Multi-Image Super-Resolution for Fisheye Video Sequences Using Subpixel Motion Estimation Based on Calibrated Re-Projection

Michel Bätz, Andrea Eichenseer, and André Kaup
Multimedia Communications and Signal Processing
Friedrich-Alexander University Erlangen-Nürnberg (FAU), Cauerstr. 7, 91058 Erlangen, Germany
Email: {michel.baetz, andrea.eichenseer, andre.kaup}@fau.de

***Abstract*—Super-resolution techniques are a means for reconstructing a higher spatial resolution from low resolution content, which is especially important for automotive or surveillance systems. Furthermore, being able to capture a large area with a single camera can be realized by using ultra-wide angle lenses, as employed in so-called fisheye cameras. However, the underlying non-perspective projection function of fisheye cameras introduces significant radial distortion, which is not considered by conventional super-resolution techniques. In this paper, we therefore propose the integration of a fisheye-adapted motion estimation approach that is based on a calibrated re-projection into a multi-image super-resolution framework. The proposed method is capable of taking the fisheye characteristics into account, thus improving the reconstruction quality. Simulation results show an average gain in luminance PSNR of up to 0.3 dB for upscaling factors of 2 and 4. Visual examples substantiate the objective results.**

I. Introduction

Increasing the spatial resolution is desirable in many applications such as automotive or video surveillance systems where it is very important to be able to detect faces, license plates, or street signs. The utilization of camera sensors with an increased resolution is one way to obtain images with a higher resolution. These sensors, however, are more expensive, require a larger amount of memory for storing the images, and consume more energy which is especially critical in remote or mobile systems. As an alternative, so-called super-resolution (SR) approaches can be employed to obtain a high resolution (HR) image from a low resolution (LR) input. SR methods can generally be divided into two groups: single-image SR (SISR) using only a single input image and multi-image SR (MISR) exploiting the temporal correlation of several adjacent frames of a video sequence. In principle, MISR approaches are capable of recovering more detail due to the incorporation of additional information from adjacent frames. However, the reconstruction quality of MISR methods highly depends on an accurate motion estimation. Moreover, it is necessary that a subpixel motion is present in the sequence for MISR methods to perform well at all [1].

In the literature, many SISR approaches exist that range from simple edge-directed interpolation [2] to more complex example-based methods trying to establish a relationship between HR and LR patch pairs [3], [4], [5], [6]. Likewise, a considerable number of MISR methods can be found in the literature [1]. These techniques range from spatial and frequency domain approaches to arbitrarily complex deterministic or stochastic regularization approaches [7], [8]. Additionally, some methods do not need an explicit motion estimation at all [9] or combine both SISR and MISR approaches to form a hybrid solution [10].

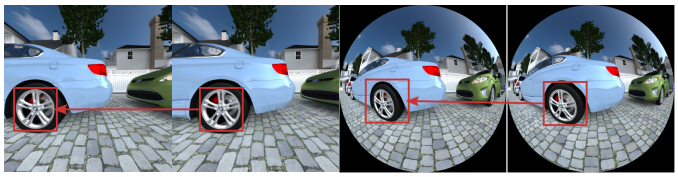

Figure 1. Translational motion as captured by both a conventional perspective camera (left half) and by a fisheye camera (right half).

The second important aspect in video surveillance or automotive systems is the desire of being able to capture the largest possible area with a single camera. Typically, this is realized by employing so-called fisheye cameras which make use of ultra-wide angle lenses that are capable of capturing a field of view (FOV) of 180 degrees or even more [11]. Several processing strategies that make use of fisheye cameras have been presented in the literature [12], [13], especially for driver's assistance systems or surveillance cameras. However, fisheye cameras introduce a large radial distortion into the images as the underlying projection function is not perspective. This underlying mapping is typically not considered for processing techniques such as motion estimation (ME). Fig. 1 illustrates how a translational motion is captured by a conventional camera (left half) and by a fisheye camera (right half). It is evident that a conventional ME based on block-matching is not capable of finding good candidate blocks as radial distortion is not taken into account. In [14], a hybrid fisheye-adapted ME technique is presented that uses the equisolid projection model to improve the ME itself. However, a real fisheye camera does not necessarily follow a perfectly equisolid model, making a calibration necessary as elaborately analyzed in [15]. In this paper, we therefore propose the integration of that calibrated hybrid ME approach into a multi-image super-resolution framework for fisheye video sequences and extend the approach to also consider subpixel shifts. The proposed SR method is capable of taking the characteristics of fisheye images into account and hence improves the reconstruction

