

Improved matching criterion for frame rate upconversion with trilateral filtering

X. HoangVan, J. Ascenso and F. Pereira

Frame rate upconversion (FRUC) is an important post-processing technique to enhance the visual quality of low frame rate video. A major, recent advance in this area is FRUC based on trilateral filtering which novelty mainly derives from the combination of an edge-based motion estimation block matching criterion with the trilateral filter. However, there is still room for improvement, notably towards reducing the size of the uncovered regions in the initial estimated frame, this means the estimated frame before trilateral filtering. In this context, proposed is an improved motion estimation block matching criterion where a combined luminance and edge error metric is weighted according to the motion vector components, notably to regularise the motion field. Experimental results confirm that significant improvements are achieved for the final interpolated frames, reaching PSNR gains up to 2.73 dB, on average, regarding recent alternative solutions, for video content with varied motion characteristics.

Introduction: Frame rate upconversion (FRUC) plays an important role in multiple applications such as film and video conversion, displays technology and video transmission. A large number of solutions have been proposed in recent years to improve the FRUC performance [1–4]. Among these solutions, FRUC with trilateral filtering [3] has shown superiority owing to its simple architecture and high performance. The trilateral filtering FRUC approach brings improvements due to the following key contributions: 1. unidirectional motion estimation (ME) with an edge-based matching criterion; and 2. unreliable pixel refinement with trilateral filtering. The unidirectional ME approach (in both temporal directions) can avoid the bidirectional ME problems when motion is nonlinear, but introduces undesirable holes (uncovered pixels) and overlapping regions (unreliable pixels); the trilateral filter is able to efficiently estimate these uncovered and unreliable pixels. However, the trilateral filtering FRUC solution may still be improved by reducing the amount of uncovered pixels, thus increasing the quality of the so-called *initial estimated frame*, X_t^I , which strongly affects the final interpolated frame quality. Moreover, the final estimated frame is also improved due to the reduction of the number of unreliable pixels.

In this context, this Letter proposes a novel unidirectional ME matching criterion that aims to reduce the size of uncovered and unreliable regions, leading to significant improvements in the final interpolated frame quality. In the proposed criterion, the motion vector components (MVCs) are exploited to favour the motion vectors (MVs) closer to the search centre, regularising the motion vector field (MVF). The proposed solution brings major quality improvements with negligible additional complexity and can be easily integrated in the trilateral filtering FRUC architecture.

Trilateral filtering FRUC: From the FRUC solutions available in the literature, the trilateral filtering FRUC solution proposed by Wang *et al.* generates high quality interpolated frames [3] (see architecture in Fig. 1) and thus has been adopted as the starting point in this Letter. In this solution, motion estimation and compensation are implemented in both the forward and backward directions and fused together to create the initial estimated frame. Then, this frame is refined by means of a so-called *trilateral filter* where both temporal and spatial correlations are exploited to estimate the uncovered and unreliable pixels.

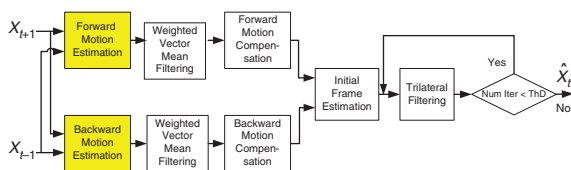


Fig. 1 Frame rate upconversion with trilateral filtering architecture, high-lighting improved modules

The first step in the adopted FRUC architecture is forward and backward ME in order to obtain two MVFs between the two neighbouring frames (X_{t-1} , X_{t+1}) and the target frame that needs to be estimated. In

the unidirectional ME block matching criterion, both the mean absolute difference (MAD) of the luminance and edge information are used [3]. The two MVFs are then refined using a weighted MV mean filter that further improves the MV accuracy. After, motion compensation creates the forward and backward estimated frames which are combined to create the initial estimated frame; Fig. 2 shows two initial estimated frame examples. Then, trilateral filtering generates the final interpolated frame by first identifying the unreliable and uncovered pixels with a mask expressing the reliability of each pixel in X_t^I . After, weighting parameters are computed according to the distance and correlation between matching blocks. Finally, the blocks containing uncovered and unreliable pixels are improved using temporal and spatial adjacent blocks and their associated weighting parameters.

