COPYRIGHT © 2011 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 01, ISSUE 02, MANUSCRIPT CODE: 110104

High Frequency Component Matching Motion Estimation: A Revolutionary Improvement for Real-Time Telemedicine Applications

Morsalin Uz Zoha

Abstract—In Telemedicine applications, for diagnostic purposes, it is essential that the video compression process causes no tangible loss of detail and introduces no noticeable artifacts which could be otherwise misinterpreted as being pathological in nature. On the other hand, due to the limitation of storage and transmission and the huge amount of medical video, high compression ratio is often required. Video coding exploits the high correlation between successive image frames to improve coding efficiency, which is achieved by motion estimation and motion compensation techniques. This paper demonstrates the feasibility of incorporating the high frequency component matching block motion estimation in H.264 in order to improve video compression performance in telemedicine applications. Implementation using the H.264 codec and thereby facilitating performance analysis, the experimental result indicates that high frequency component matching introduces a revolutionary improvement as a matching criterion for motion compensated predictive coding in real-time telemedicine applications.

Index Terms—High frequency component matching, H.264, Motion Estimation, Telemedicine.

1 Introduction

IN telemedicine applications, the principal requirement is to achieve high processing speed and a low computing time simultaneously without scarifying in medical video quality, which requires advanced video compression technologies. H.264/AVC is the latest video coding standard [1], [4] developed by ISO/IEC MPEG and ITU-T VCEG, which aims at providing functionality similar to existing video coding standards such as H.263+ and MPEG-4 but with significantly better performance and improved support for reliable transmission. The encoding process is the most important part in the whole video transmission procedure. The strategy for motion estimation of objects in video sequences plays a very important role in achieving outstanding coding performance. Motion estimation is the most intensively computational and the most time-consuming part. So it would significantly affect the efficiency of whole video coding systems and the quality of reconstructed video sequence.

H.264 adopts the block-based motion estimation. A very popular technique for block-based motion estimation is the block matching algorithm (BMA). The BMA compares square block of image data from subsequent frames of video to determine the correlation between the luminance values of the data. If data in one block is closely correlated to the data in a different location on different frames, it is assumed that movement of recorded objects has occurred between the frames. Alternatively the similarities between the two blocks can be used to compress the video images. If the similarities are known between frames, only the data information between each frame needs to be saved.

The performance of BMA involves two aspects: search algorithms and block matching methods. For BMA the simplest and most accurate strategy is to use the full search algorithm (FS), sometimes referred to as the exhaustive search or the brute force search. The full search gives optimal solution, in terms of prediction quality, by exhaustively searching over all possible blocks within the search window. However, the computational complexity of full search has motivated a host of sub optimal but faster search strategies.

For the block matching methods which is actually to find a matched block to the current block or region that minimizes the energy in the motion compensated residual, there are many measurements of the energy, such as SAD (Sum of Absolute Difference), SSD (Sum of Square Difference), MSE (Mean Square Error), SATD (Sum of Absolute Transform Difference). FS estimation is guaranteed to find the minimum SAD in the search window. The FS-SAD motion estimation approach for a block size of $N \times N$ pixels can be expressed in the form of

$$\underset{k,l}{SAD}(m,n) = \sum_{i=kN}^{(k+1)N-1} \sum_{j=lN}^{(l+1)N-1} \left| I_t(i,j) - I'_{t-1}(i+m, j+n) \right|$$
$$-s \le m,n \le s-1 \qquad (1)$$

where $\underset{k,l}{SAD}(m,n)$ shows the SAD value computed for a displacement of (m,n) pixels for the block with index (k,l);

- ***Morsalin Uz Zoha*** *is with the University of Chittagong (www.cu.ac.bd), as a Researcher, Chittagong-4331, Bangladesh. E-mail: mzoha@acm.org*

$I_t(i,j)$ and $I'_{t-1}(i,j)$ are luminance values of the current and previously reconstructed frames, respectively, and s defines the search window range. The motion vector (MV) for any block is obtained as the horizontal and vertical displacements resulting in the lowest total absolute difference value.

Most approaches used for motion estimation do not take properties of the transform encoder into account. In [2], it is noted that when conducting motion estimation using MAD, the block that can minimize the distortion is selected without considering how many bits are required to code the residue, while the following transform coding tries to code the residue using minimum number of bits. It is proposed in [2] to take this point into consideration using a novel block matching criterion that aims at eliminating the mismatch between motion estimation and the following transform coding. The smooth constrained matching criterion considers the difference between the maximum residue and the minimum residue within the residue block, denoted as MMD, and incorporates this smoothness measure into the MAD. If formulated using the SAD criterion, the smooth constrained SAD measure of a macroblock can be expressed as

$$SAD_{sc} = SAD + \alpha \sum_{m=1}^{4} MMD_m \quad (2)$$

where α is a weighting factor, and MMD_m shows the MMD measure of the m^{th} 8×8 block of the corresponding 16×16 macroblock. It is shown in [2] that although the computational load of motion estimation increases, the overall the video quality is improved at constant bit-rate.