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

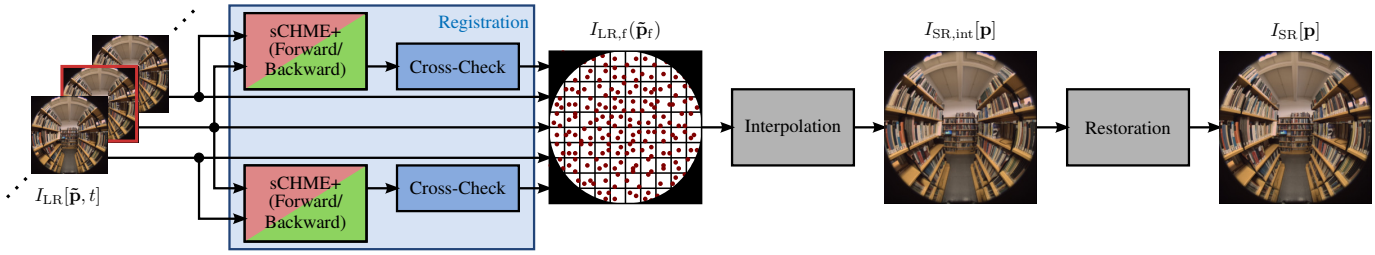

Figure 2. Proposed multi-image super-resolution framework using a subpixel-accurate hybrid motion estimation technique (sCHME+, red/green) based on calibrated re-projection. The workflow can be divided into 3 major parts: registration (blue), interpolation, and restoration.

quality. Moreover, we also analyze the effects of a varying number of utilized input frames on the reconstruction quality.

The remainder of the paper is structured as follows. In Section II, the employed MISR approach is briefly described. The proposed subpixel-accurate fisheye ME technique applied for the registration is given in Section III. Section IV presents the simulation results and Section V concludes this paper.

II. Multi-Image Super-Resolution Approach

As mentioned in the previous section, multi-image SR methods try to reconstruct an HR image from several LR input images. The observation model [1] relating the desired HR image $\mathbf{I}_{\mathrm{HR}}$ to the corresponding k-th observed LR video frame $\mathrm{I}_{\mathrm{LR},k}$ can basically be formulated as

$$\mathbf{I}_{\mathrm{LR},k} = \mathbf{DBM}_k\mathbf{I}_{\mathrm{HR}} + \mathbf{n}_k, \quad \text{for } 1 \leq k \leq M. \qquad (1)$$

The matrices $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{M}_k$ denote the downsampling, blurring, and motion between the frames, respectively. The vector $\mathbf{n}_k$ represents the noise and M expresses the number of LR images.

To obtain a super-resolved image, the non-uniform interpolation approach mentioned in [1] is a straightforward solution. The basic workflow of such an interpolation-based SR technique is comprised of three major steps: registration, interpolation, and restoration. These steps can be seen in Fig. 2 where the registration is highlighted in blue.

In the registration step, a motion estimation between all utilized LR frames $I_{\mathrm{LR}}[\tilde{\mathbf{p}}, t]$ needs to be performed. The LR pixel position is denoted by $\tilde{\mathbf{p}} = (\tilde{m}, \tilde{n})$, the point in time is given by t. In this paper, the employed auxiliary LR frames are symmetrically chosen around the LR frame to be super-resolved: $t \in \mathcal{T} = \{\tau - \lceil (N-1)/2 \rceil, \ldots, \tau + \lfloor (N-1)/2 \rfloor\}$. The number of utilized frames including the central frame at $t = \tau$ that is to be super-resolved is represented by N. All pixels are subsequently compensated into a common floating point accurate mesh $I_{\mathrm{LR,f}}(\tilde{\mathbf{p}}_{\mathrm{f}})$ with $\tilde{\mathbf{p}}_{\mathrm{f}}$ being the floating point accurate positions corresponding to the target LR frame. For the registration of real-world video sequences containing both global and local motion, a multitude of approaches can be applied. We chose a block-based motion estimation method [16] as a basis for further adaptations towards the underlying fisheye characteristics. Moreover, to increase the robustness of our approach against inaccurate motion vectors (MV), both a forward and backward ME is computed between the LR frames. This procedure then allows for a cross-check between the MVs as mentioned in [10] and thus successfully removes coarse outliers in the motion vector field (MVF).

