LETTER

Joint Motion-Compensated Interpolation Using Eight-Neighbor Block Motion Vectors

Ran LI†(a), Student Member, Zong-Liang GAN†, Zi-Guan CUI†, and Xiu-Chang ZHU†, Nonmembers

SUMMARY Novel joint motion-compensated interpolation using eight-neighbor block motion vectors (8J-MCI) is presented. The proposed method uses bi-directional motion estimation (BME) to obtain the motion vector field of the interpolated frame and adopts motion vectors of the interpolated block and its 8-neighbor blocks to jointly predict the target block. Since the smoothness of the motion vector field makes the motion vectors of 8-neighbor blocks quite close to the true motion vector of the interpolated block, the proposed algorithm has the better fault-tolerance than traditional ones. Experiments show that the proposed algorithm outperforms the motion-aligned auto-regressive algorithm (MAAR, one of the state-of-the-art frame rate up-conversion (FRUC) schemes) in terms of the average PSNR for the test image sequence and offers better subjective visual quality.

key words: frame rate up-conversion (FRUC), 8-neighbor blocks, minimum mean square error (MMSE), Tikhonov regularization

1. Introduction

Frame rate up-conversion (FRUC) refers to the technique of constructing a high frame rate video by periodically inserting new frames into an input lower frame rate video. Traditional FRUC algorithms based on the motion compensation consist of three steps, motion estimation (ME), motion vector analysis (MA) and motion compensated interpolation (MCI). ME algorithms mostly use full search and so suffer some mistakes in the motion vector field whether the advanced ME algorithms are performed or not. At a result, after ME, there will always be some mistakes in the motion vector field whether the advanced ME algorithms are performed or not. At a result, after ME, we usually perform MA process to alleviate the outliers found in the motion vector field [3]–[5]. Traditional MCI method introduces blocking artifacts since block edges may not always be consistent with heterogeneous objects. By extending traditional MCI, overlapped block motion compensation (OBMC) [6] and adaptive OBMC (AOBMC) [7], can provide improved prediction accuracy at the expense of increased computational complexity. Recently, Zhang et al. [8] blazed a new trail for FRUC by introducing auto-regressive (AR) model in which each pixel in the interpolated frame is approximated by a linear combination of the pixels in a square neighborhood in the reference frames; the proposed motion-aligned auto-regressive (MAAR) algorithm is one of the state-of-the-art FRUC schemes.

In this Letter, we present a novel joint motion-compensated interpolation (8J-MCI) algorithm. Based on two assumptions that is, temporal symmetry between blocks in previous and following frames and smoothness of motion vector field, the proposed method uses motion vectors of the interpolated block and its 8-neighbor blocks to jointly make a prediction. By combining 8J-MCI with BME, we improve both the subjective and objective quality of the interpolated frame.

2. Proposed Algorithm

The BME algorithm [5] is used to estimate the motion vector field of the interpolated frame, which consists of the two steps that the forward motion estimated using block matching criterion – the sum of absolute difference (SAD) and the bi-directional motion vector refinement using bi-directional SAD (SBAD). After BME, each block in the interpolated frame has a single motion vector, but there will always be some mistakes in the motion vector field filed so as to produce blocking artifacts by performing traditional MCI. Depending on the smoothness of motion vector field, the interpolated block can be predicted by considering the motion vectors of its 8-neighbor blocks in which it is quite possible that there are several motion vectors approaching to the true motion vector of the interpolated block, as shown in Fig. 1.

Our BME-based 8J-MCI method is described as fol-

---

Fig. 1 Joint motion-compensated interpolation using motion vectors of the interpolated block and its 8-neighbor blocks.
lows. Firstly, we find the candidate matching blocks $\varphi_i$ \((i = 1, 2, \ldots, 9)\) of the interpolated block from the previous frame $f_{r-1}$, and the prediction $x_{ip}$ can be calculated by the linear weighted sum of these candidate blocks, that is,

$$x_{ip} = \sum_{i=1}^{9} \alpha_i \varphi_i + n_1 = \Phi \alpha + n_1.$$

(1)

