Frame Interpolation Using Phase Information and Guided Image Filtering

Fahim Arif[1], Sundas Amin[1], Abdul Ghafoor[1](✉), and M. Mohsin Riaz[2]

[1] National University of Sciences and Technology (NUST), Islamabad, Pakistan
fahim@mcs.edu.pk, sundusamin@yahoo.com, abdulghafoor-mcs@nust.edu.pk
[2] COMSATS University, Islamabad, Pakistan
mohsin.riaz@comsats.edu.pk

Abstract. Videos with low frame rate lacks the visual quality element and unable to meet the standards of new multimedia systems. In this paper, a technique for frame interpolation utilizing phase information is proposed. Phase information gives the intuition that the motion of signals can be depicted as a phase shift. The two consecutive input frames of video are passed through a guided filter to preserve edges of objects in frames. These frames then decompose into multi scale pyramid and the difference in each pixel is calculated to compute the phase difference which then used to interpolate the in-between frame. The proposed technique can be used to increase the frame rate of videos. Subjective and objective comparison is performed with the state of art existing technique to prove the significance of proposed technique.

1 Introduction

Frame interpolation becomes the significant part of video processing. Frame interpolation is the process of generating intermediate frame based on the existing frames from the video sequence. Video frame interpolation plays an important function in increasing the frame rate of video and presents the smoother video playback. Traditional approaches to frame interpolation used two steps to interpolate a frame; first step is to compute the correspondences using stereo methods or leveraging (optical flow), and perform image warping based on correspondences. Generally correspondences computation is called motion estimation and image warping and blending of input frame is frame interpolation. Intermediate frame generation is a very complex task because of moving objects, object occlusion and sudden light changing.

Kim et al. [1] proposed hierarchical motion estimation method to track the motion. Modified 3-D recursive search along with the motion vector refining process is performed to achieve the higher efficiency. *Gracewell et al.* [2] proposed fast forward motion estimation method for the frame rate up conversion. Overlapped blocked motion compensation approach is used to reduces the blocking artifacts. Motion vectors are refined using spatial correlation between the neighbouring blocks. *Philip et al.* [3] proposed a modification in block matching technique

E. Ricci et al. (Eds.): ICIAP 2019, LNCS 11752, pp. 249–259, 2019.
https://doi.org/10.1007/978-3-030-30645-8_23

to increase the quality of video playback. Dynamic looping refined the motion vectors acquired from block matching. *Dikbas et al.* [4] proposed a true motion estimation technique with the low computational complexity using predictive search. Blocking artifacts catered at the interpolation stage are minimized by extracting the dense motion field in backward and forward direction. *Liu et al.* [5] trained a deep network by combining the benefits of optical flow methods and neural network methods. Missed frames were calculated by flowing pixel values in the available frames. The blurring artifacts were removed in most of the cases but constructing frames by using RGB differences failed to interpolate the true frame when scenes are repetitive.

Li et al. [6] proposed optical flow based video interpolation technique and used Laplacian cotangent mesh to preserve the smoothness in interpolated frames by minimizing differentials in mesh. Creating accurate mesh for the image of interest is complex and takes more computation time. *Huang et al.* [7] used the predictive square search motion estimation technique to refine the motion vectors. Markov random field correction and block refinement were incorporate to enhance the interpolated frames quality. The algorithm maintains to achieve better quality but it expend extensive amount of resources. *Xiao et al.* [8] proposed an algorithm to lessen the ghost and blurring artifacts. The algorithm performed classification of the edge blocks and flat blocks based on depth map threshold value and then the motion estimation was performed distinctively on blocks. Hole filling based interpolation was effectuated to remove the holes artifact from the interpolated frame.