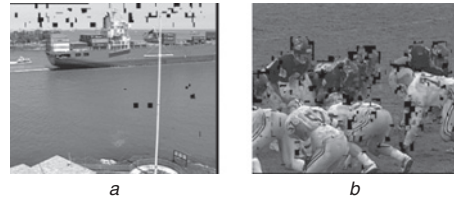


Fig. 2 Initial estimated frames examples (CIF at 30 Hz)

a Container, first estimated frame
b Football, first estimated frame

Proposed ME block matching criterion: As shown in Fig. 2, the initial estimated frames obtained with the original trilateral filtering FRUC solution still contain a large amount of uncovered pixels which leads to lower final interpolated frame quality. These uncovered pixels mainly derive from the ill-posed unidirectional ME problem and errors in the MVF estimation. To reduce the size of the uncovered regions and simultaneously improve the MVF accuracy, this Letter proposes a novel ME block matching criterion with two major components:

- Combined luminance and edge error metric:** The luminance and some edge information are combined to define an error metric. The edges correspond to salient information in the image and can help to improve the MVF accuracy, thus increasing the quality of the forward and backward estimated frames.
- Motion vectors based weighting term:** The combined luminance and edge error metric is weighted with a function of the MVC to favour MVs closer to the centre of the search area as this improves the interpolated frame quality, notably for uniform background and regular textured regions. Thus, a penalty is introduced when MVs assume extreme positions in the search range as these extreme positions are responsible for many uncovered regions. This approach was inspired by the weighted mean absolute difference (WMAD) criterion proposed in [5] with a different purpose, notably to reduce the computational complexity.

In this context, the proposed unidirectional ME matching criterion, named *motion vector weighted mean absolute difference* (MVMAD), corresponds to the minimisation of the following error metric:

$$\rho_{MVMAD}(\Delta x, \Delta y) = \frac{1}{N} \left(\sum_{x,y \in B} \{ |X_{t-1}(x, y) - X_{t+1}(x - \Delta x, y - \Delta y)| + \alpha |DX_{t-1}(x, y) - DX_{t+1}(x - \Delta x, y - \Delta y)| \} * (1 + k * (\Delta x^2 + \Delta y^2)) \right) \quad (1)$$

where (x, y) denotes the pixel location in block B of X_{t-1} (for forward ME) and X_{t+1} (for backward ME), $(\Delta x, \Delta y)$ are the MV components and (DX_{t-1}, DX_{t+1}) are the high frequency components extracted using a highpass filter D over the two reference frames X_{t-1} , X_{t+1} [3]. N is the number of pixels in block B and α are a pre-defined parameter allowing to trade-off the impact of luminance and the edge information that has been experimentally set to 0.3 [3].

In the proposed MVMAD criterion, the sum of the squared MVC is used to weight the combined MAD of the luminance and edge information with the help of a k parameter that is used to adjust the impact of the MVC penalty on the overall error metric. A higher k value will attract more of the MVs to the search centre and vice versa; the simulations suggest that k should be set around 0.02. This proposed

ME block matching criterion is integrated both in the forward and backward ME processes. As there is no significant additional complexity associated to this technique, it is very suitable for integration in the trilateral filtering FRUC solution.

Performance evaluation: In this Letter, the proposed solution (labelled Proposed) is compared against the solutions proposed in [3] and [4] (labelled Ref [3] and Ref [4]). The Ref [4] solution is based on MV smoothing, selective average based motion compensation and weighted hole interpolation [4]. Moreover, due to the lack of information for some main parameters for Ref [3] in [3], this Letter also includes results for an implementation by the authors of Ref [3], using the parameters corresponding to the obtained best performance (labelled as Replica [3]). For the performance evaluation, six video sequences at CIF spatial resolution (352×288) and 30 Hz, with 101 frames each, were evaluated. The search range was set to ± 50 as in [3].

The results in Table 1 show that a significant reduction is achieved in the percentage of uncovered pixels with the proposed matching criterion (no results are shown for Ref [3] and Ref [4] as they are not available). This can be confirmed in Fig. 3 where the uncovered regions in the initial estimated frames have been greatly reduced.