In [3], high frequency component matching criterion has been proposed, which also takes the transform coding properties into account, and aims to improve overall video quality but at the cost of lower computational load. For this purpose, basic features of the transform coding process are considered first. The discrete cosine transform (DCT) is typically used to encode the prediction error in motion compensated predictive coding (MCPC) systems, and typically a nonuniform quantization matrix, which uses a quantization table obtained from psychovisual experiments to exploit the limits of the human visual system, is utilized to quantize the DCT coefficients before entropy coding. The nonuniform quantization matrix has lower step sizes in the low-frequency range and higher step sizes in the high-frequency range. The standard SAD measure given in (1) does not take into consideration the fact that during transform and quantization, low-frequency detail is retained more effectively compared to high-frequency detail. It is shown in [3], MCPC video coding performance can be improved by taking the features of the transform and quantization properties into account at the motion estimation /compensation stage. Instead of the standard SAD matching, high-frequency components are matched, and compensation for low-frequency differences is left to transform coding.

This paper mainly concentrates on improving the medical video quality using the high-frequency component matching block motion estimation in H.264/AVC codec. The rest of the paper is organized as follows. The next section briefly reviews the H.264 standard and the high frequency component matching motion estimation. Section 3 presents experimental results on medical videos comparing high frequency component matching motion estimation with others utilizing the standard AVC codec. This paper concludes with Section 4.

2 A Brief Review of H.264 and High Frequency Component Matching Motion Estimation

2.1 Review of H.264

The recent international video coding standard, H.264/AVC [1], [4] has been approved by ITU-T as Recommendation H.264 and by ISO/IEC as International Standard 14 496-10 (MPEG-4 part 10) Advanced Video Coding (AVC). It is not a fundamentally different method, but rather a significant refinement of well-established methods. The elements common to all video coding standards are present in the current H.264/AVC recommendation. Some new techniques, such as spatial prediction in Intra coding, adaptive block size motion compensation, 4×4 integer transformation, multiple reference pictures (up to seven reference pictures) and content adaptive binary arithmetic coding (CABAC), are used in this standard. The testing results [4], [5], [6] have shown that H.264 has achieved substantial superiority of video quality over that of H.263, MPEG-2, and MPEG-4. It has achieved up to 50% in bit rate saving compared to H.263 or MPEG-4 coding schemes. This means that H.264 offers significantly higher coding quality with the same bit rates [7]. Therefore, H.264 is a serious contender for a variety of telemedicine applications.

2.2 High Frequency Component Matching (HFCM)

Instead of directly matching the pixel values to obtain motion vectors, Ertürk proposed to match the high-frequency components only [3]. To obtain the high-frequency components of an image frame, the Laplacian of Gaussian (LoG) filter is utilized.

The LoG filter is defined as:

$$LoG(i,j) = -\frac{1}{\Pi\sigma^4}\left[1 - \frac{i^2 + j^2}{2\sigma^2}\right] e^{-\frac{i^2+j^2}{2\sigma^2}} \quad (3)$$

where σ shows the Gaussian standard deviation. It is possible to approximate the LoG filter using a discrete convolution kernel. The LoG filter is extremely successful in extracting high-frequency components. The high-frequency components of an image are obtained by filtering the image with the LoG filter using convolution in the form of

$$I_{HFC}(i,j) = LoG(i,j) * I(i,j) \quad (4)$$

where $I_{HFC}(i, j)$ shows the high-frequency components of the image, and * represents convolution.

The new high-frequency component matching criterion is composed as the sum of absolute differences of high-frequency components and can be expressed in the form of

$$SAD_{HFC}(m,n) = \sum_{i=kN}^{(k+1)N-1} \sum_{j=lN}^{(l+1)N-1} \left| I_{HFC,t}(i,j) - I'_{HFC,t-1}(i+m, j+n) \right|$$
$$-s \le m,n \le s-1 \qquad (5)$$

The high-frequency component matching criterion ensures that high-frequency components of the block are matched as well as possible and leaves the compensation for low-frequency components to the prediction error coding stage, which consists of transform, quantization, and entropy coding. Any low-frequency mismatch at the motion estimation/compensation stage will not be observed in the reconstructed image because the typical DCT transform and nonuniform quantization will successfully encode low-frequency components.