Here, the noise component $n_1$ from $n_1 = [n_{11}, n_{12}, \ldots, n_{1L}]^T$ has been generated independently and obeys a Gaussian distribution with mean zero and variance $\sigma_n^2$, that is, $\Pr(n_{1k}) = N(0, \sigma_n^2)$, where $L$ is the length of $x_{ip}$. Likewise, another prediction $x_{if}$ can also be calculated using the candidate blocks $\psi_i$ \((i = 1, 2, \ldots, 9)\) from the following frame $f_{r+1}$ as follows,

$$x_{if} = \sum_{i=1}^{9} \beta_i \psi_i + n_2 = \Psi \beta + n_2.$$

(2)

Here, $\Pr(n_{2k}) = N(0, \sigma_n^2)$. The probability distribution of residual error $e$ can be calculated by (1) and (2),

$$\Pr(e = x_{ip} - x_{if}) = \frac{1}{(2\pi(\sigma_{x_1}^2 + \sigma_{x_2}^2))^L} \exp \left\{ -\frac{\|e - (\Phi \alpha - \Psi \beta)\|^2}{2(\sigma_{x_1}^2 + \sigma_{x_2}^2)} \right\}.$$

(3)

The mean square error is that

$$E(\|e\|^2) = (\sigma_{x_1}^2 + \sigma_{x_2}^2) \cdot L + \|\Phi \alpha - \Psi \beta\|^2.$$

(4)

On the basis of the assumption that the temporal symmetry between blocks in previous and following frames, the $x_{ip}$ should be similar to the $x_{if}$. Therefore, the weights $\alpha$ and $\beta$ can be computed by the minimum mean square error (MMSE) criterion. It is noted that the weights $\alpha$ and $\beta$ can not be zero vector since one candidate block at least contributes to the prediction of the interpolated block, so we need to add constrains on $\alpha$ and $\beta$. Above all, the optimal model is as follows,

$$\begin{align*}
\{\alpha, \beta\} &= \arg \min_{\alpha, \beta} \|\Phi \alpha - \Psi \beta\|_2^2, \\
\text{s.t. } u^T \alpha &= 1, \ u^T \beta = 1.
\end{align*}$$

(5)

Here, $u$ is a full-ones vector.

However, without prior knowledge of the ‘truth’, the model (5) often produces over-fitting. To reducing the bad effects caused by over-fitting, the most common approach is to regularize the MMSE model using Tikhonov regularization which imposes an $L_2$ penalty on the norm of $\alpha$ and $\beta$, that is,

$$\begin{align*}
\{\alpha, \beta\} &= \arg \min_{\alpha, \beta} \|\Phi \alpha - \Psi \beta\|_2^2 + \lambda(\|\Gamma_\alpha \alpha\|_2^2 + \|\Gamma_\beta \beta\|_2^2), \\
\text{s.t. } u^T \alpha &= 1, \ u^T \beta = 1,
\end{align*}$$

(6)

where $\Gamma_{\alpha}$ and $\Gamma_{\beta}$ are known as the Tikhonov matrix [9]. The $\Gamma_{\alpha}$ and $\Gamma_{\beta}$ terms allow the imposition of prior knowledge on the solution $\alpha$ and $\beta$. In our case, we can exploit the approach that the candidate matching blocks using the motion vectors closer to the true motion vectors should be given larger weight than the candidate matching blocks using the motion vectors far from the true motion vectors. It is obvious that if the motion vector is closer to the true motion vector, the candidate block in the previous frame found by it, is more similar to the corresponding candidate block in the following frame on account of temporal symmetry. Therefore, we proposed the diagonal $\Gamma_{\alpha}$ and $\Gamma_{\beta}$ in the form of

$$\Gamma_{\alpha} = \Gamma_{\beta} = \text{diag}(\|\varphi_1 - \psi_1\|_2^2, \|\varphi_2 - \psi_2\|_2^2, \ldots, \|\varphi_9 - \psi_9\|_2^2).$$

(7)

With this structure, $\Gamma_{\alpha}$ and $\Gamma_{\beta}$ penalize weights of large magnitude assigned to the candidate blocks using the motion vectors which are far from the true motion vectors.