Zhao et al. [9] presented a motion vector refinement method to correct the false motion vectors for the problems of holes and overlap. Motion vectors of holes and overlapping regions were obtained by constructing label array. The algorithm achieved the better results by removing holes but the blurriness near edges still exist. *Ji et al.* [10] presented a hybrid motion estimation technique by using spatial and temporal information. The algorithm was blend of unidirectional and bidirectional sum of absolute differences performed on multi resolution frames. However, the technique has the blurring artifact when the scene changes quickly. *Matsuo et al.* [11] proposed contrast compensation based linear filtering interpolation to reduce motion blur. This technique utilized spatiotemporal contrast information to enhance the frame rate of video. The algorithm improved the quality of interpolated frames however, this method induced other artifacts. *Lim et al.* [12] proposed motion compensated method based on region segmentation. Motion estimation step was based on the shapes that created arbitrarily by using pixel intensity. This makes this method computationally complex. *Jim et al.* [13] proposed a technique to reduce block artifacts in interpolated frames by estimating the true motion vectors. Sum of absolute differences (SAD) were calculated using three consecutive frames. However, SAD alone does not give the reliable results.

Kovacevic et al. [14] presented a block matching correlation method for frame rate up conversion. True motion vectors were calculated using phase plane correlation. However, this approach does not consider occlusion handling feature.

Okade et al. [15] proposed a technique to remove outlier motion vectors by using weighted vector median filter. Motion vector field was smoothed by median filtering to remove false motion vectors. The algorithm improves motion estimation but because of computational complexity unable to use in real-time applications. *Qu et al.* [16] proposed a post processing technique to extract refined motion vectors. Combination of bilateral and unilateral motion estimation techniques was introduced and the outliers of both were removed using vector extrapolation and weighted summation. However, motion jerkiness is still the issue in videos in which scenes are swiftly changing. *Lu et al.* [17] proposed a frame rate up conversion (FRUC) method utilizing spatial and temporal information of missing frames and integrated with HEVC coding for the limited bandwidth communication channels. However, the quality of reconstructed frames depends on accuracy of spatial and temporal information. *Umnyashkin et al.* [18] proposed a motion compensation technique utilizing hexagonal blocks. Motion estimation process was performed using mesh search structure instead of full search to lessen the computational load but the motion regions still have jerkiness.

Qu et al. [19] presented a non-integer FRUC technique to reduce motion blurriness and jerkiness in fast motion regions. The algorithm based on decimal multiples interpolation which is computationally expensive. *Lu et al.* [20] proposed a motion vector processing technique using artifact information metric to extract true motion vectors. It performed processing on those motion vectors that were unreliable and resulted in artifacts in frame interpolation. However, the complexity of this method is high if large numbers of initial motion vectors are unreliable. *Lu et al.* [21] also proposed the multi-frame based FRUC method and estimates motion vectors using unidirectional approach in both forward and backward directions. The motion dubiety regions were also handled using occlusion handling process. The computational complexity of method increase with the increase in occluded regions. *Kim et al.* [22] used prediction method motion vector smoothing in motion estimation process. It uses bidirectional motion estimation as a base algorithm which lacks to estimate reliable motion trajectories. *Dong et al.* [23] presented an image interpolation method which is based on sparse representation and depends on data fidelity term that fails to restrain the image local structures that lead to the increase in missing pixels.

In this paper, a technique for frame interpolation utilizing phase information is proposed. Phase information gives the intuition that the motion of signals can be depicted as a phase shift. The two consecutive input frames of video are passed through a guided filter to preserve edges of objects in frames. These frames then decompose into multi scale pyramid and the difference in each pixel is calculated to compute the phase difference which then used to interpolate the in-between frame. The proposed technique can be used to increase the frame rate of videos. Subjective and objective comparison is performed with the state of art existing technique to prove the significance of proposed technique.

