Improving Multiple-Image Super-Resolution for Mobile Devices through Image Alignment Selection

Neil Patrick Del Gallego
De La Salle University
Taft Avenue, Malate
1004 Metro Manila,
Philippines
neil.delgallego@dlsu.edu.ph

Joel Ilao
De La Salle University
Taft Avenue, Malate
1004 Metro Manila,
Philippines
joel.ilao@delasalle.ph

ABSTRACT

Multiple-image super-resolution (MISR) attempts to recover a high-resolution (HR) image from a set of low-resolution (LR) images. In this paper, we present a mobile MISR tailored to work for a wide range of mobile devices. Our technique aims to address misalignment issues from a previous work and further enhance the quality of HR images produced. The proposed architecture is used to implement a prototype application that is freely available at Google Play Store, titled Eagle-Eye HD Camera. The system is divided into the following modules: Input Module, Edge Detection Module, Image Selection Module, Image Alignment Module, Alignment Selection Module and Image Fusion Module.

We assessed the quality of HR images produced by our mobile MISR, through an online survey, as well as compare it with other related SR works. Performance time was also measured. A total of 114 respondents have participated in the survey, where majority of respondents preferred our approach. Our approach is observed to be comparable with other SR works in terms of visual quality and performance time, and guaranteed to work in a mobile environment.

Keywords
Super-resolution, Mobile Devices, Mean Fusion, Image Alignment

1 INTRODUCTION

Multiple-image super-resolution (MISR) attempts to recover a high-resolution (HR) image from a set of low-resolution (LR) images. Figure 1 shows HR images produced by our proposed mobile multiple-image super-resolution (MMISR) system for mobile devices. To the best of the authors' knowledge, there are limited studies and implementations of super-resolution on mobile devices, presumably because of its high time and space complexity. However, mobile devices are already capable of implementing a mobile MISR system, provided that the system makes efficient use of its hardware resources. Mobile device manufacturers such as ASUS[1], Vivo[2], and OPPO[3] include a camera feature that mimics an MISR technique to capture HR images. Similar MMISR studies were observed from [Chu13], [ZC14], [ZWZ13], and [DS15].

Images obtained from mobile devices may be modeled as having undergone a series of noise, downsampling and motion blur, which is similar to the image degradation model proposed in [MPSC09]. The goal of any MMISR system is to reverse these degradation effects.

An MISR technique can be divided into the following steps [NM14]: denoising, deblurring or image selection, alignment, upsampling and image fusion. In our implementation, the steps are performed in a sequential manner and memory is being managed by our matrix pool discussed in Section 4.1. The contributions of this study can be summarized below:

[1] How to Shoot Super Resolution on ZenFone 4: `https://youtu.be/o3DFhZxzwtk`
[2] Vivo V7. 64MP Ultra HD Photos: `https://www.vivo.com/product/en/product/v7`
[3] OPPO Pure Image, Ultra HD: `https://www.oppo.com/en/technology/pure-image`

A   **B**   **A**   **B**

Figure 1: Sample HR images produced by the SR system with noticeable improvement. A: Bicubic interpolation.
B: Proposed SR method.

1. We implemented and published a free prototype application, titled Eagle-Eye HD Camera[4].

2. We presented a revised architecture for performing MMISR, that aims to address misalignment issues from the work of [DGI17].

3. We compared the quality of resulting HR images with the Bicubic baseline, and other related SR works. We also conducted an online survey to assess our HR images, where 114 respondents have participated. The survey shows promising results to further improve our work.

4. We developed a method for objectively assessing the quality of aligned images. The MMISR prototype performs alignment selection based on this assessment method.

Our paper is organized as follows: We review recent work in SR for mobile devices in Section 2. Specifically, we discuss limitations of [DGI17] and how we address these in Section 3. Lastly, we discuss our results in Section 5 and conclusions in Section 6.