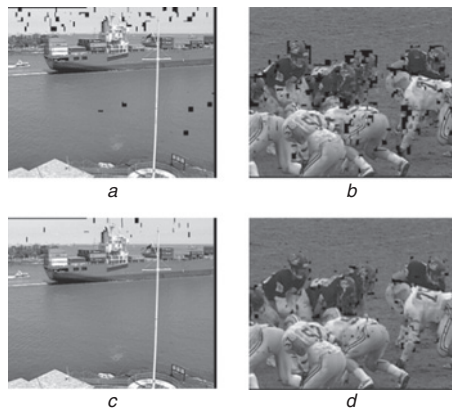


Fig. 3 Initial estimated frame quality assessment (CIF at 30 Hz)

- a Container, first estimated frame, **replica [3]**
b Football, first estimated frame, **replica [3]**
c Container, first estimated frame, **proposed**
d Football, first estimated frame, **proposed**

Table 1: Percentage of uncovered pixels in initial estimated frames

Sequences	Football	Container	Mobile	Tennis
Replica [3]	8.25	1.29	7.86	2.45
Proposed	3.38	0.21	0.85	0.81

Table 2 shows the PSNR results for the final estimated frames. The proposed criterion in (1) achieves significant improvements regarding the Ref [3], Ref [4] and Replica [3] solutions with average gains of 1.31, 1.47, and 2.73 dB, respectively. This improvement is a direct consequence of the reduction of the uncovered regions' size in the initial

estimated frames and also of the unreliable regions' size (indirectly assessed with the final estimated frame PSNR).

Table 2: PSNR of final estimated frames (dB)

Sequences	Ref [3]	Ref [4]	Replica [3]	Proposed
Football	22.74	22.17	22.18	22.79
Container	41.87	43.13	42.27	44.01
Mobile	25.09	25.98	21.31	27.19
Tennis	29.58	27.95	27.77	29.53
Flower	26.96	27.92	27.79	30.79
Paris	33.53	31.63	29.94	33.29
Average	29.96	29.80	28.54	31.27
Gains to	+1.31	+1.47	+2.73	

Conclusion: This Letter proposes an improved ME block matching criterion for a trilateral filtering FRUC solution. The proposed criterion adopts a combined luminance and edge information error metric which is weighted with a function of the MV components to regularise the MVF, notably favouring MVs closer to the search centre to reduce the size of the uncovered and unreliable regions in the initial estimated frame. The proposed criterion results in significant performance improvements in the quality of the interpolated frames with rather negligible complexity cost.

© The Institution of Engineering and Technology 2013

10 October 2012

doi: 10.1049/el.2012.3454

One or more of the Figures in this Letter are available in colour online.

X. HoangVan and F. Pereira (*Instituto Superior Técnico, Lisbon, Portugal*)

E-mail: hoang.xiem@lx.it.pt

J. Ascenso (*Instituto de Telecomunicações, Lisbon, Portugal*)

X. HoangVan and F. Pereira: Also with the Instituto de Telecomunicações

J. Ascenso: Also with the Instituto Superior de Engenharia de Lisboa, Lisbon

References

- Wong, C. K., Au, O. C., and Tang, C. W.: 'Motion compensated temporal interpolation with overlapping'. Proc. IEEE Int. Symp. Circuits and Systems, Atlanta, CA, USA, 1996, Vol. 2, pp. 608–611
- Choi, B.-T., Lee, S.-H., and Ko, S.-J.: 'New frame rate upconversion using bi-directional motion estimation', *IEEE Trans. Consum. Electron.*, 2000, **46**, (3), pp. 603–609
- Wang, C., He, Y., and Tan, Y.-P.: 'Frame rate upconversion using trilateral filtering', *IEEE Trans. Circuits Syst. Video Technol.*, 2010, **20**, (6), pp. 886–893
- Jung, H. S., Kim, U. S., and Sunwoo, M. H.: 'Simplified frame rate upconversion algorithm with low computational complexity'. Proc. 20th European Signal Processing Conf., Romania, August 2012, pp. 385–389
- Castagno, R., Haavisto, P., and Ramponi, G.: 'A method for motion adaptive frame rate upconversion', *IEEE Trans. Circuits Syst. Video Technol.*, 1996, **6**, (5), pp. 436–446