RESEARCH   **Open Access**

A resolution enhancement algorithm for an asymmetric resolution stereo video

Jing Ge[1,2], Ju Liu[1,2*], Yunlong Zhao[1] and Boyang Zhang[1]

Abstract

This paper presents a resolution enhancement algorithm in an asymmetric resolution stereo video for improving the video resolution. In this stereo video architecture, a scene is captured by two cameras to form two views, one view is a lower-resolution video and the other is a full-resolution video. The goal is to enhance the lower-resolution video to a full-resolution video. In the lower-resolution video, frames synchronized with full-resolution video are enhanced via disparity estimation algorithm, while the rest frames are improved by mono-view video super-resolution based on key frames method. The experimental results demonstrate that the proposed method is effective for both visual and objective qualities.

Keywords: Resolution enhancement; Disparity estimation; Video super-resolution

1 Introduction

Nowadays, with the development of image/video processing techniques and display techniques, high resolution and quality videos become more widespread than 20 years ago. While, due to the limited network bandwidth, especially in wireless network, raw videos should be compressed with a high compression ratio, the high-quality videos are converted to the low-quality ones. Since that, quality enhancement process for low-quality videos is needed to improve user experience. Video super-resolution is one of the promising approaches for improving visual quality. It is a process to create a high-resolution (HR) video sequence from a low-resolution (LR) video sequence, which is applying high correlation between mono-view LR frames to extract effective information for the compensation.

Video super-resolution algorithms are ubiquitously used for mono-view video enhancement [1–8]. Among the video super-resolution algorithms, one strategy is to reconstruct the HR frames in a mixed resolution framework (a mono-view mixed resolution video has both HR frames and LR frames). Brandi et al. had a super-resolution algorithm for the mixed resolution framework

[6]. He had set the HR frames as the key frames and treated the LR frames as non-key frames; he improved the resolution of non-key frames by key frames information. Motion compensation algorithm in [7, 8] employs overlapped block motion compensation (OBMC) and in [9] utilizes adaptive overlapped block motion compensation (AOBMC) to alleviate block artifacts and have better visual quality than OBMC. Ancuti et al. [10] utilized maximum *a posterior* (MAP) estimation and high-quality photographs to provide the prior information for HR frames. Najafi et al. [11] used HR frames to form a regularization function and used it for super-resolution de-blurring stage.

Tam [12] observed that, in the case of asymmetric blurring, the binocular image quality is dominated by the quality of HR view. Diogo [13] stated that the mixed resolution framework applied for multi-view video offered great data-size reduction without suffering from the significant quality degradation in stereoscopic video applications. For instance, under mobile 3D video application, the key point in design is to optimize the video bit rate to suit for the limited network bandwidth, hence the asymmetric stereoscopic video scheme is a candidate of this kind of application. Consequently, for stereoscopic video, the asymmetric stereoscopic video technique, which is utilizing a pair of lower resolution and full resolution frames to reduce the data storage requirement, is presented [14–16]. Zhang [15] stated a simultaneous method for

*Correspondence: juliu@sdu.edu.cn
[1] School of Information Science and Engineering, Shandong University, Shanda South Road, Jinan, China
[2] Suzhou Research Institute, Shandong University, Linquan Road, Suzhou, China

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 2 of 9

video super-resolution and high-quality depth estimation. The iteration process is between the result of depth estimation and the result of video super-resolution until the results of both are stable. Ankit [16] introduced spatio-temporal Markov random field model to stereo super-resolution for the mixed resolution stereo video.

The mixed resolution framework mentioned above is asymmetric resolution with the same frame rate, while in this paper, the framework is asymmetric resolution as well as asymmetric frame rate. We present a resolution enhancement algorithm for the asymmetric stereoscopic video. Our goal is to effectively exploit all information including lower resolution frames and full resolution frames for obtaining full resolution video. The pair of lower resolution and full resolution frames captured from two views at the same time are called synchronized frames, and others in the lower resolution view are termed as non-synchronized frames. This method has employed disparity estimation and disparity compensation for resolution enhancement of synchronized frames and applied video super-resolution based on key frames for resolution enhancement of non-synchronized frames.

The contributions of this paper are (1) introducing an asymmetric resolution and asymmetric frame rate stereoscopic framework, (2) applying normalized mutual information based on adaptive support weight stereo matching method to disparity estimation, and (3) providing a robust adaptive overlapped block motion compensation algorithm to mono-view video super-resolution. The rest of the paper is organized as follows. In Section 2, we introduce the asymmetric resolution framework. The resolution enhancement algorithm for the mixed architecture is presented in Section 3. We show our experimental results and conclusions in Sections 4 and 5, respectively.