2 RELATED WORK

Image super-resolution is still needed despite the advances in hardware such as the introduction of high-definition (HD) displays. Some high-end mobile devices introduced as of 2018 have a resolution of 1440 × 2560 known as quad-HD displays and camera resolution size may go as large as 40MP [YSL+16][5]. However, development of MMISR seems to be limited due to its computational cost. Mobile devices are typically equipped with burst mode capture which can be utilized to perform MISR, provided that the images captured

have substantially different pixel values. This is proven in the work of [DGI17], that there are substantial differences from images captured using the burst mode of the camera [DGI17]. [Chu13] proved that multiple images captured from mobile devices result in small motions due to high frame rates, where an affine flow model is suitable for aligning the images. A joint image alignment and deblurring approach was also proposed by [ZC14].

The MMISR system developed by [DGI17] mostly works when the user captures images steadily, or the subject have adequate lighting. Limitations and issues were observed, which are summarized below:

1. Limitation **L1**: Images become misaligned whenever images are captured with shaky hands. While the system has some tolerance for aligning images with slight angular changes, it can only work ideally when images are captured steadily. This scenario may not be practical for most end-users. An affine and perspective transformation estimation was used to align the images which proves to be an insufficient approach as discussed in Section 2.1.

2. Limitation **L2**: Inadequate lighting and changes in exposure values also affects the alignment.

3. Limitation **L3**: Using a mean fusion approach may smoothen the pixel values. While mean fusion can be effective for removing noise, it also causes some high-frequency details to be lost. The problem is that the system does not employ any regularization-based methods as observed from related approaches [MPSC09, NMG01, LHG+10, PC12, YZS12].

2.1 Misalignment Issues

The MMISR system of [DGI17] implements Affine Transformation Estimation (ATE) [Ho15] and Perspective Transform Estimation (PTE) [Ho15] sequentially. However, their technique can produce misaligned images and causes unwanted artifacts to appear in the

---

[4] Eagle-Eye HD Camera: `https://play.google.com/store/apps/details?id=neildg.com.eagleeyesr`

[5] Huawei P20 Pro: `https://consumer.huawei.com/en/phones/p20-pro`

HR image. Figure 2 exhibits blurring and distortion of texts due to incorrect transform estimation. Figure 3 shows a low-light image example taken with varying exposure values (-3, 0, +3 respectively). In this example, a warping distortion occurred due to lack of image detail or varying exposure value. There is not enough reliable keypoints available for correctly estimating the perspective transformation. Figure 4 introduces a ghosting effect. ATE and PTE is performed globally on the whole image. Hence, it cannot handle localized transformation on image regions (Image B and C).
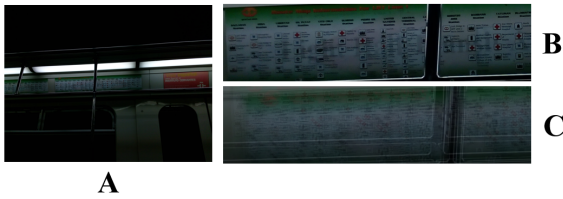

Figure 2: Misalignment example due to incorrect transform estimation. A: one of the LR image sequences. B: zoomed region on one of the images. C: zoomed region on an image with misalignment.
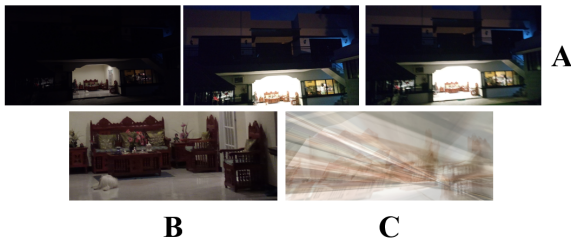

Figure 3: Low-light images with varying exposure values (EV) are prone to misalignments. A: Low-light LR images with -3, 0, +3 EV. B: zoomed region on one of the images. C: zoomed region on an image with warping distortion.

Figure 4: Misalignment example on a scenery image. A: one of the LR image sequences. B: zoomed region on one of the images. C: zoomed region on an image with ghosting effect.