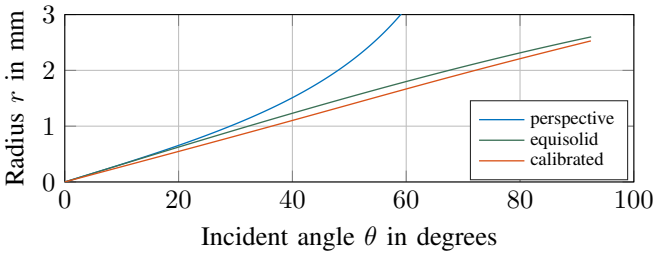

Figure 3. Projection functions of radius r over incident angle θ.

In the interpolation step, the non-uniformly sampled pixels are resampled onto the HR image grid using the well-known Delaunay triangulation and a subsequent cubic interpolation. The result is the interpolated image $I_{\mathrm{SR,int}}[\mathbf{p}]$ with $\mathbf{p} = (m, n)$ denoting the HR pixel positions. Finally, as a means of image restoration, Lucy-Richardson deconvolution [17] is used.

As a subpixel-accurate traditional ME technique (sTME) based on block matching does not consider the fisheye characteristics, it is likely to produce reconstruction artifacts due to incorrect motion vectors. We therefore propose to apply a subpixel-accurate and calibrated hybrid ME technique (sCHME+) for fisheye video sequences to the registration step in MISR. The details of the proposed sCHME+ approach are presented in the next section.

III. Calibrated Hybrid Motion Estimation

Building upon [14] where a hybrid fisheye motion estimation technique is proposed that combines traditional block-based ME with a fisheye-adapted ME based on the equisolid fisheye model [11], we also make use of such a hybrid approach. However, instead of using a perfect fisheye model which does not necessarily hold true for real-world fisheye lenses, a calibration is applied to find the mapping between the incident angle θ and the radius r as also shown in [15]. For the calibration, the OCamCalib toolbox by Scaramuzza et al. [18] is used, yielding the polynomial p that relates the radius in the fisheye domain r_{fe} to the corresponding incident angle θ. Mathematically, this polynomial can be written as

$$p(\theta) = a_n\theta^n + a_{n-1}\theta^{n-1} + \ldots + a_1\theta + a_0 = \sum_{i=0}^{n} a_i\theta^i, \quad (2)$$

where a_i denote the coefficients of the polynomial and n represents its order. Fig. 3 illustrates the relationship between

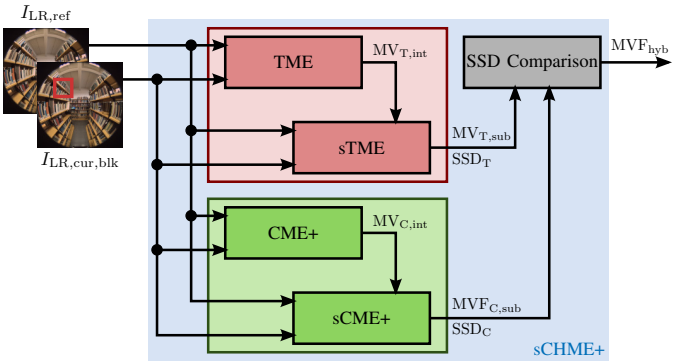

Figure 4. Proposed sCHME+ approach in detail. The upper box (red) depicts a traditional hierarchical ME, while the lower box (green) illustrates the novel calibrated hierarchical ME part. Combining both results in a hybrid solution.

