Enhanced motion compensated frame interpolation using object layer inference

T.-S. Wang, K.-S. Choi, H.-S. Jang and S.-J. Ko

In motion compensated frame interpolation (MCFI), an interpolated frame is produced by placing objects according to their motion estimated between two neighbouring frames. However, multiple motion vectors associated with the background (BG) and foreground (FG) objects can pass through a single location in the interpolated frame. In such motion ambiguity (MA) regions, annoying artefacts can be induced by adopting the motion of the BG object instead of the FG object. Introduced is a novel MCFI algorithm which can accurately infer the motion of the FG object at the MA region with low complexity. Simulation results show that the proposed method achieves higher visual quality for the interpolated frame without annoying artefacts as compared with conventional MCFI algorithms.

Introduction: In various motion compensated frame interpolation (MCFI) techniques, it is critical to find the true motion trajectory of objects in a scene [1–3]. However, even though reliable motion is estimated in motion estimation (ME), the interpolated frame can be severely degraded by annoying artefacts such as blockiness and flickering artefacts. This is mainly because motion compensated interpolated (MCI) following ME in MCFI can utilise wrong motion information around the occlusion region and the motion ambiguity (MA) region. Since the occlusion including the covered and uncovered regions belongs to the background (BG) and is visible in only one of the two successive frames, unidirectional interpolation using the BG motion is more effective than bidirectional interpolation. For the MA region, the motion associated with the foreground (FG) object should be determined and used to interpolate the frame.

In [4], a motion vector (MV) having the lowest sum of absolute differences (SAD) is selected among the multiple MV candidates to solve MA on the assumption that the block having the FG object produces the lowest SAD. In [5], instead of selecting an MV explicitly, the median of pixel values associated with the multiple MV candidates is exploited for the interpolated pixel value. However, those conventional methods frequently produce the aforementioned annoying artefacts since they cannot actually discriminate the FG and BG layers.

In this Letter, we propose a simple but effective MA solution which can accurately infer the motion of FG and BG by exploiting the occlusion near the MA region. MCI using the proposed MA solution can significantly improve the visual quality of the interpolated frame by determining the accurate motion and adopting a relevant interpolation scheme for both the occlusion and MA regions.

Proposed MCI method: In Fig. 1b, it can be observed that a part of the BG layer associated with one of the multiple MVs (voc1) is always located between the occlusion xOC and the FG object x1 in both f−1 and f+1. This spatial configuration occurs whenever the magnitude of relative motion between the FG and BG layer, ∥vFG − vBG∥, is larger than the size of the FG object, since the FG object always moves away from the occlusion belonging to the BG layer. Based on the observation, in our proposed MCI method, the MA is solved by explicitly inferring the FG and BG layers from the detected occlusion information. Let OBG and OFG denote a set of the pixel positions corresponding to the occlusion and a set of the positions which the multiple MVs causing MA pass through in the following frame, respectively. We first find a position within the occlusion, xOC ∈ OBG, such that |xOC − x1|, x1 ∈ OFG, is minimal. Then, the element of OFG, farthest from xOC, is associated with the FG object. For the occlusion regions where no motion passes, the MV of the BG layer is also easily estimated using the neighbouring MV. For the example in Fig. 1, vOC1 is determined as the MV of the FG and used for the MA region, since |xOC − x1| < |xOC − x1|. Then, the hole caused by the uncovered region in f+1 is unidirectionally interpolated with vBG.

However, such a straightforward approach to the distance calculation is computationally expensive, since there are a number of pixels in OBG. For practical implementation, we propose a low complexity distance calculation method using a simple morphological operator as follows.

1. The detected occlusion area is first expanded by repeating the morphological dilation operation with the 3 x 3 structuring element. At each iteration, the dilated pixels are labelled with the number of the iteration as shown in Fig. 1c. This label value indicates the distance from the occlusion area.
2. When the MA occurs, OFG is obtained and the distance label at each position in OFG is examined. The pixel corresponding to the largest label value is utilised to interpolate the frame.
3. Step 2 is repeated until the overlapped area is completely interpolated.