Algorithm 1. Frame interpolation algorithm using phase information

```
Input frames $F_1, F_2$
Interpolation parameter $\alpha$
Decomposition using steerable pyramid: $(PD_1, PD_2) \leftarrow decompose(F1, F2)$
Phase computation: $(\phi_1, \phi_2) \leftarrow phase(PD_1, PD_2)$
Amplitude extraction: $(A_1, A_2) \leftarrow amplitude(F1, F2)$
Phase difference computation: $\phi_{diff} \leftarrow phase(\phi_1, \phi_2)$
$\phi_{diff} = attan2(sin(\phi_1 - \phi_2), cos(\phi_1 - \phi_2))$
$l=$ current level; $L=$ total levels of steerable pyramid
for all $l$ = $L$-1:1 do
    $\tilde{\phi}^{l}_{diff} \leftarrow$ correction of shift $(\tilde{\phi}^{l+1}_{diff})$
end for
to achieve smooth interpolation between phase $\phi_1$ and phase $\phi_2$:
$\hat{\phi}_{diff} \leftarrow phaseadjustment(\phi_{diff}, \tilde{\phi}_{diff})$
Interpolation of phase: $\phi_\alpha \leftarrow phaseinterpolation(\phi_1, \hat{\phi}_{diff}, \alpha)$
Output frame $F_{12}$
```

2 Proposed Methodology

The algorithm takes two consecutive frames input and performed three major steps (guided filtering, phase based interpolation and restoration).

2.1 Guided Filtering

Edge preserving guided filtering [24] transmit the guidance image structure to the filtering image to preserve edges. The proposed algorithm takes input of two consecutive frames. The filter is applied on first frame F_1 and extract the base layer and then it is applied on second frame F_2 to extract base layer of second frame. After that respective input frames are used to extract the detail layers using base layers. The edge-preserving smoothing is performed by decomposition of an frame F_1 into two layers i.e. [24],

$$F1 = O_x + t \tag{1}$$

where O_x is filtered output image and t is texture image. O_x is known as a base layer and t is a detail layer. The concept of image guided filter is based on the assumption that filtering output image O is a linear transform of guidance image G in a window w_y which is centred by the pixel y,

$$O_x = a_y G_x + b_y \forall x \in w_y \tag{2}$$

where a_y and b_y are linear coefficients and are assumed to remain constant in w_y. The efficient frame interpolation algorithm used the guided filter to decompose input frames into base layers and detail layers. The base layer is obtained by

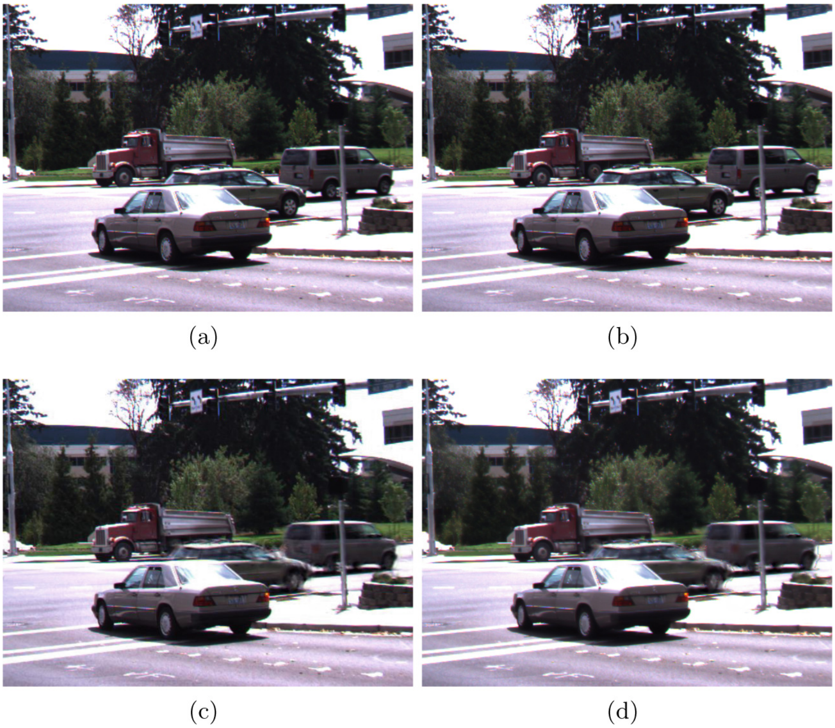


Fig. 1. Example 1: (a, b) input frames (c) *Meyer et al.* [25] (d) proposed

applying guided filter to the input frame. After this, base layer O_x is used to obtained the detail layer t as,

$$t = F1 - O_x \tag{3}$$

In the next step, extracted base layers and detail layers of input frames are forwarded to the phase information algorithm to interpolate the intermediate frame.