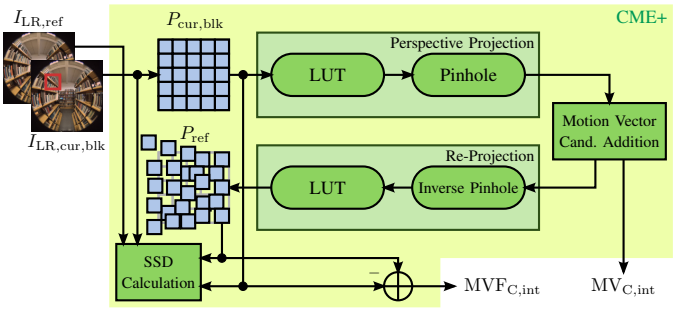

Figure 5. Workflow of the calibrated fisheye approach CME+. P refers to the pixel positions and LUT is a lookup table relating the radius r to the incident angle θ. sCME+ works analogously yielding $\mathrm{MV}_{\mathrm{C,sub}}$ and $\mathrm{MVF}_{\mathrm{C,sub}}$.

radius r and incident angle θ for the perspective, equisolid, and calibrated case. From these curves, it is evident that a perspective lens can only capture a limited FOV.

Additionally, due to the need for subpixel-accurate MVs in multi-image super-resolution, we extend the hybrid motion estimation scheme by a hierarchical subpixel step. The basic workflow of the subpixel-accurate calibrated hybrid motion estimation (sCHME+) is depicted in Fig. 4. The red box consists of a traditional motion estimation (TME) based on block-matching followed by a subpixel-accurate extension (sTME) around the best integer motion vector $\mathrm{MV}_{\mathrm{T,int}}$ to obtain a subpixel-accurate motion vector $\mathrm{MV}_{\mathrm{T,sub}}$. The green box contains the calibrated fisheye-adapted ME (CME+) followed by its corresponding hierarchical subpixel extension (sCME+) around the best integer MV candidate $\mathrm{MV}_{\mathrm{C,int}}$, which then yields the pixelwise MV field $\mathrm{MVF}_{\mathrm{C,sub}}$. The sum of squared differences (SSD) between the current block and the motion compensated blocks obtained by applying both $\mathrm{MV}_{\mathrm{T,sub}}$ and $\mathrm{MVF}_{\mathrm{C,sub}}$ are then compared and the best hybrid solution $\mathrm{MVF}_{\mathrm{hyb}}$ is found by selecting the path corresponding to the smaller SSD value. If $\mathrm{MV}_{\mathrm{T,sub}}$ is chosen, it is copied to each pixel position of the block to create a pixelwise MV field.

In the following, the actual CME+ approach, which represents the core of the ME approach, is explained in detail. The principle workflow for evaluating one MV candidate is illustrated in Fig. 5. First, the pixel positions of the current block $P_{\mathrm{cur,blk}}$ are extracted for further processing. These positions are subsequently converted from Cartesian coordinates to polar coordinates, providing the radius in the fisheye domain r_{fe} of each pixel. From the polynomial p obtained from the calibration, a lookup table (LUT) is created which maps r_{fe} to the corresponding incident angle θ. The perspective radius r_{pp} can then be calculated according to the pinhole model as

$$r_{\mathrm{pp}} = f\tan(\theta), \qquad \text{with } \theta = p^{-1}(r_{\mathrm{fe}}), \tag{3}$$

where f denotes the focal length and θ is obtained from the LUT. In the perspective domain, the actual motion vector candidate addition is conducted as usual on the Cartesian coordinates corresponding to the polar coordinates $(r_{\mathrm{pp}}, \varphi)$. To perform the SSD comparison, the shifted positions have to be re-projected into the fisheye domain as in [14]. To that end, the shifted perspective radius r'_{pp} has to be mapped to the incident angle θ by use of the inverse pinhole model followed by the mapping of the LUT to obtain r'_{fe}. This re-projection can be written as

$$r'_{\mathrm{fe}} = p(\theta), \qquad \text{with } \theta = \arctan\left(\frac{r'_{\mathrm{pp}}}{f}\right). \tag{4}$$