2 Asymmetric resolution framework

The asymmetric resolution framework is described as follows, the left view is a lower resolution with full frame rate video; the right view is a full resolution with lower frame rate video. To get the full resolution video, the resolution enhancement algorithm of the left view for asymmetric resolution framework is achieved with two steps: synchronized frames enhancement by disparity estimation and disparity compensation and non-synchronized frames enhancement by video super-resolution. The process is illustrated in Fig. 1. The left view outputs lower resolution frames with full frame rate, the right view outputs full resolution frames with lower frame rate. The purpose of this architecture is to reconstruct full resolution frames of the left view by using all information of the two views.

The whole enhancement process is described in the following. Firstly, the left view frames I_L with lower resolution are classified into frames $I_{L,LR1}$ (gray frames) synchronized with the right view full resolution frames, and non-synchronized frames $I_{L,LR2}$ (white frames). I_R is the full resolution frames in the right view. Disparity estimation is employed between the up-sampled $I_{L,LR1}$ and I_R to obtain the disparity map. Considering the tradeoff of the up-sampling quality and the computational complexity, we select the bicubic interpolation as the up-sampling method. The disparity estimation in this method utilizes normalized mutual information
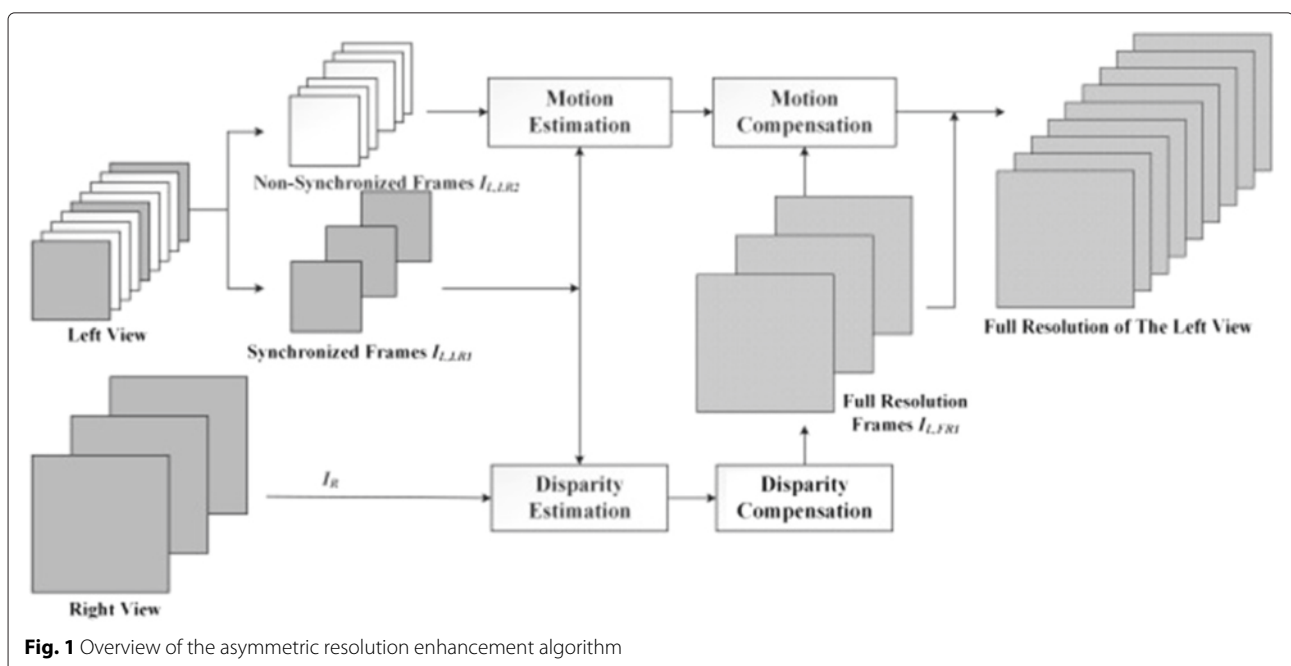

Fig. 1 Overview of the asymmetric resolution enhancement algorithm

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 3 of 9

(NMI) algorithm based on adaptive support weighting (ASW).

Then, using the disparity map, the corresponding high frequency information of $I_{L,LR1}$ is warped geometrically from the full resolution I_R. $I_{L,FR1}$, the full resolution of $I_{L,LR1}$, is consisted of two parts, one part is the high frequency information of $I_{L,LR1}$, and the other one is up-sampled low resolution of $I_{L,LR1}$.

At last, for the left view, the non-synchronized frames $I_{L,LR2}$ use the mono-view video super-resolution method to improve the resolution. In our method, the full resolution frames $I_{L,FR1}$ are regarded as key frames, and the remaining target frames apply the enhancement algorithm of AOBMC algorithm based on key frames to retrieve the lost information of $I_{L,LR2}$. Thus, $I_{L,FR2}$, the full resolution of the non-synchronized frames $I_{L,LR2}$, has been finished. The process of enhancement is elaborated in the next section in detail.

