Implementation of motion estimation algorithm based on Intracoding and Refresh with Video epitomic Priors

K. Yoga Ramya Krishna[1], T.Sravanthi[2]
[1]student M.Tech (CS,) Department of ECE, VBIT, Andhra Pradesh, India
[2]Ms. T.Sravanthi, Assistant Professor, Department of ECE, VBIT, Telangana State, India

***Abstract*— in this paper, we are interested in video epitomic analysis. In video coding, intra coding plays an important role in both coding efficiency and error resilience. There are many motion estimation algorithms for video compression such as full search, diamond search, cross search. But the complexity of these algorithms is high. The solution to this problem is video epitomic analysis. The extension to this video epitomic analysis is compression oriented video epitomic analysis. In compression oriented analysis we use bilateral filter to get the image quality.**

Index Terms— Intra coding, Epitomic analysis, Image epitome, Refresh.

I. INTRODUCTION

In current existing coding standards, three basic picture (slice) coding types are designed: I frame (intra prediction coded frame a.k.a. intra frame), P frame (forward prediction coded frame), and B frame (bi-directional prediction coded frame). I frame has the feature of independent encoding and decoding, while the following P frames or B frames directly or indirectly take I frame as reference frame. In some special applications such as video editing, video streaming, and high-definition movie compression, I frame appears in a very high frequency, affording the main roles of compression tasks. Considering all the above, we used H.264/AVC in this paper.

The epitome of a video sequence is a spatially and/or temporally compact representation of the video that retains the video's essential textural, shape, and motion components. With just a few frames to work with, the video epitome models multiple motion patterns simultaneously within its frames. While the video epitome can itself be useful for visual purposes, its true power arises when used within a larger model for applications such as motion analysis, super-resolution, video imprinting, and compression.

II. TECHNICAL APPROACH

Video compression is applied on a series of consecutive images in a video stream, making use of similarities between neighboring images. Video compression uses a similar method as that of still image compression. However, it adds compression between the frames to further reduce the average file size.

The latest video compression standard, H.264 (also known as MPEG-4 Part 10/AVC for Advanced Video Coding), is used in this paper. 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. There are three types of frames used in H.264 compression they are I frames (intra frames), P frames (predicted frames), and B frames (predicted bi-directional frames).

1. I frames are compressed without reference to any other frames.
2. P frames are encoded with reference to a previous frame, called forward prediction.
3. B frames are encoded with reference to both the previous frame and the next frame. This is called forward and backward prediction.

A. Full Search Algorithm

Generally, in this method all possible modes are checked with performing the motion estimation for every blocks and calculating the sum of absolute difference for all block. First of all frames are divided into 16X16 macro blocks. Each macro block in reference frame is compared to all macro blocks in other frames if match founds it takes distance difference along x-direction and as well as in y-direction. Then sum of these two differences is called as sum of absolute difference (SAD). This process is continued for all macro blocks in reference frame. From all these sum of absolute difference we find the minimum absolute difference (MAD) and it is treated as motion vector

B.Diamond Search Algorithm

The *DS* algorithm employs two search patterns. The first pattern consists of nine checking points from which eight points surround the centre, forming a diamond shape. This pattern is referred to as large diamond search pattern (LDSP). The second patterns referred to as small diamond search pattern (SDSP). The SDSP consists of five checking points.

A).Large Diamond Search Pattern (LDSP)

1. Large diamond search pattern (LDSP) is centered at the search center, and all the checking points of LDSP are tested. If the center position gives the minimum SAD, go to Step 3 otherwise go to Step 2.

2. The center of LDSP moves to the point where the minimum SAD was obtained in the previous step, and all

ISSN: 2278 – 7798

the points on LDSP is tested. If the center position gives the minimum SAD, go to Step 3 otherwise recursively repeat this step.

3. Switch the search pattern from LDSP to SDSP. The point that yields the minimum SAD is the final solution of the motion vector.

b) Small Diamond Search Pattern (SDSP)

1. Small diamond search pattern (SDSP) is centered at the search center, and all the checking points of SDSP are tested. If the center position yields the minimum SAD (i.e., no motion), then the center represents the motion vector otherwise go to Step 2.

2. The center of SDSP moves to the point where the minimum SAD was obtained in the previous step, and all the points on SDSP is tested. If the center position yields the minimum SAD, then the center represents the motion vector otherwise, recursively repeat this step.

C.Adaptive search

The cross search algorithm follows the ideas similar to that of logarithmic search, but some differences lead to fewer computational search points.

Steps to find motion vector are

- Assume ICP is having minimum distortion creates a cross search pattern and takes the centre of cross search pattern is the motion vector.
- I f it is not having minimum distortion then find the distortion at other points of cross search pattern and if at (s,0) minimum distortion occurs compare to ICP then search goes in that direction.
- If the distortion at (8s,0) is increased than compared to the previous point then search stops at that point and come back to the previous point and create another cross search pattern.

