Motion-Compensated Frame Interpolation Using a Parabolic Motion Model and Adaptive Motion Vector Selection

Kang-Sun Choi and Min-Chul Hwang

We propose a motion-compensated frame interpolation method in which an accurate backward/forward motion vector pair (MVP) is estimated based on a parabolic motion model. A reliability measure for an MVP is also proposed to select the most reliable MVP for each interpolated block. The possibility of deformation of bidirectional corresponding blocks is estimated from the selected MVP. Then, each interpolated block is produced by combining corresponding blocks with the weights based on the possibility of deformation. Experimental results show that the proposed method improves PSNR performance by up to 2.8 dB as compared to conventional methods and achieves higher visual quality without annoying blockiness artifacts.

Keywords: Motion-compensated interpolation, frame rate up-conversion, parabolic motion model.

I. Introduction

Recently, motion-compensated frame interpolation (MCFI) has been studied to improve temporal resolution with reduced motion jerkiness and jitter by increasing the frame rate of the video. In conventional MCFI methods, the motion vector (MV) field between successive frames is estimated, and then corresponding motion-compensated (MC) blocks are located at the midpoint of the MV in the interpolated frame under the equivelocity motion assumption. Since this approach causes the hole and overlapping problem, methods such as dominant block selection [1] and bilateral motion estimation (ME) [2], which determine the MV of each non-overlapping block in the interpolated frame, have been employed.

However, if an object in a scene has accelerated or decelerated during the interval between successive frames, the MVs of blocks inside the object can be determined inaccurately. In this case, the conventional MCFI methods produce ghost artifacts by averaging the shifted versions of the object [3]. Moreover, if a moving object in the scene deforms during the frame interval, averaging two versions of the deformed object also produces the ghost artifacts, regardless of the correctness of the object motion.

In this letter, we propose an MCFI method using a parabolic motion (PM) model which determines an asymmetric backward/forward MV pair (MVP) to estimate the true motion trajectory more accurately. To avoid the hole and overlapping problem, we introduce a simple but effective MVP reliability measure by which the most reliable MVP is selected among the estimated neighboring MVPs. By weighing the bidirectional MC blocks in accordance with the possibility of deformation, natural interpolated frames can be produced.

The rest of this letter is organized as follows. Section II presents the proposed algorithm in detail. Section III provides experimental results, and section IV, the conclusion.

II. Proposed MCFI Method

In the proposed method, we exploit a constant acceleration motion model to estimate the true motion trajectory of moving objects. Figure 1 shows the diagram of the proposed MCFI

Manuscript received June 11, 2010; revised Aug. 23, 2010; accepted Sept. 7, 2010.
Kang-Sun Choi (phone: +82 41 560 1307, email: ks.choi@kut.ac.kr) was with the Department of Electrical Engineering, Korea University, Seoul, Rep. of Korea, and now is with the School of Electrical, Electronics & Communication Engineering, Korea University of Technology and Education, Cheonan, Rep. of Korea.
Min-Chul Hwang (email: mc09.hwang@samsung.com) is with the Advanced R&D Group, Visual Display Division, Samsung Electronics, Suwon, Rep. of Korea.
doi:10.4218/etrij.11.0210.0203

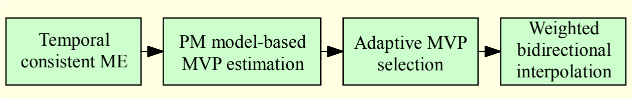

Fig. 1. Diagram of proposed MCFI method.

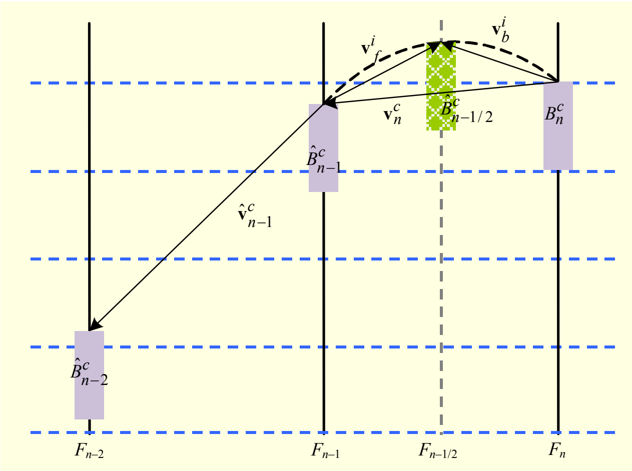

Fig. 2. Calculation of MVs based on PM model.