Experimental results: In the proposed MCFI method, a slightly modified ME algorithm presented in our previous work [3] is performed to provide the true motion field for the layer inference-based MCI. To demonstrate the performance of correct MV inference of the proposed algorithm, the visual quality of its interpolated frames is compared subjectively to those of two conventional MCI methods, Wang’s algorithm [4] and Haan’s algorithm [5], for two test sequences ‘Tractor’ and ‘Basketball’. For the sequence ‘Tractor’, the chimney of a tractor has little motion whereas the background moves fast, thus the MA occurs on the chimney as shown in Fig. 2. The conventional methods without inferring the image layer information yield blockiness in the MA region and ghost artefacts in the occlusion. On the other hand, the proposed method successfully produces pleasing results in both the MA and occlusion regions. It is seen in Fig. 3 that the small fast-moving ball is accurately interpolated in the proposed method, whereas a ghost of the ball is produced or the ball even disappears in the conventional methods.

Fig. 1 Example of occlusion and MA

Large displacement of BG and FG in opposite direction produces MA
a 2D representation
b 1D representation
c Label value of each pixel by dilation operator

Problem statement: To construct the MV field, we employ the hierarchical ME approach [3] which can estimate the motion of objects with variable sizes between successive frames. Moreover, we utilise the occlusion estimation method [6] to identify the occlusion and to remove the inaccurate MVs caused by the occlusion. Thus, we can assume that the final MV field consists of accurate and reliable MVs. However, occasionally multiple MVs associated with the BG and FG objects can pass through a single location in the interpolated frame, and thereby result in MA. For normal regions associated with reliable MVs except the multiple MVs, the conventional bidirectional interpolation yields pleasing results without visible artefacts. However, the occlusion or MA regions should be carefully interpolated.

Fig. 1a demonstrates an example of the occlusion and MA where the BG moves left while the FG object slightly moves in the opposite direction. For better visualisation, MVs in the horizontal line is temporally illustrated in Fig. 1b. The multiple MVs, vFG and vBG in the MA region intersect each other even though the motion of both the BG and the FG is reliable. Such MA is frequently observed in the real video scene where a small object moves fast or the BG moves rapidly as the camera rotates around the main object.

To interpolate the MA region correctly, the motion of the FG object should be correctly chosen. If the motion of the BG layer is selected, severely annoying blockiness and flicker artefacts can be observed since FG object passing through the MA region suddenly disappear. In the case that the motion of the BG and FG is randomly selected for the region, the interpolated frame can contain a mix of FG and BG.
Conclusions: In this Letter, we propose a novel layer inference-based MCI method using the occlusion information. Based on the observation on the spatial configuration around the occlusion, pixels belonging to the FG layer can be inferred accurately from multiple pixel candidates using the distance between the pixel and the closest occlusion position. Experimental results confirm that the proposed method achieves much higher visual quality of the interpolated frame without annoying artefacts in both the MA and occlusion regions compared with the conventional MCI algorithm.

Acknowledgments: This research was supported by the Seoul Future Contents Convergence (SFCC) Cluster established by the Seoul R&BD Program (10570), by the Mid-career Researcher Program through the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (2010-0000449), and by a Korea University Grant.

References

Fig. 2 Result for ‘Tractor’ sequence

- a Original frame and its MVF overlay of enlarged red square region
- b Part of interpolated frame by Wang’s method
- c Part of interpolated frame by Haan’s method
- d Part of interpolated frame by proposed method

Fig. 3 Result for ‘Basketball’ sequence

- a Original frame and its MVF overlay of enlarged red square region
- b Part of interpolated frame by Wang’s method
- c Part of interpolated frame by Haan’s method
- d Part of interpolated frame by proposed method