2.2 Loss of Detail after Mean Fusion

As mentioned in **L3**, it is observed from the method of [DGI17] that the nature of the mean fusion process smoothens out some of the high-frequency details of the images as illustrated in Figure 5.
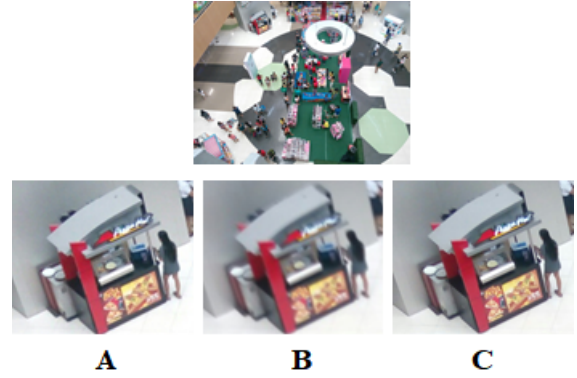

Figure 5: A: cubic interpolation. B: SR method by [DGI17]. C: ground-truth. Loss of detail is exhibited in result of B after performing mean fusion.

3 ADDRESSING THE LIMITATIONS

This section discusses our proposed approaches that aim to address **L1**, **L2** and **L3**. Our proposed MMISR system selects well-aligned images from the outputs of two alignment algorithms, PTE and MTB (Median Threshold Bitmap alignment [War03]). The choice of these alignment techniques are influenced by the following factors:

1. PTE can easily be performed on a mobile device because of its low computational cost and fast processing time (see performance time discussion). Images captured from mobile devices typically have a resolution of 8MP or more, and these techniques can handle images with large resolution.

2. MTB is observably the fastest and most reliable technique for aligning images captured on a mobile device [War03].

To verify the quality of image alignment algorithms, an experiment was performed that compared ATE, PTE and MTB through a fitness score. It is observed that misaligned images introduce additional edges when merged with the reference image as shown in Figure 6. With this observation, and because there are no standard mesures for assessing how an image is well-aligned, we propose a technique that measures the density of edges through Sobel derivatives [Sob68]. Correctly aligned images should not introduce additional edges from the reference image. To detect this observation, the following steps are performed:

1. Let L_0 be the first reference LR image, $\{A_1...A_N\}$ are aligned image sets produced by image alignment technique A.

2. Count the non-zero elements of the edge image for L_0 to produce an integer measure, e_0.
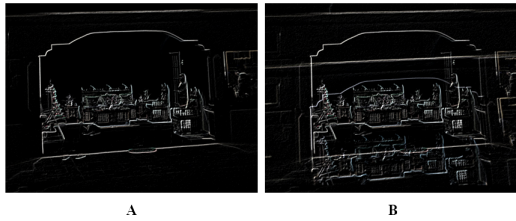
Figure 6: Misaligned images introduce additional edges. A: reference image edge. B: misaligned image merged with the reference image.

3. For $a \in \{A_1...A_N\}$, add a with L_0 to produce the edge image, as seen in Image B in Figure 6. Let this be $\{\bar{A_1}...\bar{A_N}\}$

4. Count the non-zero elements of the edge images for $\{\bar{A_1}...\bar{A_N}\}$. Let this be $\{\bar{e}_1...\bar{e}_N\}$.

5. Compute for the integer measures, $\{\varepsilon_i...\varepsilon_N\}$ by simply subtracting $\bar{e}_i$ to e_0, where i is $1...N$. Label this as *SobelMeasure*.

$\{\varepsilon_i...\varepsilon_N\}$ refers to the corresponding *SobelMeasure* values of aligned image set, $\{A_1...A_N\}$. A low value indicates that minimal edges were introduced when attempting to combine the aligned images to the reference image.