method. When the MV for each block is initially obtained, ME utilizing temporal consistency is proceeded to improve the reliability of the MV. We note that the estimated MV represents the motion of the block during a frame interval. In order to determine an accurate bidirectional MVP at the middle of the interval, PM-based MVP estimation is performed. Figure 2 illustrates how the proposed MVP estimation method obtains a more accurate MVP for the interpolated frame using the PM model, where the horizontal dotted lines represent the block grid on the frames. Let B_n^c and $\hat{B}_{n-1}^c$ denote the c-th block of the current frame F_n and its corresponding MC block in F_{n-1} obtained by the MV, $\mathbf{v}_n^c$, respectively. Given $\mathbf{v}_n^c$ and $\hat{\mathbf{v}}_{n-1}^c$, an MVP $(\mathbf{v}_f^i, \mathbf{v}_b^i)$ can be uniquely determined using the PM model which will be explained later. Since $\hat{B}_{n-1/2}^c$, the motion-compensated interpolated (MCI) block, is not aligned with the block grid in most cases, as shown in Fig. 2, the hole and overlapping problems can take place in the interpolated frame. To avoid these problems, we re-estimate an MVP $(\mathbf{v}_f^c, \mathbf{v}_b^c)$ for each grid-aligned block using the proposed MVP reliability measure. Then, each block of the interpolated frame $F_{n-1/2}$ is interpolated using its corresponding MC blocks for the MVP.

1. Motion Estimation Considering Temporal Consistency

Typically, objects in a scene move consistently in a short period of time. In the proposed method, when the MV for each block is initially obtained, the reliability of the MV is examined using the temporal consistency of the motion trajectory that is given by

$$TD(\mathbf{v}_n^c) = \left|\hat{\mathbf{v}}_{n-1}^c - \mathbf{v}_n^c\right|. \quad (1)$$

If $TD(\mathbf{v}_n^c)$ is larger than a threshold τ_n which indicates the limit of acceleration/deceleration, $\mathbf{v}_n^c$ is regarded as an *outlier* and replaced by another MV associated with the next minimum sum of absolute difference (SAD). This outlier elimination is repeated until $TD(\mathbf{v}_n^c)$ is less than the threshold. In our work, τ_n for the frame F_n is determined by

$$\tau_n = \max\left(\left\{\left|\mathbf{v}_{n-1}^c\right| \mid c \in C\right\}\right), \quad (2)$$

where C is a set of all the blocks in a frame.

Since $\hat{\mathbf{v}}_{n-1}^c$ has not been usually determined at time (n–1) when $\hat{B}_{n-1}^c$ has not been defined, $\hat{\mathbf{v}}_{n-1}^c$ is approximated with available MVs at n–1, specifically the MVs of each B_{n-1}^k which overlap $\hat{B}_{n-1}^c$, as follows:

$$\hat{\mathbf{v}}_{n-1}^c = \sum_{k=1}^{N} w^k \cdot \mathbf{v}_{n-1}^k, \quad w^k = P^k \Big/ \sum_{i=1}^{N} P^i, \quad (3)$$

where P^k is the number of overlapping pixels of B_{n-1}^k, and N is the number of the overlapping blocks.

2. Bidirectional Motion Vector Pair Estimation for the Interpolated Frame

A. Parabolic Motion Model-Based Motion Vector Estimation

The quadratic equation modeling the PM trajectory is given by

$$(x, y) = (a_x t^2 + b_x t + c_x, a_y t^2 + b_y t + c_y), \quad (4)$$

where t is the time index and (x, y) denotes the position of each block in F_t. Since we have three positions on the motion trajectory of B_n^c at t=n, n–1, and n–2, and since each position contains two coordinates, the model parameters, a_x, b_x, c_x, a_y, b_y, and c_y can be determined with $\mathbf{v}_n^c$ and $\hat{\mathbf{v}}_{n-1}^c$. Without loss of generality, assume that n=0 and B_n^c is located at the origin (0, 0). Then, c_x and c_y are zero. Using t=n–1=–1 and the position of the MC block $\hat{B}_{n-1}^c$, which is $\mathbf{v}_n^c$, the quadratic equation is expressed as

$$\mathbf{v}_n^c = (a_x - b_x, a_y - b_y)\ . \quad (5)$$

Similarly, assigning the position of $\hat{B}_{n-2}^c$, $\mathbf{v}_n^c + \hat{\mathbf{v}}_{n-1}^c$, to the quadratic equation with t=n–2=–2 leads to

$$\mathbf{v}_n^c + \hat{\mathbf{v}}_{n-1}^c = (4a_x - 2b_x, 4a_y - 2b_y). \quad (6)$$

By using (5) and (6), the bidirectional MVP of the block $\hat{B}_{n-1/2}^c$ at time $t = n - 1/2 = -1/2$ can be calculated by

$$\mathbf{v}_b = \left(a_x/4 - b_x/2, a_y/4 - b_y/2\right) = 0.625\mathbf{v}_n^c - 0.125\hat{\mathbf{v}}_{n-1}^c, \quad (7)$$

$$\mathbf{v}_f = \mathbf{v}_b - \mathbf{v}_n^c = -0.375\mathbf{v}_n^c - 0.125\hat{\mathbf{v}}_{n-1}^c. \quad (8)$$