This process is continued until the search centre matched to the minimum distortion

D. Video Epitomic Analysis

Image epitome preserves enough texture and shape characteristics in image. The size of image epitome is only A fraction of that of original image or even less. The image epitomes are derived from the MLE with Expectation-Maximization algorithm. The steps in EM-algorithm are,

1. The original image is divided into a set of image patches with different sizes.
2. First define an energy function,

$$\varnothing(\mathrm{x},\mathrm{z}) = \frac{1}{\left|X^{+}\right|}\sum_{p\in X^{+}}\left|x_p - z_{p^*}\right|^2$$

In the above equation, x represents the original image, z for image epitome. X^+ denotes image patch set sampled from original image, and the variable p represents the coordinate set of image patch in X^+, and p^*stands for the coordinate set of best match in image epitome. X_p is the pixel value vector in the coordinates of p while z_{p*} is the vector composed of pixel values in the coordinates of p^*

3. Then find best image epitome by using,

$$Z_{best} = \arg\min\varnothing(\mathrm{x},\mathrm{z})$$

4 After random initialization of z by the EM algorithm,

E-step: Searching for the best matching block z_{p*} for every x_p in X^+.

M-step: Updating z to minimize the energy function $\Phi(x,z)$ according to z_{p*} and x_p using,

$$z(\mathrm{i},\mathrm{j}) = \left(\sum_{p\in X^{+},(\mathrm{i},\mathrm{j})\in \mathrm{p}^{*}}\delta_{p*}(\mathrm{i},\mathrm{j})\right)^{-1}\left(\sum_{p\in X^{+},(\mathrm{i},\mathrm{j})\in \mathrm{P}^{*}}\overline{x_p}(\mathrm{i},\mathrm{j})\right)$$

Where p and $p*$ are the matching coordinate set pair through E-step, and $x_p(i, j)$ represents the pixel values of the corresponding pixel in x_p with the location of (i, j) in p^*. The function $\delta_p(I,j)$ is the delta function which is given as,

$$\delta_p(\mathrm{i},\mathrm{j}) = \begin{array}{l} 1 \text{ If } (i,j) \in P^* \\ 0 \text{ if } (\mathrm{i},\mathrm{j}) \in \mathrm{P}^* \end{array}$$

E.Bilateral filter

The visual quality of image epitome is as important as the characteristics for coding performance for this application. At the mean time, the limited space (size) resource requires similar texture to merge as much as possible, and this will bring the loss of significant contours and edges in image and visual quality would be degraded greatly. To solve such problem, the detail information can be extracted and used to boost patch search to get more accurate results which preserve much more detail information. To avoid the quality degradation, detail information can be taken as detail channel, which is added to YUV channel to highlight detail content in original image. Detail channel is extracted through bilateral filtering.

III. RESULT ANALYSIS

A. Outputs of motion estimation algorithms

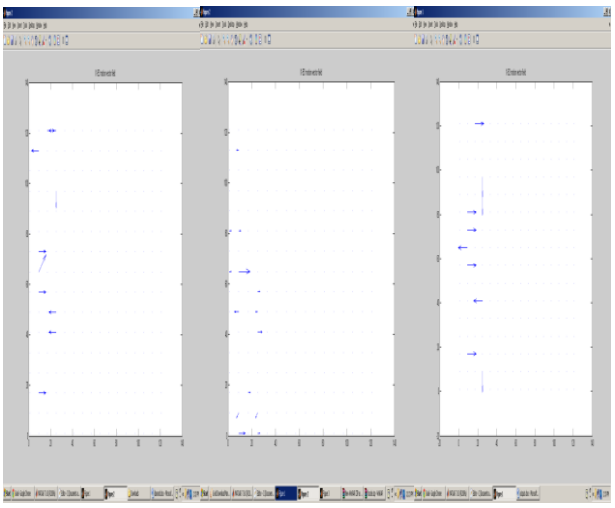


Figure 1a) Full Search **b)** Diamond Search **c)** Adaptive Search motion vectors

ISSN: 2278 – 7798

B. comparision of video compression techniques without video epitomic analysis for macro block size 16X16

Type	Elapsed time (sec)	TCP	Quality(dB)
Full search	24.008249	408608	87.7888
Diamond search	21.324490	113.4609	90.3273
Adaptive search	11.684163	61.9336	93.4028