2.2 Frame Interpolation Using Phase

After separating base layer and detail layer of input frames using guided filtering next step is interpolation. The phase based algorithm [25] takes base layers of frames as an input and then performs decomposition of frames using steerable pyramids. The steps of phase based algorithm are summarized in the Algorithm 1.

2.3 Joint Image Restoration

After frame interpolation, post-processing step is performed to improve the visually appealing quality of frame. To cater the light issues gradient loss and vari-

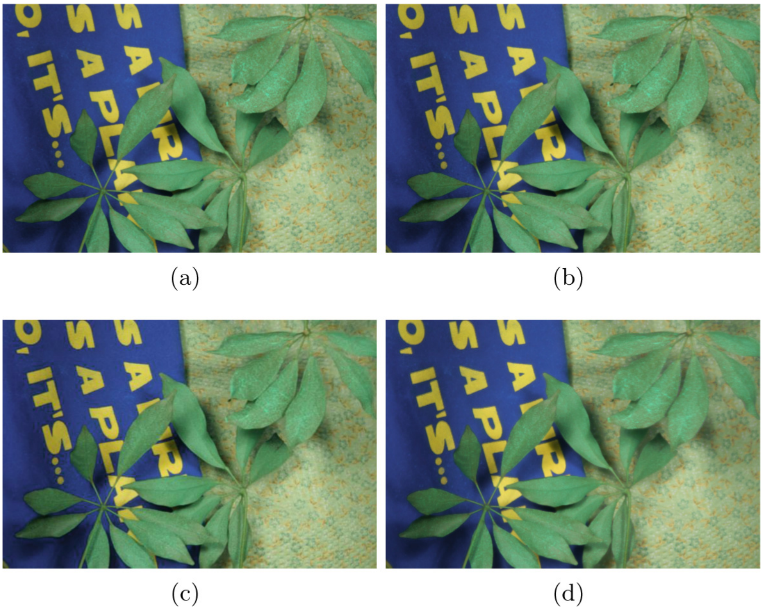

(a) (b) (c) (d)

Fig. 2. Example 2: (a, b) input frames (c) *Meyer et al.* [25] (d) proposed

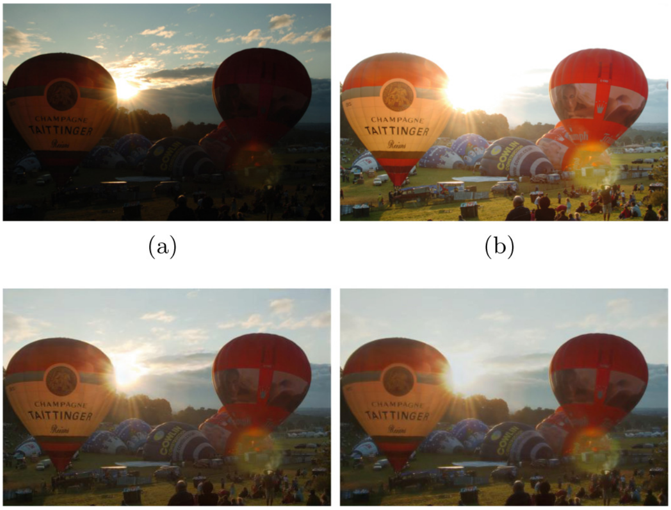
ation in gradient magnitude. The proposed efficient technique used joint image restoration algorithm [26]. The problems produced by intrinsic discrepancy in structure of images are called cross field problem [26]. Simple image filtering algorithm might create weak edges because of smoothing property of filter, even if gradients of reference image transferred to the noisy field outcome might appears artificial.

To solve this problem algorithm [26] construct a map. This map apprehend the structure discrepancy between both images. On the basis of this map an optimal scale map is derived to preserve edges and manipulate reference strength in the frames when sudden change in colors or brightness occur.