3 Experimental Results

The evaluation of experimental result includes the comparison among SAD_{HFC}, SAD_{SC} and SAD to apply the better one in medical video compression i.e. telemedicine applications. Here, H.264/AVC codec is utilized to evaluate the performance of high frequency component matching motion estimation and four different sized i.e. CIF (352 x 288), QCIF (176 x 144) and larger sizes (640 x 480 and 512 x 512) 4:2:0 YUV format medical video sequences are tested. H.264 is implemented with H.264 reference model version JM6.1e [8]. Table 1 lists the test video sequences used in the experiment. In the experiment the frame rate of H.264 is 30 fps, slice mode is OFF, rate control is OFF, RDO is OFF, hardmard is OFF, search range is 32, search is not restricted, symbol mode is CABAC, no partition mode and out file mode is Annex B.

In the first test, the interframe quantization parameter (QP) is set to a constant value and the average PSNR is observed. Table 2 shows the average PSNR versus QP results for the sequences. It is seen that SAD_{HFC} and SAD_{SC} -based motion estimation always outperform SAD-based motion estimation in terms of PSNR, while SAD_{HFC} and SAD_{SC} perform similarly. The gain of SAD_{HFC} over SAD increases with coarser quantization (i.e., high QP), while the gain is lower for smaller QP values, as in this case, high-frequency detail can be transform coded with more success.

TABLE 1
TEST SEQUENCE

Sequence	Format	Size	Frame Number
X-ray computed tomography	YUV	512x512	100
Echocardiography	YUV	640x480	100
Coronary angiogram	YUV	352x288	100
Mammography	YUV	176x144	100

In the second test, the bit rate is set to a constant value, and the average PSNR is observed. In this case, the encoder changes the interframe QP to achieve the target bit-rate. The average PSNR versus bit-rate results are provided in Table 3. Note that bit-rates below200 kbps are not taken into consideration as the encoder is observed to randomly drop frames to achieve the target bit-rate. It is seen that, the gain what SAD_{HFC} provides compared to SAD, reduces for higher bit-rates as high-frequency details are transform encoded more successfully at higher bit-rates. At lower bit-rates, sometimes the SAD_{HFC} performance can fall slightly below SAD performance. The main reason for the high gain that is obtained in the case of constant QP, not being reflected onto bit-rate, is observed to be the entropy encoding procedure that is typically optimized to encode SAD prediction errors. SAD_{SC} is observed to perform slightly better than SAD_{HFC} on average. However, the computational load of SAD_{HFC} is lower compared to SAD_{SC}, because in SAD_{HFC}, the LoG is computed once for the current frame and the previously decoded frame, whereas SAD_{SC} needs to compute the MMD measure for each search location. Therefore, for telemedicine applications where real time medical video transmission with acceptable quality is required, High Frequency Component Matching motion estimation proves itself as a revolutionary improving factor.

These results are in accordance with what has been described in [3]. In [3] Ertürk left the tasks of implementation using the standard AVC/H.264 codec and thereby facilitating performance analysis of high frequency component matching motion estimation for larger frame sizes as future work; here in this paper these tasks are experimented incorporating high frequency component matching motion estimation in H.264 for telemedicine applications where the test medical sequences' frame sizes include 512 x 512, 640 x 480, CIF and QCIF formats.

TABLE 2
AVERAGE PSNR VERSUS QP FOR SAD, SAD_{SC} AND SAD_{HFC} MOTION ESTIMATION

QP		5	10	15	20	25	30
		Average PSNR					
X-ray computed tomography	SAD_{HFC}	60.21	56.54	52.11	47.85	42.59	37.99
	SAD_{SC}	60.22	56.55	52.12	47.84	42.62	38.00
	SAD	59.59	55.06	50.25	44.87	39.86	36.15
Echocardiography	SAD_{HFC}	55.59	51.32	46.88	41.53	36.77	32.53
	SAD_{SC}	55.59	51.30	46.78	41.60	36.78	32.53
	SAD	54.02	50.23	44.92	39.73	34.88	29.92
coronary angiogram	SAD_{HFC}	58.91	53.98	48.34	45.67	42.14	36.84
	SAD_{SC}	58.95	54.08	48.35	45.65	42.14	36.83
	SAD	58.21	53.48	46.97	44.85	40.33	33.98
Mammography	SAD_{HFC}	61.23	58.12	53.75	48.39	44.25	40.22
	SAD_{SC}	61.24	57.99	53.76	48.38	44.26	40.25
	SAD	60.96	57.34	52.09	46.12	42.08	39.01