3 Resolution enhancement algorithm
3.1 Synchronized frames enhancement by disparity estimation

Because the asymmetric video has asymmetric frame rate, we apply stereo matching method for disparity estimation on a pair of synchronized full resolution and lower resolution frames. Stereo matching algorithms have been classified into global methods and local methods [17, 18]. Since the local algorithms are more efficient and simpler than global methods, the local methods are chosen as stereo matching algorithm in our method. Moreover, as stereo matching result is sensitive to radiometric variation, it is necessary to find a stable and robust matching cost. Mutual information, based on information theory, can be a good candidate [19]. Another local method, adaptive support weighting (ASW) algorithm [20] proposed by Yoon et al., was an effective binocular vision algorithm and could alleviate block artifacts effectively, which selected a fixed size and shape support window for each targeted pixel. The weighting of each pixel in this fixed support window was adaptive to different color similarity and different spatial distance compared with the targeted pixel. In our method, we have presented a mutual information matching measurement based on ASW in the stereo matching method.

For one pixel p in the left reference frame $I_{L,LR1}$, the corresponding pixel with the disparity d in the right reference frame I_R is p_d. In the asymmetric resolution framework, the left and the right reference frames $I_{L,LR1}$ and I_R are the synchronized frames of the left and right views. In the algorithm for fixed support window, normalized mutual information (NMI) C_{NMI} is treated as the matching measurement for the two symmetrical support windows and is expressed as (1), where N_p and N_{p_d} are the support windows for pixel p and pixel p_d, respectively, $H(N_p)$ and

$H(N_{p_d})$ are the entropy of these two support windows, respectively, $P_{N_p}(x)$ is the probability of arbitrary pixel x in the support window N_p, $P_{N_{p_d}}(y)$ is the probability of arbitrary pixel y in the support window N_{p_d}, $P_{N_p,N_{p_d}}(x,y)$ is the joint probability, and

$$
\begin{aligned}
C_{\text{NMI}}\left(N_p, N_{p_d}\right) &= \frac{H\left(N_p\right) + H\left(N_{p_d}\right)}{H\left(N_p, N_{p_d}\right)} \\
&= \frac{-\sum_x P_{N_p}(x) \log P_{N_p}(x) - \sum_y P_{N_{p_d}}(y) \log P_{N_{p_d}}(y)}{-\sum_x \sum_y P_{N_p,N_{p_d}}(x,y) \log P_{N_p,N_{p_d}}(x,y)}.
\end{aligned}
\tag{1}
$$

Within the allowable range of disparity $S_d = \{d_{\min}, \cdots, d_{\max}\}$, the disparity of pixel p is

$$
d(p) = \arg\min_{d \in S_d} C_{\text{NMI}}\left(N_p, N_{p_d}\right).
\tag{2}
$$

The C_{NMI} is the matching cost for the support window, while the weighting using in ASW is related with the pixel pair. So that, the NMI measurement above has changed to suit the pixel pair. A measurement C_{nmi}, based on the pixel-wise NMI measurement, is proposed and formulated as (3).

$$
C_{\text{nmi}}(q, q_d) = \frac{h(q) + h(q_d)}{h(q, q_d)} = \frac{-P(q) \log P(q) - P(q_d) \log P(q_d)}{-P(q, q_d) \log P(q, q_d)}
\tag{3}
$$

Here, $P(q)$ is the probability of pixel q in the support window N_p of the pixel p in the left reference frame; $P(q_d)$ is the probability of pixel q_d in the support window N_{p_d} of the pixel p_d in the right reference frame. $P(q, q_d)$ is the joint probability of pixel q and pixel q_d. $h(q, q_d) = -P(q, q_d) \log P(q, q_d)$ is the joint entropy of pixel q and q_d. As the left support window is fixed, $P(q)$ is a constant. The value of $P(q_d)$ is small changes with d by-pixel shift, such that $P(q_d)$ is a constant approximately. In other words, the greater the value of the entropy probability of two corresponding pixels is, the less value of C_{nmi} is.

Fig. 2 First frame of the left view from each video. From left to right, the first row is *Book_Arrival* and *Bullinger*; the second row is *Car* and *Stone*

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 4 of 9

Table 1 Average and maximum PSNR (dB) comparisons of the synchronized frames

Test video	MRF [16]		ASW [20]		SS-AW [23]		Proposed method	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
Book	31.45	32.22	30.98	31.35	31.17	31.61	32.21	34.47
Bullinger	36.28	37.49	35.33	37.22	35.64	37.34	35.73	37.22
Car	35.01	35.31	36.03	36.37	36.01	36.42	36.22	36.64
Stone	28.93	29.07	29.80	29.98	29.84	29.97	30.09	30.26

Combining the ASW method and the pixel-wise NMI algorithm, the dissimilarity function is formulated as (4), where the C_{nmi} is the matching cost function. It has added the weighting matrices as filter to alleviate artifacts rising from the matching error. The weights in the weighting matric are according to the difference of the two pixels in the same position within two symmetrical support windows. $w(p,q)$ and $w(p_d,q_d)$ are weighting matrices and are taken as [20] mentioned.