The model (6) can be rewritten equivalently as

$$\hat{w} = \arg \min_{w} \left\{ \|Xw\|_2^2 + \lambda(\|\Gamma w\|_2^2) \right\},$$

$s.t. \ p^T w = 1, \ q^T w = 1,$

(8)

where $X = [\Phi, -\Psi]$, $\Gamma = \text{diag}(\Gamma_{\alpha}, \Gamma_{\beta})$, $w = [\alpha^T, \beta^T]^T$, $p = [u^T, 0^T]^T$, and $q = [0^T, u^T]^T$. Then, $\hat{w}$ can be calculated directly by the usual Tikhonov solution,

$$\hat{w} = -\frac{1}{2}(X^T X + \lambda \Gamma^T \Gamma)^{-1} (\mu_1 p + \mu_2 q),$$

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \left[ \begin{array}{c} p^T (X^T X + \lambda \Gamma^T \Gamma)^{-1} p \\ q^T (X^T X + \lambda \Gamma^T \Gamma)^{-1} q \end{array} \right]^{-1} \left[ \begin{array}{c} p^T \\ q^T \end{array} \right].$$

(9)

Here, $\lambda$ is a balance factor, and we use $\lambda = 0.25$ from point on.

Given the solution $\hat{w}$ of model (8), the interpolated block $x_i$ is constructed using the following formula,

$$x_i = \frac{1}{2}(\Phi \alpha + \Psi \beta) = \frac{1}{2}(\Phi, \Psi)\hat{w}.$$

(11)

Finally, the vector $x_i$ is rearranged as block $B$, and repositioned to the interpolated frame.

3. Experimental Results

In our simulations, a full-search block matching algorithm is used for BME between two adjacent frame images. The block size is set 16 x 16, the radius of the search range in forward ME is 16 pixels and the radius of the search range in BME refinement is 2 pixels. The performance of 8J-MCI has been evaluated using 4 test sequence, Mobile, Bus, Football and Foreman, which are in the standard CIF format and 30Hz frame rate. Every even frame of the first 100 frames in each test sequence is dropped and interpolated by the proposed algorithm.

3.1 Evaluation of the Robustness to Wrong Motion Vectors

In order to evaluate fault-tolerance for motion vectors, the
Table 1  The evaluation of fault-tolerancy for the proposed 8J-MCI.

|               | Mobile | Bus | Football | Foreman | Average |
|---------------|--------|-----|----------|---------|---------|
| MCI           | 26.19  | 25.68| 21.92    | 33.82   | 26.90   |
| MCI+MA in [3] | 26.39  | 26.40| 22.20    | 34.24   | 27.31   |
| MCI+MA in [4] | 28.77  | 25.11| 21.82    | 33.19   | 27.22   |
| MCI+MA in [5] | 28.79  | 26.69| 22.23    | 34.40   | 28.03   |
| 8J-MCI        | 28.67  | 28.02| 22.96    | 34.81   | 28.62   |

8J-MCI is compared with the combination of traditional MCI and popular MA algorithms in [3]–[5], and these compared methods use the same motion vector field estimated by BME as 8J-MCI. The average peak signal-to-noise ratios (PSNRs) of the 50 interpolated frames within each test sequence are presented in the Table 1. It can be observed that the performance of 8J-MCI is a little better than traditional MCI using MA algorithms in addition to the Mobile sequence. Since the Mobile sequence containing simple and slow motions has a flat motion field, the motion vector of the interpolated block is quite close to the motion vectors of its 8-neighbor blocks so that the interpolated block is less affected by its surrounding motion vectors. Therefore, for the Mobile sequence, the performance of our methods is basically close to the compensation method using the median vector filter [4] or weighted median vector filter [5].