The algorithm takes input of interpolated frame F_{12} which might have noise and a reference frame R. In proposed algorithm $F2$ is considered as a reference frame. The algorithm recovered a frame from F_{12} with retained structure and removed noise. A map m is introduced with size of R. To estimate the map m and restore frame F, the main objective function is expressed as

$$DT(m, F) = DT_1(m, F) + \lambda DT_2(F) + \beta DT_3(\nabla m) \tag{4}$$

where DT_1, DT_2 and DT_3 are data terms used to remove various outliers. λ and β are parameters, λ is a confidence constraint and β controls smoothness of map m. DT_1 is used to remove outliers between map m and frame F using reference frame R. DT_2 is defined to control the wild difference between the noisy input frame F_{12} and restore frame F. DT_3 is used to produce the smooth



(a) (b)

(c) (d)

Fig. 3. Example 3: (a, b) input frames, (c) *Meyer et al.* [25], (d) proposed

map m with strong edges. The iterative method is proposed [26]. To solve function $DT(m, F)$, iterative re-weighted least squares (IRLS) [26] approach is used. After this step the proposed efficient frame interpolation algorithm is completed and the final interpolated frame having better edges is achieved especially the proposed approach handles the abrupt light changes more sophistically.

3 Experiment and Results

To test the abrupt change in brightness and light MEF datasets [27] is used which includes the images captured belong to the same scene but under different light condition. Quantitative analysis of the proposed and existing techniques are performed using Peak Signal to Noise Ratio (PSNR) and Structure Similarity Image Measure (SSIM).

Figure 1(a, b) shows the input frames of dump-truck sequence. Figure 1(a) is a frame in which the cars and truck are visible, the frame is challenging because it has brightness which is difficult to handle while interpolation because of not showing the correct pixels. Figure 1(b) is a frame in which cars and truck are move little forward with respect to their position. Both input frames have trees which shows the correct interpolation should also interpolate the trees movement in a natural way. Figure 1(c, d) shows the results of phase based approaches. Figure 1(c) shows the frame of existing technique *Meyer et al.* [25] interpolation result, which have blurriness around the edges of cars. Trees are also lacking the same colors of input frames. Figure 1(d) shows the proposed interpolation result in which the edges of cars and truck is more clear and blurriness is visibly reduced. The lines on roads are also more clear. Figure 2(a, b) show input

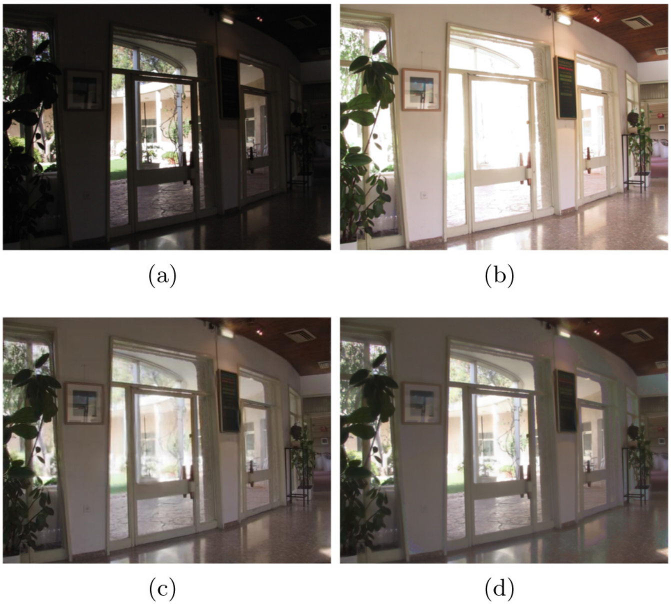


Fig. 4. Example 4: (a, b) input frames, (c) *Meyer et al.* [25], (d) proposed

frames of leaves which are placed on a two types of surface text and design of very short flowers, the frame have variety of different colors. Figure 2(b) is a frame in which leaves and the surface is moving clockwise, because the movement of text surface and the design surface it is a challenging frame sequence to interpolate. Figure 2(c, d) shows the interpolation results of phase based approaches. Figure 2(c) represents the frame of existing technique *Meyer et al.* [25] result, which shows blurry edges around the leaves. Surface also lacks the clarity of edges specially near text. Figure 2(d) represents the proposed interpolation result in which leaves edges are more sharp and blur factor is significantly reduced. The surface of text is also interpolated visually better result.

