Jo Yew Tham, Surendra Ranganath, Maitreya Ranganath and Ashraf Ali Kassim, “A Novel Unrestricted Centre Biased Diamond Search Algorithm for Motion Estimation”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 8, No. 4, August 1998.

[7] M. Ghanbari, “The cross search algorithm for motion estimation,” *IEEE Trans.commun.,*vol.38 , no.7, pp.950-953, July 1990

[8] Jong-Nam Kim, Sung-Cheal Byun, Yong-Hoon Kim, and Byung-Ha Ahn, “Fast full search motion estimation algorithm using early detection of impossible candidate vectors”, *IEEE Transaction on Signal Processing*, vol. 50, no. 9, September 2002.

[9] Antonio Chimienti, Claudia Ferraris, and Danilo Pau, "A complexity-bounded motion estimation algorithm", *IEEE Transactions on image processing*, vol. 11, no. 4, April 2002.

[10] B.Liu and A.zaccarin, " New fast algorithm for estimation of block motion vector," *IEEE Trans. Circuits Syst. Video Technol.*, vol.3, pp 148-157, Apr 1993

[11] L.W.Lee, J.F. Wang, J.Y.Lee, and J.D. Shie, " Dynamic search window adjustment and interlaced search for block- matching algorithm," *IEEE Trans. Circuits Syst. Video Technol.*, vol.3, pp. 85-87, Feb. 1999

[12] J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, pp. 1799–1808, Dec. 1981.

[13] R. Srinivasan and K. R. Rao, "Predictive coding based on efficient motion estimation," *IEEE Trans. Commun.*, vol. 33, pp. 1011–1015, Sept. 1995.

[14] Hanan, suneer, mohsen and magdy bayoumi, "A new efficient block-matching algorithm for motion estimation" *Journal of VLSI Signal processing system*, vol. 42, issue1, Jan. 2006.

[15] Yu Liu and King Ngi Ngan, "Fast multiresolution motion estimation algorithms for wavelet-based scalable video coding," *Elsevier Signal processing Image communication* vol.22, March 2007

VII BIOGRAPHIES

Mukesh A. Zaveri received the B.E. degree in electronics engineering from Sardar Vallabhbhai Regional College of Engineering and Technology, Surat, India, in 1990, the M.E. degree in electrical engineering from Maharaja Sayajirao University, Baroda, India, in 1993, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology Bombay, Mumbai, in 2005. He is currently an Assistant Professor and the Head of the Computer Engineering Department, Sardar Vallabhbhai National Institute of Technology. His current research interests include the area of signal and image processing, multimedia, computer networks, sensor networks, and wireless communications

Deepak J.Jayaswal was born on November 22, 1969. He received the B.E degree in electronics from Walchand College of engineering, Shivaji University, Kolhapur, Maharashtra, India in 1991, M-Tech degree in communication from IIT Bombay, 2002.

He is working in St. Francis Institute of Technology, Borivali(w) Mumbai from the year 2005 as Assistant Professor in Electronics and Telecommunication department, His research area includes Image and Video processing.

.