### 3.2 Subjective Evaluation

In comparison with traditional compensation methods MCI, OBMC [6] and AOBMC [7], the subjective quality of the interpolated frame predicted by 8J-MCI is shown in Fig. 2. Note that all compared compensation methods use the same motion vector field estimated by BME as 8J-MCI. It can be obviously seen from Fig. 2 that the 13th frame containing medium complexity motions in Foreman recovered by MCI has some blocking artifacts (highlighted in red circle). Both OBMC and AOBMC suppress blocking artifacts to some degree but result in blurring or over-smoothing artifacts. The proposed 8J-MCI not only reduces blocking artifacts but also has no over-smoothing artifacts. In addition, 8J-MCI also obtains the highest PSNR in all compensation methods.

### 3.3 Objective Evaluation

Three FRUC algorithms are selected as benchmarks, including the well-known 3DRS [1], DualME [2] and MAAR (one of the state-of-the-art FRUC schemes) [8]. In the experiments, the motion search range for MAAR and DualME is set to $17 \times 17$. The ME block size used in these three benchmarks is $8 \times 8$. The average PSNRs of the 50 interpolated frames are shown for each test sequence in Table 2. It can be observed that 8J-MCI has the higher PSNR than other algorithms for all sequences except Foreman. The Foreman sequence contains slight wobbles in the background, but our method does not take the global motion in account so that the quality of the interpolated frame decays in a certain degree. The average PSNRs of different algorithms are also presented in Table 2. It can be seen that the proposed method 8J-MCI obtains the highest PSNR among all algorithms.

### 4. Conclusions

In this paper, we presented a novel motion compensation method for FRUC. The proposed algorithm does not use single motion vector to perform block-based MCI unlike most existing ones, but adopts motion vectors of the interpolated block and its 8-neighbor blocks to jointly interpolate the target block. The joint compensation approach can interpolate the target block while correcting the mistakes existing in the motion vector field. Experimental results show that the proposed method is not only robust to motion vector estimated wrongly, but also can to reduce blocking artifacts in comparison with existing popular compensation methods. In addition, the FRUC algorithm comprised of BME and our joint compensation method outperforms MAAR (one of the state-of-the-art FRUC schemes) in the average PSNR for the test image sequence.

### Acknowledgments

This work was supported by the Graduate Student Innovation Project of Jiangsu Province China (CXZZ12_0466 and CXZZ11_0390), National Natural Science Foundation of China (61071091), Natural Science Foundation of the
Higher Education Institutions of Jiangsu Province China (12KJB510019), and Technology Research Program of Hubei Provincial Department of Education (D20121408).

References

[1] G.D. Haan, P.W.A.C. Biezen, and H. Huijgen, et al., “True motion estimation with 3-D recursive search block matching,” IEEE Trans. Circuits Syst. Video Technol., vol.3, no.5, pp.368–379, 1993.

[2] S. Kang, S. Yoo, and Y. Kim, “Dual motion estimation for frame rate up-conversion,” IEEE Trans. Circuits Syst. Video Technol., vol.20, no.12, pp.1909–1914, 2010.

[3] B.T. Choi, S.H. Lee, and S.J. Ko, “New frame rate up-conversion using bi-directional motion estimation,” IEEE Trans. Consum. Electron., vol.46, no.3, pp.603–609, 2000.

[4] J. Zhai, K. Yu, and S. Li, “A low complexity motion compensated frame interpolation method,” Proc. IEEE International Symposium on Circuits and Systems, pp.4927–4930, Kobe, Japan, May 2005.

[5] J. Ascenso, C. Brites, and F. Pereira, “Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding,” 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services, pp.1–6, Smolenice, Slovak Republic, July 2005.

[6] M. Orchard and C. Sullivan, “Overlapped block motion compensation: An estimation-theoretic approach,” IEEE Trans. Image Process., vol.3, no.9, pp.693–699, 1993.

[7] B. Choi, J. Han, and C. Kim, 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.7, pp.407–416, 2007.

[8] Y. Zhang, D. Zhao, and S. Ma, et al., “A motion-aligned autoregressive model for frame rate up-conversion,” IEEE Trans. Image Process., vol.19, no.5, pp.1248–1258, 2010.

[9] A.N. Tikhonov and V.Y. Arsenin, ed., Solutions of ill-posed problems, V.H. Winston & Sons, Washington D C, USA, 1977.