Subsequently, the actual SSD comparison can be performed using the pixel values of $I_{\mathrm{LR,ref}}$ at the shifted positions P_{ref} and the pixel values of $I_{\mathrm{LR,cur,blk}}$ at the original positions $P_{\mathrm{cur,blk}}$. The motion vector candidate is stored in $\mathrm{MV}_{\mathrm{C,int}}$. The actual shifts, however, are different for each pixel due to the nature of the used projections and can be obtained by simply subtracting the positions $P_{\mathrm{cur,blk}}$ from P_{ref} to obtain the MV field $\mathrm{MVF}_{\mathrm{C,int}}$. This procedure has to be repeated until the best MV is found. Analogously, sCME+ yields the subpixel-accurate $\mathrm{MV}_{\mathrm{C,sub}}$ and $\mathrm{MVF}_{\mathrm{C,sub}}$. The pixelwise shifts $\mathrm{MVF}_{\mathrm{C,sub}}$ are then further used for the super-resolution. Note that the whole projection and re-projection step is only performed on the pixel positions themselves and hence does not require any interpolation of the luminance values.

Due to the nature of the tangent, however, radii corresponding to an angle larger than 90 degrees cause a problem, i. e., generate a negative radius r_{pp}. We therefore propose a compensation for ultra-wide angles that requires three steps. The first necessary modification needs to be conducted during the motion vector candidate addition. For all positions that correspond to an angle larger than 90 degrees, the shifts need to be multiplied by -1. Mathematically, this is represented by

$$\mathbf{m}_{\mathrm{cand}} = \begin{cases} (-\Delta x_{\mathrm{pp}}, -\Delta y_{\mathrm{pp}}), & \forall r_{\mathrm{pp}} < 0 \\ (\Delta x_{\mathrm{pp}}, \Delta y_{\mathrm{pp}}), & \text{otherwise} \end{cases}, \tag{5}$$

where $(\Delta x_{\mathrm{pp}}, \Delta y_{\mathrm{pp}})$ denotes the Cartesian representation of the MV candidate $\mathbf{m}_{\mathrm{cand}}$. As a second modification for the ultra-wide angle compensation, the angle of the shifted polar coordinates φ' has to be mirrored according to

$$\hat{\varphi}' = \varphi' - \pi, \qquad \forall r_{\mathrm{pp}} < 0, \tag{6}$$

where $\hat{\varphi}'$ represents the manipulated polar angle. Finally, for all re-projected radii r'_{fe} corresponding to $\theta > 90°$, we add twice the distance to the radius corresponding to 90°, namely $r_{\mathrm{fe},90}$, as

$$\hat{r}'_{\mathrm{fe}} = r'_{\mathrm{fe}} + 2(r_{\mathrm{fe},90} - r'_{\mathrm{fe}}), \qquad \forall r_{\mathrm{pp}} < 0, \tag{7}$$

Table I
AVERAGE $\mathrm{PSNR_Y}$ RESULTS (IN DB) FOR BOTH N=7 AND N=15 AND A RESOLUTION ENHANCEMENT OF U=2 AND U=4.

Sequence	Motion Type	Frames	2× Upscaling					4× Upscaling				
			Bicubic	SR using sTME		SR using sCHME+		Bicubic	SR using sTME		SR using sCHME+	
			N=1	N=7	N=15	N=7	N=15	N=1	N=7	N=15	N=7	N=15
LibraryA	Camera Zoom	#11-40	34.27	35.42	35.85	35.52	**36.03**	28.83	29.08	29.56	29.23	**29.71**
LibraryC	Camera Pan	#11-40	35.65	36.94	36.95	**37.36**	37.29	30.32	31.67	31.80	31.97	**32.06**
LibraryE	Camera Shaky	#11-40	34.38	36.59	36.67	36.86	**37.14**	28.99	29.47	30.04	29.70	**30.23**
LectureA	Camera Pan	#11-40	37.03	37.78	37.86	37.84	**37.87**	32.18	32.67	32.83	32.80	**32.93**
LectureB	Camera Translation	#11-40	33.20	34.20	34.16	**34.44**	34.37	28.48	29.07	29.22	29.28	**29.43**
TestchartB	Object Translation	#11-40	33.84	35.01	35.36	35.36	**35.66**	27.82	27.93	28.19	28.06	**28.36**
AlfaC	Object Shaky	#11-40	33.77	36.10	36.17	36.68	**37.05**	27.73	28.95	29.36	29.51	**30.32**
CarparkA	Object Complex	#11-40	32.27	32.76	32.41	**32.78**	32.42	27.94	27.96	27.75	**28.00**	27.81
Average gain over bicubic			−−	1.30	1.38	1.55	**1.68**	−−	0.56	0.81	0.78	**1.07**

