Frame Rate Up-Conversion Using Adaptive Bilateral Motion Estimation

Tsung-Han Tsai , Hong-Guang Chen and Hsueh-Yi Lin
Department of Electrical Engineering
National Central University
No.300, Jhongda Rd., Jhongli City, Taoyuan County
Taiwan
han@ee.ncu.edu.tw

Abstract: - Frame rate up-conversion is a post processing tool to convert the frame rate from a lower number to a higher one. This paper presents a motion compensated frame rate up-conversion algorithm that uses the expanded range of the adaptive bilateral motion estimation to enhance the accuracy of motion vectors. In experiments using benchmark test sequences, the proposed algorithm improves the average PSNR of interpolated frames by up to 1.91 dB compared to the conventional algorithm1.

Key-Words: Frame rate up conversion, Motion estimation, motion compensated interpolation.

1 Introduction

Frame rate up conversion (FRUC) is a technique to convert lower frame rate video to higher frame rate. It is applied in low bandwidth applications, where half of the frames are dropped in the encoding stage. The dropped frames are later interpolated during the decoding process. Recently, FRUC is frequently applied to reduce motion blur of liquid crystal display (LCD). Hold-type displays, such as LCD, display input image by rotating internal crystal to correct angle. The mechanism introduces motion blur, which degrades image quality. To improve the visual quality, a common approach is to convert input video from traditional 60 frames per second (fps) to 120 fps. It is proven effective in maintaining output image sharpness, as specified in [1].

FRUC is divided broadly into two fields. The first field is to combine video frames without taking motion activity into account, such as frame repetition, black data insertion, and frame averaging. These algorithms are perfectly applied in the sequences without lots of motion. However, moving objects in a video are jerked and blurred, especially in high resolution video.

The second category uses advanced conversion technique that employs motion. The interpolator that generates the missing frames can be either linear (Haar filter) or nonlinear (median operation [2]). The linear approach, referred to as motion compensated frame rate up-conversion (MC-FRUC), reconstructs better quality of intermediate frames than the nonlinear one by taking into account the motion of individual objects.

MC-FRUC consists of motion estimation (ME) and motion compensated interpolation (MCI), as shows in Fig.1. ME is the process of determining motion vectors (MV) of the neighboring frames, and the MVs are employed in interpolating a new frame by MCI. Among existing ME algorithms, the block matching algorithm (BMA) is the most popular one because it can be easily designed with low computational complexity. However, BMA suffers from block artifacts such as compensation area overlap and the gap that can not be compensated. Recently, bilateral ME [3] is proposed as a way to avoid these BMA artifacts.

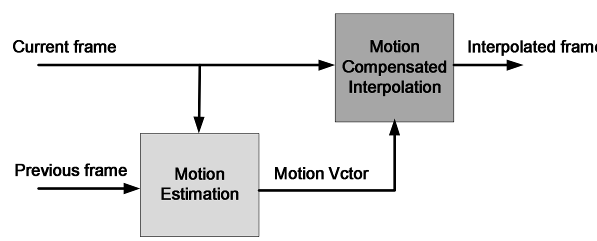

Fig. 1. Block diagram of MC-FRUC.

In this paper, Adaptive bilateral ME for frame rate up-conversion is proposed. Based on the observation of previous methods, a parameter is added for motion estimation. When two simlar regions are observed, the estimated MV might be wrong. With the additional parameter, the false area will be recognized and removed. The results also show that the better estimation accuracy is the key in improving the image quality of interpolated frames.

This paper is organized as follows. In Section 2, we describe the proposed MC-FRUC algorithm in detail. Section 3 presents experimental results based on