B. Adaptive Motion Vector Pair Selection

To determine an appropriate MVP $(\mathbf{v}_f^c, \mathbf{v}_b^c)$ for each grid-aligned block $B_{n-1/2}^c$, we propose a simple but effective MVP reliability measure as follows:

$$R^i = P^i + \lambda \cdot \mathrm{SAD}(-\mathbf{v}_f^i, -\mathbf{v}_b^i), \quad (9)$$

where $(\mathbf{v}_f^i, \mathbf{v}_b^i)$ is an MVP of $\hat{B}_{n-1/2}^i$, which overlaps the current block $B_{n-1/2}^c$ with the overlap amount P^i, and a negative value λ is a weighting parameter obtained empirically. $\mathrm{SAD}(-\mathbf{v}_f^i, -\mathbf{v}_b^i)$ represents SAD between two MC blocks of $B_{n-1/2}^c$ for $-\mathbf{v}_f^i$ and $-\mathbf{v}_b^i$.

In order to improve the reliability of the MVP selection, geometric displacement is considered based on the observation that the more $\hat{B}_{n-1/2}^i$ overlaps $B_{n-1/2}^c$, the more $(\mathbf{v}_f^i, \mathbf{v}_b^i)$ is close to true MVP of $B_{n-1/2}^c$. For example, let us consider a scene where a moving object becomes slightly deformed with the static background. Assuming that a block in the interpolated frame is located near the object boundary, the candidate MVPs for the block probably contain both the MVPs for the object and the background. If the MVP representing the object motion passes closest to the center of the block, it can be more adequate for the MVP of the block. However, if geometric displacement is not considered, the MVP representing the background motion is always chosen since the SAD for the static background is much lower than that for the deforming object.

In the proposed MVP reliability measure, the overlap amount P^i implies the geometric displacement. The final MVP for $B_{n-1/2}^c$ is determined by selecting an MVP producing maximal reliability as follows:

$$(\mathbf{v}_f^c, \mathbf{v}_b^c) = \arg\max_{(\mathbf{v}_f^i, \mathbf{v}_b^i)} \left(R^i\right). \quad (10)$$

3. Weighted Bidirectional Interpolation

To improve the video quality of the interpolated frame, the possibility of deformation of bidirectional MC blocks is estimated from the MVP by exploiting the relative length ratio of the MVP. If an MV within an MVP is much shorter than the other MV, its MC block tends to move rigidly. Specifically, the objects in the MC block are translated to the current block without deformation. Thus, an MC block for the shorter MV is regarded to have little possibility of deformation and is weighted more heavily to produce an MCI block.

Let v_x and v_y be the row and column components of $\mathbf{v}_k$, respectively. The length of $\mathbf{v}_k$ is defined by

$$L[\mathbf{v}_k] = \sqrt{v_x^2 + v_y^2 + d_t^2}, \quad (11)$$

where d_t is a constant value representing the time interval for $\mathbf{v}_k$. Then, the interpolated frame is obtained by

$$F_{n-1/2}[\mathbf{g}] = \alpha \cdot F_{n-1}[\mathbf{g} - \mathbf{v}_f^c] + (1-\alpha) \cdot F_n[\mathbf{g} - \mathbf{v}_b^c], \quad (12)$$

where $\alpha = L[\mathbf{v}_b^c] \big/ (L[\mathbf{v}_f^c] + L[\mathbf{v}_b^c])$.

III. Experimental Results

To evaluate the performance of the proposed method, we compared the proposed method to conventional methods, including Choi's [1] and Gan's [4] methods. Since Gan's method was designed for the video decoder where motion information can be extracted from the bitstream, we modified the method to perform ME for the comparison. Various CIF (352×288) test sequences listed in Table 1 were utilized. The constant d_t in (11) was set to 1/2, and the block of size 8×8 was used in all the different methods. In order to determine λ in (8) empirically, λ^*, which produces the highest average PSNR for the test sequences, is searched in the range [–0.5, 0] with step size 0.05. In this work, λ was determined as –0.15 through the parameter search.

We skipped even frames and performed interpolation with every pair of odd frames. The objective evaluation is compared between the skipped and the interpolated frame in terms of PSNR. Let $F_n[i, j]$ and $\hat{F}_n[i, j]$ denote intensity value at (i, j)th pixel in the skipped frame and that in the interpolated frame, respectively. Then, PSNR is defined as

$$\mathrm{PSNR} = 10\log_{10} \frac{255^2}{\frac{1}{W \times H}\sum_{j=0}^{H-1}\sum_{i=0}^{W-1}\left(F_n[i,j] - \hat{F}_n[i,j]\right)^2}, \quad (13)$$

where $W \times H$ is the spatial resolution of the sequence.