$$C_{\text{ASW+nmi}}(p,p_d) = \frac{\sum_{q \in N_p, q_d \in N_{p_d}} w(p,q) \cdot w(p_d,q_d) \cdot C_{\text{nmi}}(q,q_d)}{\sum_{q \in N_p, q_d \in N_{p_d}} w(p,q) \cdot w(p_d,q_d)}$$

(4)

The allowable range of disparity $S_d = \{d_{\min}, \cdots, d_{\max}\}$, the disparity of pixel p is

$$d(p) = \arg\min_{d \in S_d} C_{\text{ASW+nmi}}(p,p_d).$$

(5)

With the disparity map, the lower resolution frame $I_{\text{L,LR1}}$ is warped geometrically by the corresponding full resolution frame I_{R}. The value of one pixel at coordinate (i, j) in the full resolution frame is given in (6), where $d(i,j)$ is the disparity of pixel p at coordinate (i, j), $I_{\text{L,FR1}}$ is the full resolution frame,

$$I_{\text{L,FR1}}(i,j) = I_{\text{R}}\left(i + d(i,j), j\right).$$

(6)

3.2 Non-synchronized frame enhancement by video super-resolution based on key frames

In this section, the full resolution frames, $I_{\text{L,FR1}}$, achieved by disparity estimation in the left view and illustrated in Section 3.1, are treated as key frames as well as reference frames. The video super-resolution we employed is the

method based on adaptive overlapped block motion compensation (AOBMC) which has been detailed described in [21], and in this paper, we state the method briefly.

First of all, the motion vector field is obtained by motion estimation. The motion estimation is employing between the full resolution frames $I_{\text{L,FR1}}$ and the up-sampled low resolution frames $I_{\text{L,LR2}}$ in the left view.

Then, we need to distinguish the accuracy of each motion vector, two criteria have been presented. One criterion is the angle difference between the angle of current block and that of its four neighbor blocks. The formulators (7) and (8) give the difference of the angles.

$$\bar{\theta} = \left| \theta(\text{MV}_T) - \frac{1}{4}\sum_{n=1}^{4} \theta(\text{MV}_{F_n}) \right|$$

(7)

$$\text{Angle}\left(\text{MV}_T, \text{MV}_{\overline{F}}\right) = \begin{cases} \bar{\theta}, & \text{if } 0 < \bar{\theta} < \pi \\ 2\pi - \bar{\theta}, & \text{if } \pi \leq \bar{\theta} \leq 2\pi \end{cases}$$

(8)

In the formulators, $\theta(\cdot)$ represents a vector direction whose value ranges from 0 to 2π. MV_T and $\text{MV}_{\overline{F}}$ represent the motion vector of the current block and the mean motion vector of the neighbor (top, bottom, left, and right) blocks surrounding T, respectively. Another measurement is the normalized sum of absolute difference $(\text{SAD}_{\text{norm}})$. The SAD_{norm} is defined as follows.

$$\text{SAD}_{\text{norm}}(T,R) = \frac{\sum_{i,j} \left| I_{\text{L,LR2}}(T(i,j)) - I_{\text{L,FR1}}(R(i,j)) \right|}{\sum_{i,j} I_{\text{L,LR2}}(T(i,j))}$$

(9)

T is the current block of the current frame $I_{\text{L,LR2}}$, and R is the matching block in the adjacent full resolution frame $I_{\text{L,FR1}}$. The value of SAD_{norm} is maintained in the range from 0 to 1 and would not occur big change if the difference of the two blocks is great. The reliability of one