Using the proposed technique, 33 image sets where gathered and aligned, where each image set consists of 10 images captured using burst mode. The resulting average *SobelMeasure* of aligned image sets are visualized in Figure 7. Out of 33 image sets tested, PTE works best on 28 image sets, ATE works best for 3 image sets, and MTB alignment works best for 2 image sets. PTE is the ideal image alignment technique for images taken from mobile devices. However, MTB sometimes aligns an image sequence better than PTE. This is where we propose an alignment selection technique. We select an aligned image by selecting the lesser difference in *SobelMeasure*. Suppose P is the aligned image using PTE, and M is the aligned image using MTB, for some $\{L_0...L_N\}$ image sequence. If *SobelMeasure* of $P \leq M$, then P will be selected as the aligned image for that image sequence. Otherwise, M will be selected. This is demonstrated in Figure 8 where artifacts are severely reduced in the final image.

For addressing **L3**, an L_1-norm minimization approach proposed by [FREM04] may be applied or other regularization-based approaches in recent works [LHG+10, PC12, YZS12]. However, such a technique may result in a huge computational time for a mobile device due to its iterative nature. Based from this assumption, we simply applied a sharpening operation, unsharp masking, to individual LR images which proves to be effective in preserving edges while also removing noise as observed in Figure 9. Edges as
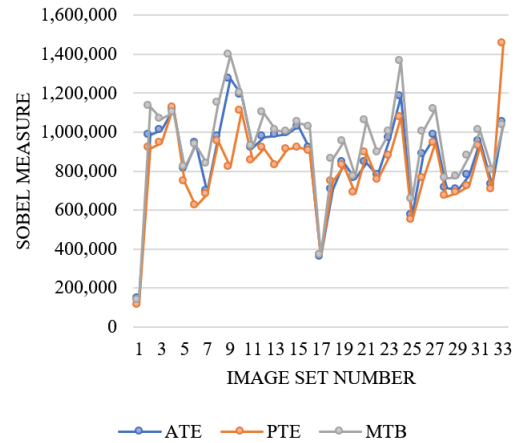

Figure 7: *SobelMeasure* values from 33 image sets visualized as a line chart. A lower value indicates that the image sequences are more well-aligned to its reference image. PTE has the lowest average *SobelMeasure* of **818,056**.

Figure 8: A: alignment using Perspective Transform Estimation. B: Best Alignment Technique. Aligning images with varying exposure values is a clear limitation of the MMISR system. This results in severe warping distortion only if (A) was applied. Misalignment and warping distortion is reduced if (B) was applied.

well as noise gets amplified but performing a mean fusion to combine all unsharp masked LR images will create an image where edges are preserved while also minimizing noise.

4 OUR PROPOSED ARCHITECTURE

Our system architecture borrows from the architecture presented in [DGI17], but modified that image alignment technique to address **L1** and **L2** and applied unsharp masking to LR images to mitigate **L3**. The system architecture is shown in Figure 10.

The system accepts a set of LR images wherein the first LR image serves as the reference LR image. In the Edge Detection and Image Selection Module, the LR images undergo the same feature-selection scheme proposed in [DGI17]. The Image Selection Module produces a filtered set of LR images, $\{L_0...L_N\}$ where an Unsharp Masking operator is applied to the images, to

Figure 9: Unsharp masking illustration. A: LR image. B: LR image with unsharp masking applied. C: Resulting image after performing mean fusion. Observe that noise and other artifacts seen in B was suppressed, while having the edges preserved in C.

address **L3**. Unlike [DGI17], non-local means denoising [BCM11] can be applied optionally, to conserve computation time. L_0 is upsampled using bicubic interpolation, which becomes the initial HR image $\hat{H}$.

The subset $\{L_1...L_N\}$ undergoes PTE [Ho15], and MTB alignment [War03]. This produces warped images, $\{P_1...P_N\}$ for the PTE-aligned images and $\{M_1...M_N\}$ for MTB-aligned images.

The image sets $\{P_1...P_N\}$ and $\{M_1...M_N\}$ enter the Image Alignment Module where an image that introduces the least error in alignment will be chosen, which addresses **L1** and **L2**. The selected aligned image, $\{W_k\}_{k=1}^N$, can either be $\{P_k\}_{k=1}^N$ or $\{M_k\}_{k=1}^N$ respectively.