Figure 3(a) shows the frame which is underexposed and have very less information and Fig. 3(b) shows the frame which reveals the severe illumination change occur in scene. The first input frame is not able to show the writing on the balloons and colors of balloons because of very less lighting are also not visible in the first input frame. The second input frame shows the information better than the first frame but it has the effect of severe lighting and the edges of balloons exposed to sun are also not visible. Figure 4(c) shows the result of the *Meyer et al.* [25] and Fig. 3(d) shows the result of the proposed algorithm. It can be perceived that the frame created by proposed algorithm handles light change sophistically.

Figure 4(a) shows a frame which shows a frame from the inside of the building and some of the outside information is also visible. It has least information most of the scene is covered under darkness and the region around the door

Table 1. PSNR and SSIM comparison

Examples	*Meyer et al.* [25]		Proposed	
	PSNR	SSIM	PSNR	SSIM
Example 1	26.4819	0.9583	27.2222	0.9494
Example 2	31.4658	0.9749	32.0195	0.9868
Example 3	19.4597	0.8675	21.2116	0.8916
Example 4	22.6511	0.8714	22.4293	0.8500

is not showing proper edges. The inside plants are also not visible. It shows the region which is outside the building. Figure 4(b) shows the frame which has overexposure to light and most of the information is invisible because of light. The region outside the building is totally not visible because of the severe sun exposure. The region inside the building in clear in this frame and indoor plants are also clearly visible. This abrupt change in light becomes a challenge to interpolate in-between frame. Figure 4(c) *Meyer et al.* [25] represent the result of phase based approach which generate the better result and the Fig. 4(d) represents the proposed algorithm which represents the better edges, cater the sudden light exposed regions better as compared to existing phase based approach. The region around the door is more smooth and exhibiting sharp edges near light changing regions. The information shown outside the frame is also clearly visible in the proposed interpolated frame (Table 1).

4 Conclusion

A technique for frame interpolation utilizing phase information is proposed. Phase information gives the intuition that the motion of signals can be depicted as a phase shift. The two consecutive input frames of video are passed through a guided filter to preserve edges of objects in frames. These frames then decompose into multi scale pyramid and the difference in each pixel is calculated to compute the phase difference which then used to interpolate the in-between frame. The proposed technique can be used to increase the frame rate of videos. Subjective and objective comparison is performed with the state of art existing technique to prove the significance of proposed technique.

References

1. Kim, D., Park, H.: An efficient motion-compensated frame interpolation method using temporal information for high-resolution videos. J. Disp. Technol. **11**(7), 580–588 (2015)
2. Gracewell, J., John, M.: Motion compensation based multiple inter frame interpolation. In: International Conference on Signal Processing and Communication, pp. 261–265 (2016)