C. Outputs of Video Epitomic Analysis

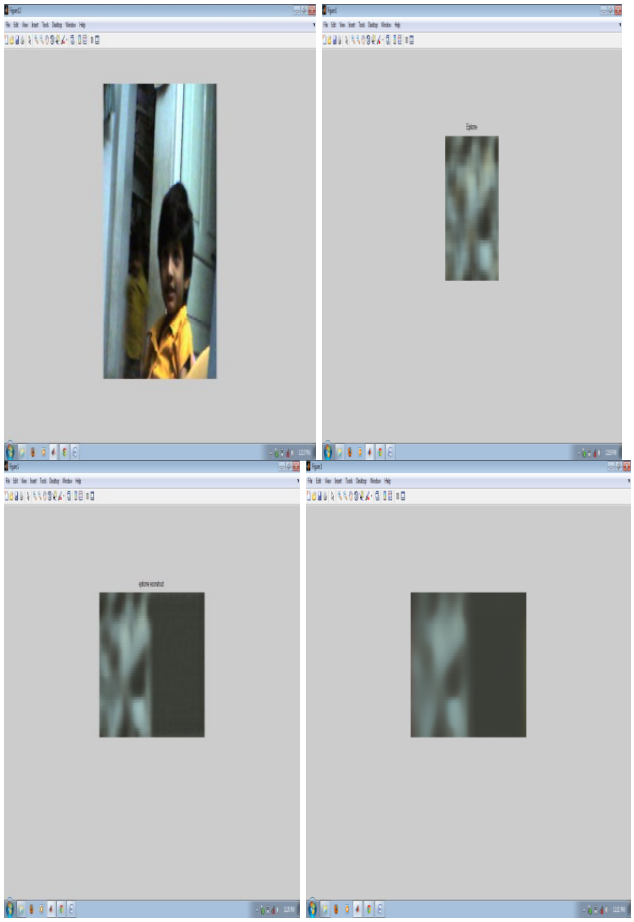
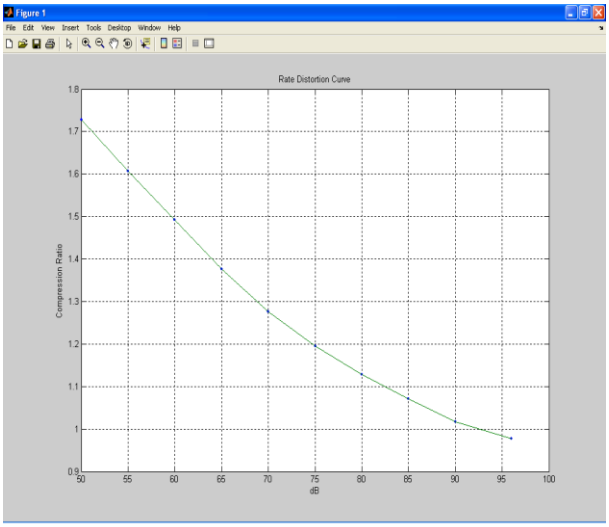


Fig2: a) Original image **b)** Epitome **c)** reconstructed epitome **d)** bilateral filter output

- Compression Ratio=32:1
- PSNR obtained from Video Epitomic Analysis=58.19dB

D: Rate distortion curve for reconstructed epitome



IV Conclusion

In this research different motion estimation algorithms have been analyzed. In compression oriented video epitomic analysis EM-algorithm is used to obtain the epitome. The results shows that by using video epitomic analysis complexity is decreased compared to other algorithms and also quality is increased for adaptive search algorithm.

V REFERENCES

[1] Advanced video coding for generic audiovisual services, ITU-T Rec. H.264 and ISOIIEC 14496-10 (MPEG4-AVC) Standards, May 2003

[2] Heiko Schwarz, Detlev Marpe, Thomas Wiegand, Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, IEEE Trans. Circuits Syst. Video Techn, 17(9), 2007: 1103–1120.

[3] Thiow Keng Tan, Choong Seng Boon, Yoshinori Suzuki, Intra prediction by template matching, in Proc. IEEE International Conference on Image Processing (ICIP), pp.1693–1696, 2006.

[4] S.-L. Yu and C. Chrysafis, New intra prediction using intra macroblock motion compensation, JVT-C151, 3rd meeting of Joint Video Team (JVT), May 2002.

[5] Tomasi.C. et al. “Bilateral filtering for gray and color images”, in Proceedings of IEEE Conference on Computer Vision (ICCV), pp.839-846, 1998.

[6] N. Jojic, B. J. Frey, A. Kannan, Epitomic analysis of appearance and shape, in Proc. IEEE Int. Conf. Comput. Vis.

ISSN: 2278 – 7798

(ICCV), pp: 34–41, 2003.image & video coding, in Proc. IEEE International Conference on Image Processing (ICIP), pp.2881–2884, 2007.

VI Author's profile

First Author

K. Yoga Ramya Krishna pursuing M.Tech at Vignana Bharathi Institute of Technology (VBIT). Her areas of interest are image registration, and video compression.

Second Author

Ms. T. Sravanthi working at Vignana Bharathi Institute of Technology (VBIT). She qualified ME from Chaitanya Bharathi Institute of Technology (Affiliated to Osmania University) & B.Tech from Bhoj Reddy Engineering College for Women (Affiliated to Jawaharlal Nehru Technology University, Hyderabad).Her areas of interest are image registration, tracking, data fusion and video compression.

ISSN: 2278 – 7798