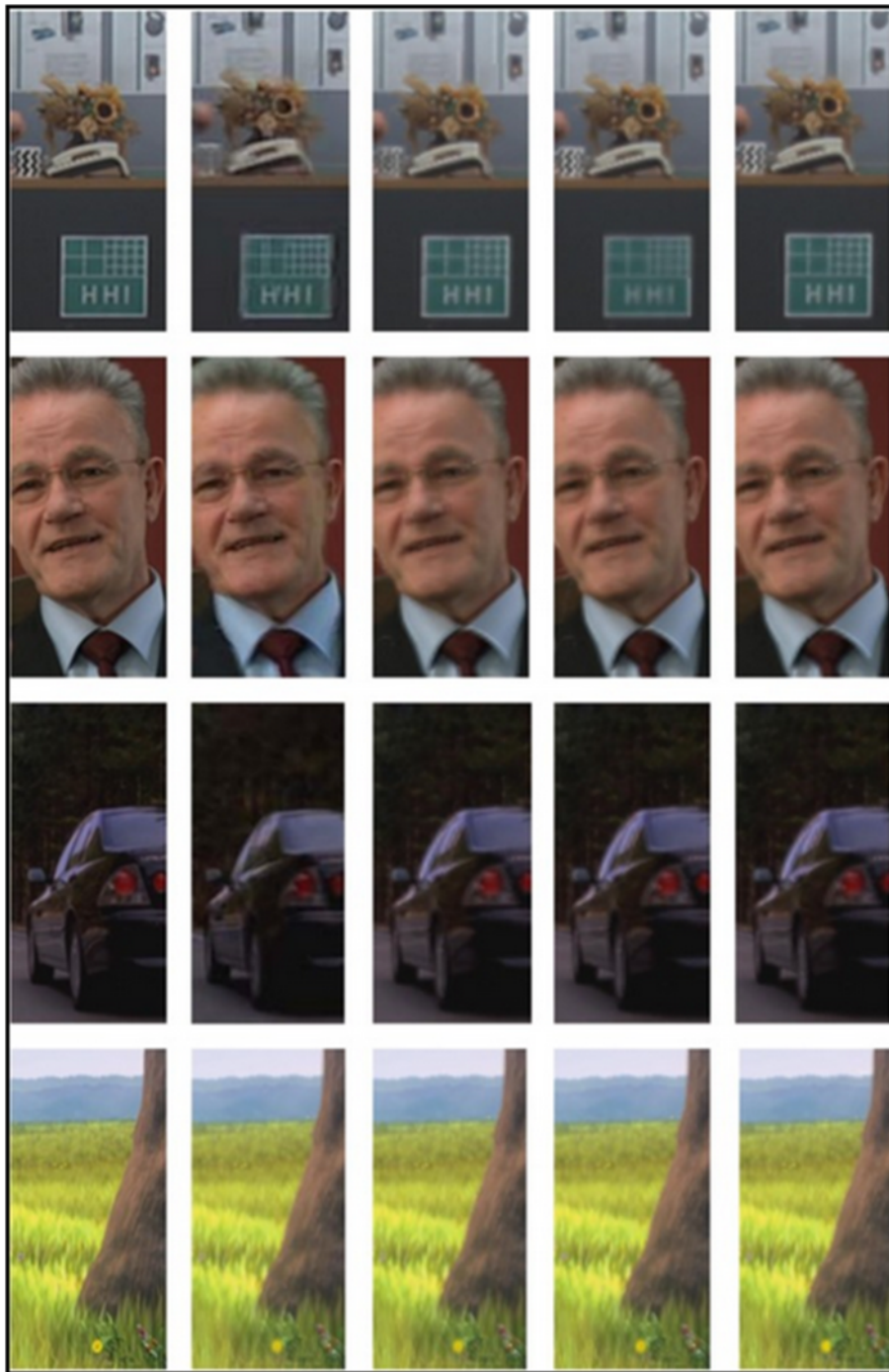**Table 2** Average and maximum SSIM comparisons of the synchronized frames

Test video	MRF [16]		ASW [20]		SS-AW [23]		Proposed method	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
Book	0.928	0.936	0.931	0.934	0.932	0.936	0.947	0.949
Bullinger	0.948	0.954	0.948	0.958	0.950	0.960	0.950	0.960
Car	0.919	0.923	0.944	0.950	0.944	0.950	0.946	0.952
Stone	0.764	0.767	0.822	0.824	0.822	0.823	0.830	0.832

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 5 of 9

motion vector is distinguished by the value of *Angle* and SAD_{norm}.

Moreover, motion vectors with the worst reliability have been corrected by vector median filter. Hence, with the different reliability of each motion vector, different weighting matric are employed according to the reliability for compensation. The compensated value of pixel in one block is

$$T_{\text{high}}(i,j) = \frac{\sum_{n=1}^{5} w_n(i,j) \cdot R_n(i,j)}{\sum_{n=1}^{5} w_n(i,j)}. \tag{10}$$

Fig. 3 Zoomed portion in the left view. From left to right, resolution enhancement results of different methods (ground truth, method in [16], in [20], in [23], and the proposed method). From up to down test videos (*Book_Arrival*, *Bullinger*, *Car*, *Stone*)

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 6 of 9

Table 3 Average and maximum PSNR (dB) comparisons of the non-synchronized frames

Test video	Bicubic		OBMC [7]		Proposed method	
	Mean	Max	Mean	Max	Mean	Max
Book	30.30	30.57	31.30	31.91	32.01	32.86
Bullinger	36.62	37.88	36.63	38.04	36.94	38.44
Car	35.75	35.98	36.15	36.50	36.60	36.97
Stone	30.71	30.93	30.92	31.11	31.10	31.26

where $R_n(n = 1, 2, 3, 4, 5)$ represent the matching block of the current block and that of the top, bottom, left, and right blocks of the current block in the adjacent frame $I_{L,FR1}$; $R_n(i,j)$ is the value of the high frequency information of the matching block; w_n are weighting matric and the weightings are taken as [21].

At last, the pixel value in the full resolution frame $I_{L,FR2}$ is achieved by adding the high-frequency information $T_{high}(i,j)$ into the up-sampled $I_{L,LR2}(i,j)$, which is described in the following formula,

$$I_{L,FR2}(i,j) = T_{high}(i,j) + I_{L,LR2}(i,j). \tag{11}$$

From the four stages mentioned above, the enhancement of the non-synchronized frames has been completed, and the left view has been improved to the full resolution version.