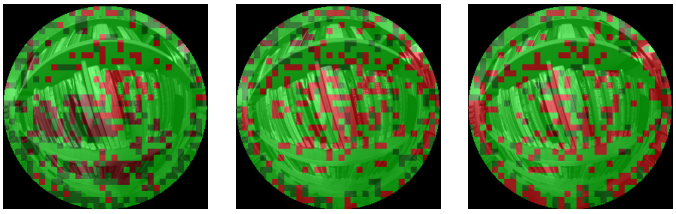

Figure 6. ME decision masks for different frame distances. Blocks where sCME+ is selected are depicted in green, while those where sTME is chosen are given in red. Subpixel-accurate MVs are marked with darker colors.

where $\hat{r}'_{\mathrm{fe}}$ denotes the manipulated radius. These three adaptations lead to a correct motion estimation for ultra-wide angles.

IV. Simulation Results

In this section, the effects of integrating the proposed sCHME+ approach into a multi-image SR framework are analyzed with regard to objective quality and visual examples. The objective quality is measured by the luminance PSNR that is evaluated inside the 185 degree circle, which corresponds to the camera FOV. That way, the black regions not containing any relevant information are neglected. For this paper, the following eight real-world fisheye video sequences are taken from the fisheye data set provided in [19]: *LibraryA*, *LibraryC*, *LibraryE*, *LectureA*, *LectureB*, *AlfaC*, *TestchartB*, and *CarparkA*. For the purpose of calibration, the checkerboard images *Calibration* also available in the above-mentioned data set are used. These sequences all have an original resolution of 1150×1086 and show different types of motion and content. For further processing, these sequences were cropped in horizontal dimension and zero-padded in vertical dimension to form a spatial resolution of 1088×1088. For the purpose of simulation, LR sequences were created in order to analyze resolution enhancement factors U of 2 and 4. To that end, a degradation by a Gaussian blur of size 3×3 (5×5) and corresponding variance 0.75 (1.0) as well as a downsampling by a factor of 2 (4) in each dimension was applied to the original sequences. The SR results were calculated for 30 consecutive frames of each video sequence with a number of utilized frames N ranging from 3 to 15. For ME itself, a block size of 8 and a search range of 32 was chosen.

As a first evaluation on how well the proposed ME technique performs, the block decision on whether a traditional integer MV (red) / subpixel MV (dark red) or a fisheye integer MV (green) / subpixel MV (dark green) is chosen as the best candidate is illustrated in Fig. 6 for the sequence *LibraryC* and different frame distances. It is evident that the majority of the blocks is chosen from the fisheye path indicating that the fisheye-adapted methods performs well. The larger the distance to the reference frame, the more traditional MVs are selected near higher values of θ. The reason for that is an insufficient search range in the perspective domain as pixel positions spread further away.

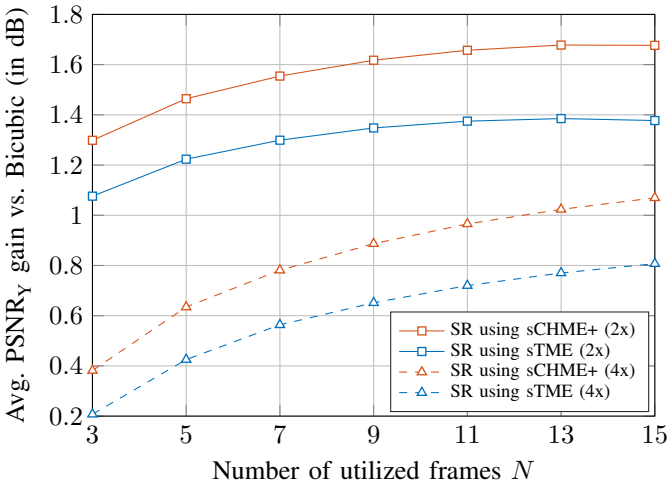

Figure 7. Average luminance PSNR gains for SR using sTME (blue) and for the proposed SR using sCHME+ (red) compared to bicubic interpolation realizing an upscaling of both U=2 and U=4.