Table 1. Comparison of average PSNR of interpolated frames.

Sequence	Choi's [1] (dB)	Gan's [4] (dB)	Proposed (dB)
Foreman	30.48	30.32	32.37
Silent	33.70	34.88	35.44
Football	20.74	21.03	21.32
Soccer	24.85	26.13	26.46
Mobile	21.41	24.65	24.92
News	33.77	35.29	35.60
Akiyo	40.81	43.25	43.67
Hall	32.52	34.14	34.46

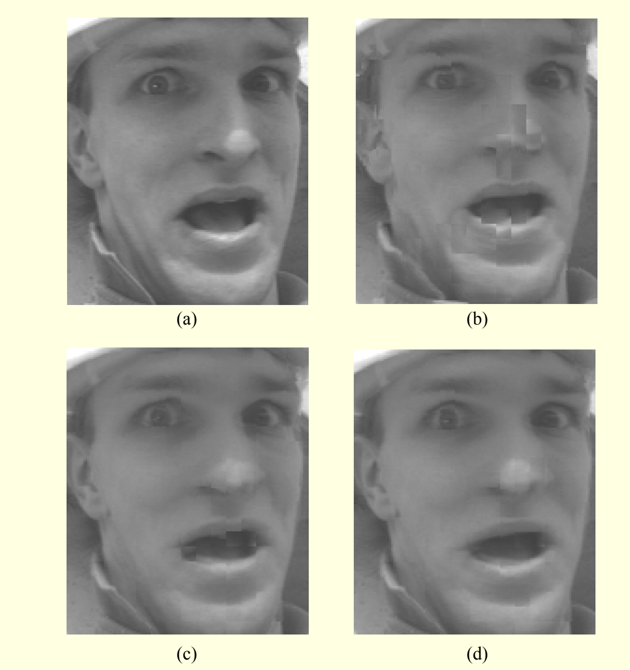

Fig. 3. Comparison of objective quality of interpolated frames for Foreman sequence: (a) part of an original frame, (b) Choi's method, (c) Gan's method, and (d) proposed method.

Table 2. Comparison of average processing time (s/frame).

	Choi's [1]	Gan's [4]	Proposed
Motion estimation	0.294	0.127	0.129
Motion compensation	0.028	0.009	0.009
Total	0.322	0.136	0.138

Table 1 shows the comparison of the average PSNR of the interpolated frames for various test sequences. The proposed method achieves higher PSNR performance than the conventional methods. In the Foreman sequence, Gan's method results in the lowest PSNR even though it also utilizes the acceleration motion model, whereas the proposed method overcomes the weak point of Gan's method by using the adaptive MVP selection. Consequently, the PSNR performance is increased up to about 2 dB.

Figure 3 compares the subjective quality of the interpolated frames for the Foreman sequence. Choi's method fails to obtain the accurate MVs so that the severe blocking artifacts are caused as shown in Fig. 3(b). Gan's method produces better quality frames than the Choi's. However, blockiness artifacts due to MV inconsistency are generated in the interpolated image as shown in Fig. 3(c). The proposed method achieves higher visual quality without blocking artifacts as opposed to the conventional methods.

Table 2 compares the average processing times for the MCFI methods. Choi's method takes a much longer time than the other methods due to segmentation and overlapped block motion compensation. The additional computation of the proposed method is ignorable as compared with Gan's method.

IV. Conclusion

In this letter, we proposed the improved motion-compensated frame interpolation technique for frame rate up-conversion. The PM model is exploited to estimate accurate MVPs. To solve the hole and overlapping problems, the MVP for the block on the grid of the interpolated frame is adaptively selected among MVPs obtained by the PM model based on the proposed MVP reliability measure. The adaptive weighted bidirectional interpolation considering block rigidity enhances the final image quality of the interpolated frames. Experimental results show that the proposed method outperforms the conventional methods both objectively and subjectively.

References

[1] B.-D. Choi et al., "Frame Rate Up-Conversion Using Perspective Transform," *IEEE Trans. Consum. Electron.*, vol. 52, no. 3, Aug. 2006, pp. 975-982.

[2] B.-D. Choi et al., "Motion-Compensated Frame Interpolation Using Bilateral Motion Estimation and Adaptive Overlapped Block Motion Compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 4, Apr. 2007, pp. 407-416.

[3] S.-J. Kang et al., "Multiframe-Based Bilateral Motion Estimation with Emphasis on Stationary Caption Processing for Frame Rate Up-Conversion," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, Nov. 2008, pp. 1830-1838.

[4] Z. Gan, L. Qi, and X. Zhu, "Motion Compensated Frame Interpolation Based on H.264 Decoder," *Electron. Lett.*, vol. 43, no. 2, Jan. 2007, pp. 96-98.