TABLE 3
AVERAGE PSNR VERSUS BIT-RATE FOR SAD, SAD_{SC} AND SAD_{HFC} MOTION ESTIMATION

BIT-RATE		200	300	400	500	600	700	800	900	1000
		Average PSNR								
X-ray computed tomography	SAD_{HFC}	38.71	40.88	43.22	46.91	48.19	50.54	53.06	56.89	58.12
	SAD_{SC}	38.72	40.89	43.23	46.91	48.20	50.55	53.07	56.89	58.13
	SAD	38.70	39.56	42.01	45.65	47.70	50.11	52.87	56.33	58.12
Echocardiography	SAD_{HFC}	35.52	38.76	40.21	42.85	45.69	48.47	51.88	54.12	56.55
	SAD_{SC}	35.52	38.76	40.21	42.86	45.70	48.48	51.88	54.11	56.55
	SAD	35.53	38.60	39.98	41.12	44.45	47.89	51.44	54.01	56.54
coronary angiogram	SAD_{HFC}	33.88	34.57	36.55	39.54	42.94	46.48	49.33	52.24	55.02
	SAD_{SC}	33.89	34.58	36.56	39.45	42.94	46.48	49.34	52.24	55.01
	SAD	33.88	33.98	35.89	38.25	41.28	45.95	49.01	51.99	55.01
Mammography	SAD_{HFC}	35.99	38.51	40.11	42.78	44.09	47.49	50.22	53.44	56.22
	SAD_{SC}	36.00	38.52	40.11	42.79	44.10	47.50	50.23	53.44	56.23
	SAD	36.00	37.14	39.79	41.20	43.21	47.01	49.97	53.12	56.21

4 CONCLUSION

Motion estimation and compensation is an important part of video compression. Video compression efficiency mainly depends on the efficient detection of motion vectors, so extra care should be taken to choose best motion estimation technique for a particular application. Unlike in some other video compression applications where video quality could be compromised with compression ratio, it is a very challenging task to compress medical videos as it requires both high quality and high compression ratio. In this paper it is shown that incorporating high frequency component matching motion estimation in H.264 introduces a radical improvement for medical video compression to be used in real-time telemedicine applications.

ACKNOWLEDGMENT

The author wishes to thank Dr. Md. Shafiqul Kabir Khan from Dhaka Medical College Hospital, Bangladesh and Dr. Hasina Akter from Rajshahi Medical College Hospital, Bangladesh for their participation in order to check whether the reconstructed / decompressed frames of the

test sequences used in the experiment are medically valid for diagnosis after using H.264/AVC codec incorporating SAD, SAD_{SC} and SAD_{HFC} motion estimation for compression.

REFERENCES

[1] Information Technology-Coding of Audio-Video Objects-Part 10: Advanced Video Coding, ISO/IEC FDIS 14496-10, 2003

[2] X. Jing, C. Zhu and L. P. Chau, "Smooth constrained motion estimation for video coding", *Signal Process.*, vol. 83, no. 3, pp. 677–680, Mar. 2003.

[3] Sarp Ertürk, "A New Perspective to Block Motion Estimation for Video Compression: High-Frequency Component Matching", *IEEE Signal Processing Letters, Vol. 14, No. 2, February 2007*,pp 113-116.

[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[5] A. Joch, F. Kossentini, and P. Nasiopoulos: A performance analysis of the ITU-T draft H.26L video coding standard. *Proc. 12th Int. Packetvideo Workshop*, Pittsburgh, PA, Apr. 2002.

[6] S. Saponara, C. Blanch, K. Denolf, and J. Bormans: The JVT advanced video coding standard: complexity and performance analysis on a tool-by-tool basis. *In Proc. 13th Int. Packetvideo Workshop*, Nantes, France, Apr. 2003.

[7] T. Hallbach, "H.264 video compression standard", *in Proceedings of 6th Nordic Signal Processing Symposium*, Bergen, Norway, Oct. 2003.

[8] Reference Model JM6.1e, JVT. [Online]. Available: ftp://ftp.imtcfiles.org

Morsalin Uz Zoha has been serving as a Researcher at the University of Chittagong, Bangladesh. He is now engaged in two ministry projects and some other NGO projects. He is a member of IEEE Bangladesh Section. He plays an active role to organize seminars and technical sessions in IEB (Institution of Engineers, Bangladehs) and BCS (Bangladesh Computer Society). He has several international publications. His research interest includes Telemedicine, Bio-informatics and Computer Vision.