For the actual SR tests, the following three reconstruction methods are compared to each other: bicubic interpolation, SR using sTME, and the proposed SR using sCHME+. The average luminance PSNR gains over all frames and sequences compared to bicubic interpolation for both U=2 and U=4 as well as a varying number of utilized frames N are depicted in Fig. 7. As can be seen, the proposed method can achieve a gain for all combinations of U and N. The corresponding PSNR values for all sequences, upscaling factors, and N=7 as well as N=15 are summarized in Table I. For U=2, the proposed method can achieve average gains over SR using sTME of 0.25 dB and 0.30 dB using N=7 and N=15, respectively. For U=4, an average gain of 0.22 dB and 0.26 dB is obtained.

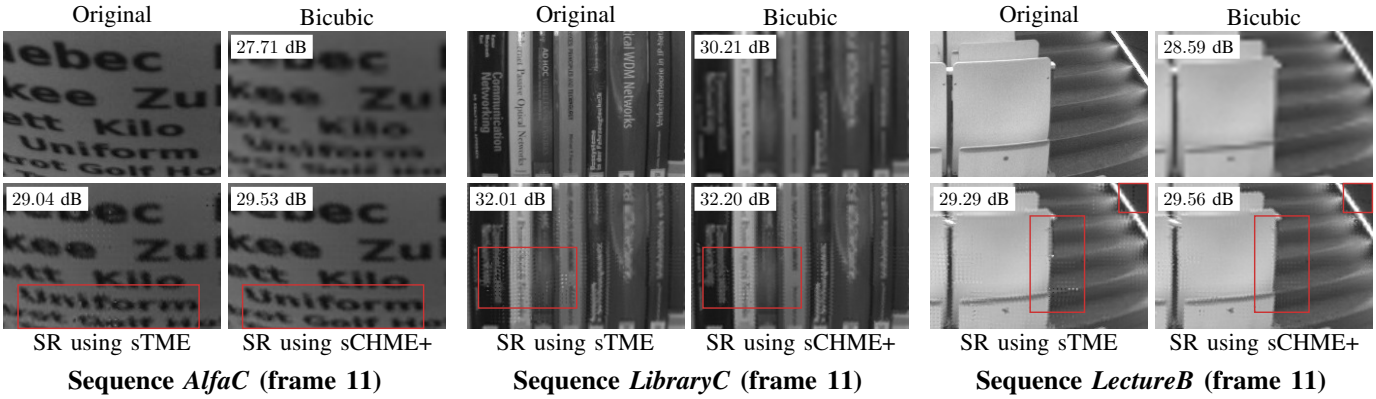

Figure 8. Visual quality comparison of image detail examples for three sequences using an upscaling of factor 4 and 15 utilized frames. Areas of interest are highlighted in red (best viewed enlarged on screen).

The aforementioned objective results are substantiated by the visual results for a resolution enhancement factor of 4 and $N = 15$ and are illustrated in Fig. 8. Image detail examples for the three sequences *AlfaC*, *LibraryC*, and *LectureB* are presented and areas of interest are highlighted in red. It is evident that reconstruction artifacts arising from an incorrect motion estimation are successfully reduced, increasing the legibility in the left example and removing disturbing artifacts in the middle and right examples.

V. Conclusion

In this paper, a multi-image SR approach for fisheye video sequences was presented. Instead of simply employing a traditional block-based ME technique, we proposed to use a calibrated hybrid ME that takes the fisheye characteristics into account and hence is capable of reducing artifacts from an incorrect motion estimation. Our proposed method was evaluated for a varying number of utilized frames and upscaling factors of 2 and 4. It was shown that our method achieves an average gain in luminance PSNR of up to 0.3 dB for both upscaling factors. Furthermore, it successfully reduces artifacts and improves the reconstruction quality. Future work will include employing the proposed strategy on coded video sequences, developing an adaptive motion search to better cope with wider angles, and evaluating other ME and SR methods.