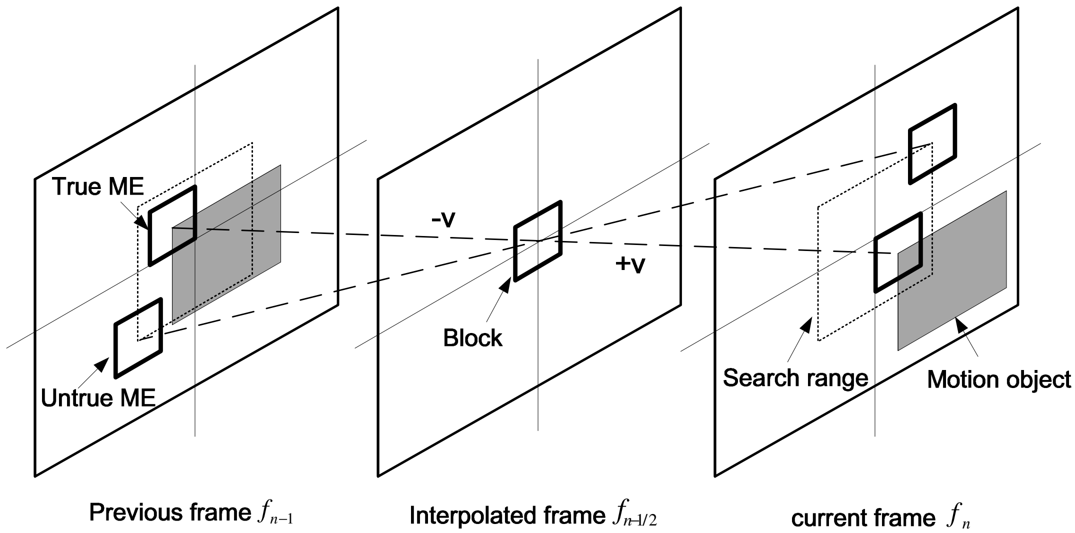

Fig. 2. Bilateral ME.

objective evaluations. Finally, Section 4 concludes the paper.

2 The Proposed Motion Compensated Frame Rate Up-Conversion Algorithm

The conventional bilateral ME calculates a motion vector using the temporal symmetry between the previous and the current frames, from the perspective of the interpolated frame. Sum of bilateral absolute difference (SBAD) is used in bilateral ME. Within the search range of motion estimation, the absolute difference of pixels (in a block) between current frame and previous frame is calculated, and corresponding MV is selected that minimizes SBAD, as shown in (1).

$$SBAD(dx,dy)=\sum_{x\in S_x}\sum_{y\in S_y}\begin{bmatrix}|f_{n-1}(x-dx,y-dy)- \\ f_n(x+dx,y+dy)|\end{bmatrix} \quad (1)$$
$$v=\arg\min_{(dx,dy)\in S_{x,y}}\{SBAD(dx,dy)\}$$

where (dx,dy) stands for the motion vector candidate. f_{n-1} and f_n denote the previous and current frames, respectively. S_x and S_y denote the horizontal search range and vertical search range, respectively. v denotes the selected MV for MCI. The bilateral ME computes the MV using the positions of the candidate blocks with the minimum SBAD. Since the interpolated frame is used as a reference for the motion vector computation, the artifacts of BMA can be avoided.

However, the bilateral ME estimates MVs using only the temporal symmetry from the perspective of the interpolated frame. If similar or identical objects within a frame are matched in the temporal symmetry from the perspective of the interpolated frame, the bilateral ME can provide a untrue MV. Take Fig. 2 as an example, the gray rectangle moves in the diagonal direction in front of the uniform background. From the interpolated frame's aspect, the block tries to find the best match within the search range. Two regions, including true ME block and untrue ME block, are with similar SBAD. Thus, it is probable that the untrue ME block is found and compensated. It causes the image quality degradation.

To solve this problem, we propose the adaptive bilateral ME in a way that assumes the true motion vectors are possible near the search center (0, 0). Adaptive bilateral ME applies Adaptive sum of bilateral absolute difference (ASBAD) for motion estimation. The definition of ASBAD is shown in (2).

$$ASBAD(dx,dy)=\sum_{x\in S_x}\sum_{y\in S_y}\begin{bmatrix}(|f_{n-1}(x-dx,y-dy)- \\ f_n(x+dx,y+dy)|+ \\ A(dx,dy))\end{bmatrix} \quad (2)$$
$$v=\arg\min_{(dx,dy)\in S_{x,y}}\{ASBAD(dx,dy)\}$$
$$A(dx,dy)=(dx^2+dy^2)/8$$

Where $A(dx,dy)$ is a parameter to reduce the probability of estimating untrue MV. The value of $A(dx,dy)$ is proportion to the distance of estimated MV.

In [9], the weighted correlation index (WCI) is applied for motion estimation, as shown in (3).

$$WCI=SBAD(1+K(dx^2+dy^2)) \quad (3)$$

where K is a parameter that can be thought of as an

ISSN: 1790-5117

ISBN: 978-960-474-075-8

elastic constant. The simulations suggest a value in the range 0.2-0.05.

In WCI, the original SBAD equation is multiplied by a factor. When estimated region is outward from estimation center, the term $(1+K(dx^2+dy^2))$ is larger than inward area. Since $(1+K(dx^2+dy^2))$ is multiplied with SBAD, outward area exhibits larger WCI compared with inward area. The difference might jeopardize estimated result. In the proposed ASBAD, the adaptive parameter *A* is added to SBAD. Thus, the inward-outward region difference on ASBAD lies solely in the parameter *A*. The mechanism ensures correct motion estimation. Besides, the additional parameter *A* is very helpful in removing untrue MV

Fig. 3 shows the snapshots of the sequence *Mobile*. Sicne the clendar is composed of a white paper and lots of words, the paper can be regarded as uniform region. Fig. 3(a) is interpolated based on SBAD. From Fig. 3(a), the characters with interpolated numbers are skewed. The character ‘1’ in the upper-left corner is totally un-distinguishable. The result, based on ASBAD in Fig. 3(b), is similar to ordinary *Mobile* sequence. All characters with number are correctly reconstructed.