4 Experimental results and analysis

In order to verify the performance of the resolution enhancement algorithm proposed for asymmetric resolution stereo video, we used four stereo videos for testing, namely *Book_Arrival*, *Bullinger*, *Car*, and *Stone* [22], with resolution 432×240, as shown in Fig. 2. Experimental environment is dual Inter Xeon E5420 4 core CPU at 2.0 Ghz and 8 Gb RAM, running under Microsoft Windows 7 $\times$64 SP1 operating system. The left view is twice spatial down-sampled of the original left view; and the right view is the odd frames extracted from the original right view.

The computational complexity is according to the whole resolution enhancement algorithm. In the stage of synchronized frames enhancement, disparity estimation is pixel-based; hence, the computational complexity of this stage is proportional to the size of resolution basically. In the stage of non-synchronized frames enhancement, each frame has been divided into blocks with same size;

each block has the same three steps, namely, motion estimation, motion vector refinement, and motion compensation. With the same block size, the computational complexity of this stage is also in proportion to the resolution size. On the whole, the computational complexity is in proportion to resolution.

Here, we test the resolution enhancement results of the left view with two parts, synchronized frames resolution enhancement by NMI method based on ASW and non-synchronized frames resolution enhancement based on AOBMC algorithm. The following two subsections, Section 4.1 and Section 4.2, have evaluate the proposed method in detail.

4.1 Evaluation of synchronized frames resolution enhancement

The proposed method has compared with the ASW [20], the segmentation-based adaptive weight method (SS-AW) [23], and the spatio-temporal Markov random field (MRF) method [16]. We employ these methods frame-by-frame to each test video and compute the mean and the maximum objective measurements of every frame. The objective peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [24], and the visual effect of the synchronized frames are shown in Tables 1 and 2 and Fig. 3, respectively. For visual results, Fig. 3 shows the zoom-in portions of one frame of each test video.

Tables 1 and 2 show the mean and maximum values of PSNR and SSIM. The proposed method has the highest PSNR and SSIM values in all test videos, except for the test video *Bullinger*. For the *Bullinger*, the mean value of the proposed method is 0.5 dB lower than that of MRF, and the maximum value is only 0.22 dB lower than that of MRF; but the proposed method has a good visual appearance.

Table 4 Average and maximum SSIM comparisons of the synchronized frames

Test video	Bicubic		OBMC [7]		Proposed method	
	Mean	Max	Mean	Max	Mean	Max
Book	0.893	0.901	0.910	0.917	0.929	0.940
Bullinger	0.954	0.958	0.952	0.957	0.954	0.958
Car	0.925	0.927	0.931	0.934	0.939	0.942
Stone	0.830	0.832	0.837	0.838	0.847	0.849

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 7 of 9

The visual results show that, compared with the three methods, the proposed method in object edge has more accurate results, like stripe on the cup and green square in *Book*, and car window edge in *Car*. As the method uses the stability of NMI and deblocking of ASW method, the results demonstrate a good visual and objective quality.
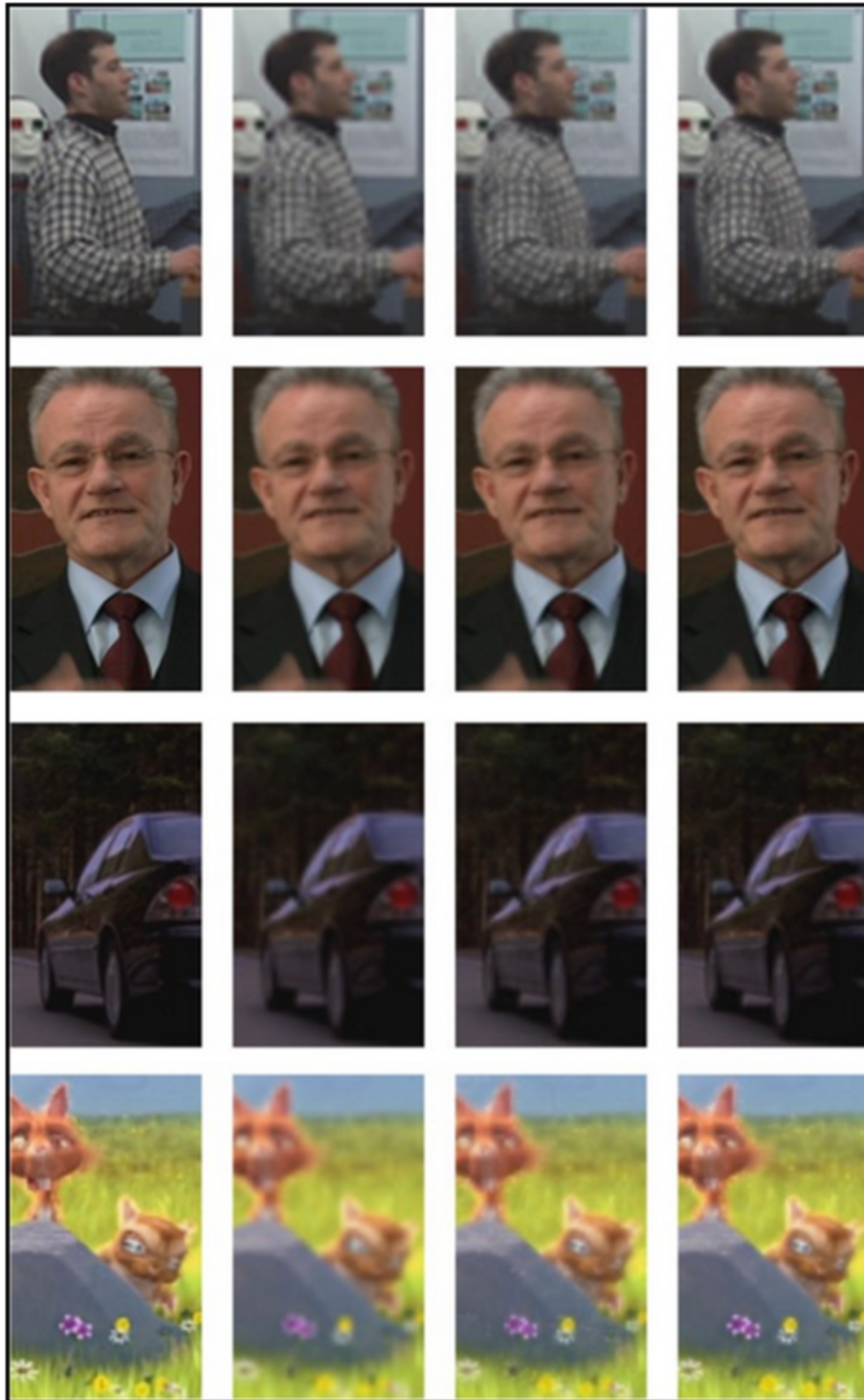
Fig. 4 Zoomed portion in the left view. From left to right, resolution enhancement results of different methods (ground truth, bicubic, method in [7], and the proposed method). From up to down test videos (*Book_Arrival*, *Bullinger*, *Car*, *Stone*)