Acknowledgement

This work was supported by Bosch Sicherheitssysteme Engineering GmbH, Nuremberg, Germany.

References

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-Resolution Image Reconstruction: A Technical Overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.

[2] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[3] K. I. Kim and Y. Kwon, "Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1127–1133, June 2010.

[4] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image Super-Resolution Via Sparse Representation," *IEEE Trans. Image Processing*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[5] T. Peleg and M. Elad, "A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution," *IEEE Trans. Image Processing*, vol. 23, no. 6, pp. 2569–2582, June 2014.

[6] Y. Zhu, Y. Zhang, and A. L. Yuille, "Single Image Super-resolution Using Deformable Patches," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June 2014, pp. 2917–2924.

[7] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Super-Resolution," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, Barcelona, Spain, Sep. 2003, pp. II–291–4 vol.3.

[8] D. Mitzel, T. Pock, T. Schoenemann, and D. Cremers, "Video Super Resolution Using Duality Based TV-L1 Optical Flow," in *Proc. 31st DAGM Symposium on Pattern Recognition*, Jena, Germany, Sep. 2009, pp. 432–441.

[9] A. Danielyan, R. Foi, V. Katkovnik, and K. Egiazarian, "Image and Video Super-Resolution via Spatially Adaptive Block-Matching Filtering," in *Proc. Int. Workshop on Local and Non-Local Approximation in Image Processing*, Lausanne, Switzerland, Aug. 2008.

[10] M. Bätz, A. Eichenseer, J. Seiler, M. Jonscher, and A. Kaup, "Hybrid Super-Resolution Combining Example-Based Single-Image and Interpolation-Based Multi-Image Reconstruction Approaches," in *Proc. IEEE Int. Conf. on Image Processing*, Quebec City, Canada, Sep. 2015.

[11] K. Miyamoto, "Fish Eye Lens," *Journal of the Optical Society of America*, vol. 54, no. 8, pp. 1060–1061, Aug. 1964.

[12] D. Dooley, B. McGinley, C. Hughes, L. Kilmartin, E. Jones, and M. Glavin, "A Blind-Zone Detection Method Using a Rear-Mounted Fisheye Camera With Combination of Vehicle Detection Methods," *IEEE Trans. Intelligent Transportation Systems*, vol. 17, no. 1, pp. 264–278, Jan. 2016.

[13] G. H. Lee, F. Faundorfer, and M. Pollefeys, "Motion Estimation for Self-Driving Cars with a Generalized Camera," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Portland, Oregon, USA, June 2013, pp. 2746–2753.

[14] A. Eichenseer, M. Bätz, J. Seiler, and A. Kaup, "A Hybrid Motion Estimation Technique for Fisheye Video Sequences Based on Equisolid Re-Projection," in *Proc. IEEE Int. Conf. on Image Processing*, Quebec City, Canada, Sep. 2015, pp. 3565–3569.

[15] A. Eichenseer, M. Bätz, and A. Kaup, "Motion Estimation for Fisheye Video Sequences Combining Perspective Projection with Camera Calibration Information," in *Proc. IEEE Int. Conf. on Image Processing*, Phoenix, AZ, USA, Sep. 2016.

[16] M. Santamaría and M. Trujillo, "A Comparison of Block-Matching Motion Estimation Algorithms," in *Proc. IEEE 7th Colombian Computing Congress*, Medellín, Colombia, Oct. 2012, pp. 1–6.

[17] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," *Journal of the Optical Society of America*, vol. 62, pp. 55–59, Jan. 1972.

[18] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A Toolbox for Easily Calibrating Omnidirectional Cameras," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Beijing, China, Oct. 2006, pp. 5695–5701.

[19] A. Eichenseer and A. Kaup, "A Data Set Providing Synthetic and Real-World Fisheye Video Sequences," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Shanghai, China, Mar. 2016, available online: `www.lms.lnt.de/fisheyedataset`.