Fig. 4 shows the snapshots of the sequence *Flower garden*. The sky is a uniform region. Fig. 4(a) is the result from SBAD. It is obvious that the lamp is broken into pieces. Fig. 4(b) shows the result interpolated by ASBAD. Compared with Fig. 4(a), the lamp is interpolated perfectly.

3 Experimental Results

In the experiments, we set the block size with 16 x 16 pixels and the search range with 8x8 pixels. We used *Flower garden*, *Mobile*, *Foreman*, and *Football* as the test sequences in the CIF (352 x 288) format. For each test sequence, 101 frames are used. First, we removed the 50 odd frames and constructed 50 new odd frames from the 51 even frames. Then, we computed the PSNRs of the constructed odd frames

(a)

(b)

Fig. 3. Interpolated frame of *Mobile* as calculated by (a) the conventional SBAD, (b) the proposed ASBAD.

(a)

(b)

Fig. 4. Interpolated frame of *Flower garden* as calculated by (a) the conventional SBAD, (b) the proposed ASBAD.

with respect to the original sequences with 101 frames, and compared the PSNR, defined in (4).

$$MES = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) \quad (4)$$

where MSE stands for the mean square error, N is the number of pixels in the image, and x_i and y_i are the gray levels of the i-th pixels in the reference and the processed images. PSNR is one of the most popular metrics used to evaluate the objective image quality in the field of the image processing.

In the experiment to evaluate the performance of the proposed MC-FRUC algorithm, we used the modified weighted correlation index (MWCI) method [4] to replace WCI. We set the quantization parameter of the MWCI to 2, as in [5]. Table 1 summarizes the average PSNRs for four test image sequences. Table 1 also indicates that the proposed algorithm outperformed MWCI by up to 1.91 dB on average for all test sequences.

Table 1 Average PSNRs of Test Sequences

Average PSNRs(dB)				
	Mobile	Foreman	Football	Flower garden
MWCI	20.75	28.43	19.10	24.48
Proposed ASBAD	22.72	31.35	20.87	25.42

4 Conclusion

In this paper, we proposed a new motion estimation method used for frame rate up conversion. The proposed method enhanced the image quality by additional adaptive parameter. When two similar regions are measured with similar sum of absolute difference value, the additional parameter removes the untrue region. The experimental results show that it improves the image quality of the interpolated frame compared with existing algorithms. In the objective evaluation, the average PSNR of the proposed MC-FRUC was 1.91 dB higher than MWCI.

From the experimental results, we conclude that the proposed algorithm, ASBAD, is indeed a novel solution in frame rate up conversion.

References:

[1] T. Kurita, “Moving picture quality improvement for hold-type AM-LCDs,” SID Int. Symposium Digest of Technical Papers, pp.986–989, June 2001.

[2] O. A. Ojo and G. de Haan, ”Robust motion compensated up conversion”, IEEE Fans. on Consumer Electronics, vo1.43, no.4, pp. 1045-1056, 1997.

[3] B. D. Choi, J. W. Han, C. S. Kim, and S. J. Ko, “Motion compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation,” IEEE Trans. Circuits and Systems for Video Technology, vol. 17, no. 4, pp. 407-416, Apr. 2007.

[4] T. Ha, S. Lee, and J. Kim, "Motion compensated frame interpolation by new block-based motion estimation algorithm," IEEE Trans. Consumer Electronics, vol. 50, no. 2, pp. 752–759, May 2004.

[5] B. D. Choi, J. W. Han, C. S. Kim, and S. J. Ko, "Frame rate up conversion using perspective transform," IEEE Trans. Consumer Electronics, vol. 52, no. 3, pp. 975-982, Aug. 2006.

[6] Suk-Ju Kang, Kyoung-Rok Cho,”Motion Compensated Frame Rate Up-Conversion Using Extended Bilateral Motion Estimation” IEEE Transactions on Consumer Electronics, Vol. 53, No. 4, NOVEMBER 2007.

[7] Ya-Ting Yang, Yi-Shin Tung, and Ja-Ling Wu, Senior “Quality Enhancement of Frame Rate Up-Converted Video by Adaptive Frame Skip and Reliable Motion Extraction” IEEE transactions on circuits and systems for video technology, vol. 17, no. 12, december 2007

[8] B. D. Choi, J. W. Han, C. S. Kim, and S. J. Ko, “Motion compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation,” IEEE Trans. Circuits and Systems for Video Technology, vol. 17, no. 4, pp. 407-416, Apr. 2007.

[9] Castagno, R., Haavisto, P., Ramponi, G.: ‘A method for motion adaptive frame rate up-conversion’, in IEEE Trans. on Circuits and Systems for Video Tech., Vol. 6, Oct. 1996, pp. 436 – 446