Ge *et al. EURASIP Journal on Image and Video Processing* (2015) 2015:23

Page 8 of 9

4.2 Evaluation of non-synchronized frames resolution enhancement

In this section, the full resolution of synchronized frames in the left view, enhanced by NMI method based on ASW, are used as key frames for the non-synchronized frames of the left view. Comparison is between the proposed method and OBMC in [7]. With the same synchronized frames results, different motion compensation algorithms are taken and results have presented in Tables 3 and 4 and Fig. 4, respectively. The motion estimation algorithm is UMHexgonS (Unsymmetrical-cross Multi-Hexagon-grid Search) algorithm, and the block size we take is 16×16.

Tables 3 and 4 show the mean and maximum value of PSNR and SSIM of the non-synchronized frames. The proposed method has the highest PSNR value as well as the highest SSIM score. In the comparison of PSNR, the mean value of the proposed method is about 0.61 dB higher than that of the other three. In the comparison of SSIM, the mean value of the proposed method is about 0.013 higher than that of the other three.

From Fig. 4, we could conclude that result combined with NMI based on ASW method and AOBMC algorithm is the best one. AOBMC algorithm is the best one since it applies different weighting matrics and motion vector classification compared with the other two. Consequently, video super-resolution based on AOBMC algorithm can be used both in mono-view and asymmetric stereoscopic video effectively, and when using in the asymmetric stereo video, the high-resolution frames should be achieved first.

5 Conclusions

In this paper, resolution enhancement algorithm in the asymmetric resolution stereo video to achieve full resolution stereo video is studied. The quality of the lower resolution view in the asymmetric framework has been improved via the disparity estimation and the motion compensation algorithm. The disparity estimation employs normalized mutual information based on adaptive support weighting; the motion compensation utilizes adaptive overlapped block motion compensation. Experimental results show that our method utilizing disparity compensation and motion compensation could give both good visual and good objective qualities.