The images $\{W_1...W_N\}$ are upsampled using bicubic interpolation to produce $\{\hat{H}_1...\hat{H}_N\}$ as initial images to be mapped to the HR grid, by mean fusion. This produces the final HR image, H.

4.1 Matrix Pooling for Optimization

We discuss matrix pooling as an optimization technique that made our MMISR implementation possible. In our implementation, we prioritized memory management over computational time to minimize the chances of out of memory errors.

Matrix pooling is heavily inspired from the object pool software design pattern, but applied to matrices. N matrices of size $H \times W$ are pre-allocated at startup. Each module may request for M matrices where $M \leq N$ for processing. After a task has been performed, M matrices are released back to the pool of N matrices. Should $M > N$, then this returns a failure. Otherwise, the tasks

in the system modules perform as is. Because of matrix pooling, memory can easily be managed and results in faster computational time, as the matrices are only instantiated during the start of an SR task, and destroyed when the SR task is completed.

5 RESULTS AND DISCUSSION

5.1 Assessment of HR Images

A preliminary survey was conducted to assess the quality of images. 114 respondents have participated. These images do not have any ground-truths as they were captured in a real-world scenario. The respondents where tasked to evaluate zoomed images. The survey is structured as follows: a total of 42 randomly selected image sets with HR images were requested to be evaluated by the respondents. There are 3 choices to choose from, **Method A**: a bicubic interpolated image, **Method B:** an HR image produced by the previous SR method [DGI17], and **Method C**: our proposed SR method. In the survey, an image thumbnail is provided followed by the 3 HR images zoomed in on a certain region, as illustrated in Figure 11. The image choices were randomized per question so that respondents will not discover any patterns in the choices.

The respondents choose one of the 3 image choices provided, followed by a confidence level rating from 1 to 5. This indicates the confidence and certainty of their chosen preferred image. A rating of 5 means that the respondent is very sure of his/her image choice. A rating of 1 means that the respondent had difficulties choosing his/her preferred image or their decision is split among the other image choices.

5.2 Preliminary Survey Results

The survey results are summarized in Table 1. The "Number of Majority Votes" column tallies the number of test images where a given technique is selected by majority. The "Average Confidence Level 5" column indicates the average percentage of Level 5 ratings for a given technique.

It is shown in Table 1 that Method C was selected as majority for 26 test images with Average Confidence Level 5 of 41.95%. It is followed by the Method A were it was selected as majority for 14 test images with Average Confidence Level 5 of 39.16%. Method B were only selected for 2 test images and Average Confidence Level 5 of 37.30%

Based from the results, it can be justified that Method C were preferred by respondents over Method A and B. Whenever the produced HR images from Method C are not preferable to respondents, the HR images produced by Method A were selected.