3. Philip, J.T., Samuvel, B.: Digital video frame rate up-conversion based on modified block matching technique. Int. J. Res. Comput. Commun. Technol. Adv. Technol. **5**(3), 135–140 (2016)
4. Dikbas, S., Altunbasak, Y.: A novel true-motion estimation algorithm and its application to motion-compensated temporal frame interpolation. IEEE Trans. Image Process. **22**(8), 2931–2945 (2013)
5. Liu, Z., Yeh, R., Tang, X., Liu, Y., Agarwala, A.: Video frame synthesis using deep voxel flow. In: International Conference on Computer Vision, pp. 4473–4481 (2017)
6. Li, W., Cosker, D.: Video interpolation using optical flow and laplacian smoothness. Neurocomputing **220**, 236–243 (2017)
7. Huang, Y., Chen, F., Chien, S.: Algorithm and architecture design of multi-rate frame rate up-conversion for ultra-HD LCD systems. IEEE Trans. Circ. Syst. Video Technol. **27**(12), 2739–2752 (2016)
8. Xiao, Y., Liu, J., Qu, A., Guo, Z.: Edge-consistency based adaptive interpolation method for stereoscopic video frame rate up conversion. In: IEEE International Conference on Multimedia and Expo Workshops, pp. 85–90 (2017)
9. Zhao, Y., Sun, G., Liu, J., Ge, J., Wan, W., Yang, X.: An improved FRUC scheme based on motion vector refinement. In: International Conference on Signal Processing, pp. 981–986 (2014)
10. Ji, B., Li, R., Wu, C.: Spatial-temporal correlation based multi-resolution hybrid motion estimation for frame rate up-conversion. In: Sun, X., Chao, H.-C., You, X., Bertino, E. (eds.) ICCCS 2017. LNCS, vol. 10603, pp. 408–417. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68542-7_34
11. Matsuo, Y., Sakaida, S.: Frame-rate conversion method by linear-filtering interpolation using spatio-temporal contrast compensation. In: IEEE International Conference on Consumer Electronics, pp. 243–244 (2017)
12. Lim, C., Park, H.: A region-based motion-compensated frame interpolation method using a variance-distortion. IEEE Trans. Circ. Syst. Video Technol. **25**(3), 518–524 (2015)
13. Kim, J., Kim, D., Choi, J., Kim, J., Han, D.: A novel frame-rate up conversion using pseudo-true motion vector. In: IEEE International Conference on Consumer Electronics, pp. 546–547 (2015)
14. Kovacevic, V., Pantic, Z., Beric, A., Jakovljevic, R.: Block-matching correlation motion estimation for frame-rate up-conversion. J. Sig. Process. Syst. **84**(2), 283–292 (2016)
15. Okade, M., Biswas, P.: A novel motion vector outlier removal technique based on adaptive weighted vector median filtering for global motion estimation. In: Annual IEEE India Conference, pp. 1–5 (2013)
16. Qu, A., Liu, J., Wan, W., Xiao, Y.: A frame rate up-conversion method with quadruple motion vector processing. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1686–1690 (2016)
17. Lu, G., Zhang, X., Gao, Z.: A novel framework of frame rate up conversion integrated within HEVC coding. In: International Conference on Image Processing, pp. 4240–4244 (2016)
18. Umnyashkin, S., Sharonov, I.: Motion compensation in video compression using hexagonal blocks. SIViP **9**(1), 213–223 (2015)
19. Qu, A., Liu, J., Xiao, Y., Wan, W., Wang, Q.: Non-integer times frame rate up-conversion using reliable analysis of motion information. In: International Conference on Signal Processing, Communications and Computing, pp. 1–6 (2016)

20. Lu, Q., Wang, Y., Fang, X.: An artifact information based motion vector processing method for motion compensated frame interpolation. J. Disp. Technol. **10**(9), 775–784 (2014)
21. Lu, Q., Xu, N., Fang, X.: Motion-compensated frame interpolation with multiframe-based occlusion handling. J. Displ. Technol. **12**(1), 45–54 (2016)
22. Kim, U.S., Sunwoo, M.H.: New frame rate up-conversion algorithms with low computational complexity. IEEE Trans. Circ. Syst. Video Technol. **24**(3), 384–393 (2014)
23. Dong, W., Zhang, L., Lukac, R., Shi, G.: Sparse representation based image interpolation with non local autoregressive modelling. IEEE Trans. Image Process. **22**(4), 1382–1394 (2013)
24. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
25. Meyer, S., Wang, O., Zimmer, H., Grosse, M., Hornung, A.S.: Phase-based frame interpolation for video. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1410–1418 (2015)
26. Yan, Q., et al.: Cross-field joint image restoration via scale map. In: IEEE International Conference on Computer Vision, pp. 1537–1544 (2013)
27. Ma, K., Zeng, K., Wang, Z.: Perceptual quality assessment for multi-exposure image fusion. IEEE Trans. Image Process. **24**(11), 3345–3356 (2015)