References
1. Lee D-b, B-Y Heo, BC Song, in *Image Processing (ICIP), 2013 20th IEEE International Conference On*. Video deblurring based on bidirectional motion compensation and accurate blur kernel estimation (IEEE Melbourne, VIC, 2013), pp. 895–899. doi:10.1109/ICIP.2013.6738185
2. SC Park, MK Park, MG Kang, Super-resolution image reconstruction: a technical overview. Signal Process. Mag. IEEE. **20**(3), 21–36 (2003). doi:10.1109/MSP.2003.1203207
3. EM Hung, RL de Queiroz, F Brandi, KF de Oliveira, D Mukherjee, Video super-resolution using codebooks derived from key-frames. Circ. Syst. Video Technol. IEEE Trans. **22**(9), 1321–1331 (2012). doi:10.1109/TCSVT.2012.2201669
4. M Shen, P Xue, C Wang, Down-sampling based video coding using super-resolution technique. Circ. Syst. Video Technol. IEEE Trans. **21**(6), 755–765 (2011). doi:10.1109/TCSVT.2011.2130390
5. MM Islam, VK Asari, MN Islam, MA Karim, Super-resolution enhancement technique for low resolution video. Consum. Electron. IEEE Trans. **56**(2), 919–924 (2010). doi:10.1109/TCE.2010.5506020
6. F Brandi, R de Queiroz, D Mukherjee, in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference On*. Super-resolution of video using key frames and motion estimation (IEEE San Diego, CA, 2008), pp. 321–324. doi:10.1109/ICIP.2008.4711756
7. BC Song, S-C Jeong, Y Choi, in *Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference On*. Key frame-based video super-resolution using bi-directional overlapped block motion compensation and trained dictionary (IEEE Paris, 2010), pp. 181–186. doi:10.1109/IPTA.2010.5586726
8. MT Orchard, GJ Sullivan, Overlapped block motion compensation: an estimation-theoretic approach. Image Process. IEEE Trans. **3**(5), 693–699 (1994). doi:10.1109/83.334974
9. J Ge, B Zhang, J Liu, F Wang, F Feng, in *Intelligent Control and Automation (WCICA), 2012 10th World Congress On*. Key frames-based video super-resolution using adaptive overlapped block motion compensation (IEEE, 2012), pp. 4712–4716
10. C Ancuti, CO Ancuti, P Bekaert, in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference On*. Video super-resolution using high quality photographs (IEEE Dallas, TX, 2010), pp. 862–865. doi:10.1109/ICASSP.2010.5495223
11. S Najafi, S Shirani, in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference On*. Regularization function for video super-resolution using auxiliary high resolution still images (IEEE Pacific Grove, CA, 2012), pp. 1713–1717. doi:10.1109/ACSSC.2012.6489325
12. WJ Tam, *Image and depth quality of asymmetrically coded stereoscopic video for 3D-TV*, (JVT-W094, San Jose, CA, 2007)
13. DC Garcia, C Dorea, RL de Queiroz, Super resolution for multiview images using depth information. Circ. Syst. Video Technol. IEEE Trans. **22**(9), 1249–1256 (2012). doi:10.1109/TCSVT.2012.2198134
14. J Tian, L Chen, Z Liu, in *Pattern Recognition (ACPR), 2011 First Asian Conference On*. Asymmetric stereoscopic image resolution enhancement (IEEE Beijing, 2011), pp. 303–306. doi:10.1109/ACPR.2011.6166679
15. J Zhang, Y Cao, Z Wang, in *Image Processing (ICIP), 2013 20th IEEE International Conference On*. A simultaneous method for 3d video super-resolution and high-quality depth estimation (IEEE Melbourne, VIC, 2013), pp. 1346–1350. doi:10.1109/ICIP.2013.6738277
16. AK Jain, TQ Nguyen, in *Image Processing (ICIP), 2013 20th IEEE International Conference On*. Video super-resolution for mixed resolution stereo (IEEE Melbourne, VIC, 2013), pp. 962–966. doi:10.1109/ICIP.2013.6738199
17. D Scharstein, R Szeliski, R Zabih, in *Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings. IEEE Workshop On*. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, (2001), pp. 131–140. doi:10.1109/SMBV.2001.988771
18. R Szeliski, *Computer Vision: Algorithms and Applications*. (Springer, London, 2010)
19. YS Heo, KM Lee, SU Lee, in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*. Mutual information-based stereo matching combined with SIFT descriptor in log-chromaticity color space, (2009), pp. 445–452. doi:10.1109/CVPR.2009.5206507
20. K-J Yoon, IS Kweon, Adaptive support-weight approach for correspondence search. Pattern Anal. Mach. Intell. IEEE Trans. **28**(4), 650–656 (2006). doi:10.1109/TPAMI.2006.70

Ge *et al. EURASIP Journal on Image and Video Processing*   (2015) 2015:23

Page 9 of 9

21. J Ge, J Liu, C Ge, X Yang, in *Signal-Image Technology Internet-Based Systems (SITIS), 2013 International Conference On*. A robust video super-resolution based on adaptive overlapped block motion compensation, (2013), pp. 187–194. doi:10.1109/SITIS.2013.41

22. 3D video database– Stereo–video. MOBILE3DTV stereo-video. http://sp.cs.tut.fi/mobile3dtv/stereo-video/, accessed date 16 March 2011

23. F Tombari, S Mattoccia, L Di Stefano, in *Advances in Image and Video Technology*. Segmentation-based adaptive support for accurate stereo correspondence, vol. 4872, (2007), pp. 427–438. doi:10.1007/978-3-540-77129-6_38

24. Z Wang, AC Bovik, HR Sheikh, EP Simoncelli, Image quality assessment: from error visibility to structural similarity. Image Process. IEEE Trans. **13**(4), 600–612 (2004). doi:10.1109/TIP.2003.819861