Figure 13 shows the best quality HR images preferred by respondents (refer to Figure 12 for the thumbnails).
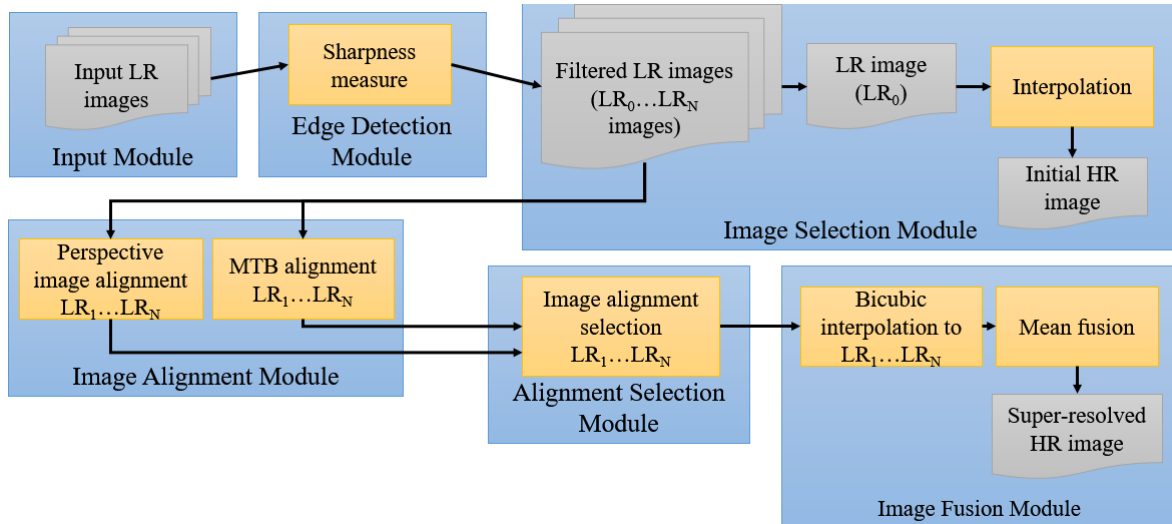
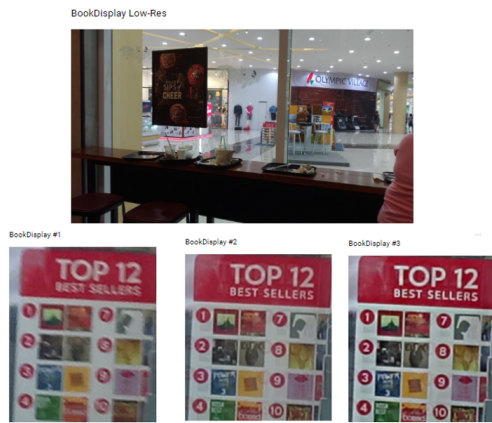Figure 10: System architecture of our MMISR system.

Figure 11: A snippet of the online survey. Image thumbnail is presented first followed by three image choices.

Table 1: Summary of survey results

Technique	Number of Majority Votes	Average Confidence Level 5
Method A	14 out of 42	39.16%
Method B	2 out of 42	37.30%
Method C	26 out of 42	41.95%

Image 1 received the highest number of votes (95.60%) for Method C, with a Level 5 Confidence of 69.30%. Image 5 has 90.40% of votes and also has the highest Level 5 Confidence Level Percentage, which is 72.80%. It can be observed from the best cases that users prefer clear texts and sharp details. Method C produced the clearest and most visible text than the other techniques in these image sets. Additional results are shown in Figure 18 and 19.

5.3 Performance Time

Processing time and space consumption were measured using a test device with 2.0 Ghz Octa-core processor, 4GB RAM, and 16MP rear camera. A total of 20 image sets were used and the processing time was averaged. The LR images have a size of 2992×5280 resolution, which is the default resolution size of the camera. The MMISR system produces 50MP HR images of 5980×10560. Figure 14 and 15 shows the average performance time and standard deviation respectively. The Image Alignment, Alignment Selection and Image Fusion module takes up at least 60 seconds to process. With this observation, the processing time of Image Alignment and Alignment Selection modules can be further reduced by having only 1 robust alignment technique implemented similar to how [ZC14] handled image alignment. Denoising has the highest standard deviation because the quantity of images selected by the Image Selection Module, varies across test sets. Each additional image also results in a huge increase in denoising time, which is why it was made an optional feature and only ideal for low-light images.

5.4 Comparison with Related SR Work

We compared our results to related SR works by using ten frames from the video provided in [FREM04]. The frames have a resolution of 49×57. In Figure 16, we compared our approach to the following: Bicubic baseline, [FREM04] because we want to compare its L_1 minimization approach with our simpler unsharp masking technique to address the "smoothening" effect of mean fusion, and [ZC14] because of its promising approach of joint alignment and deblurring. Using a scaling factor of 2, our MMISR method outperforms the Bicubic method and the method of [FREM04]. In terms of image quality and sharpness, the method of [ZC14] outperforms our MMISR. In terms of speed, MMISR

Figure 12: Thumbnails of the best quality HR images (Top 1 - 5), preferred by respondents.

Method	Image 1	Image 2	Image 3	Image 4	Image 5
A	2.60%	3.50%	7.90%	5.30%	9.60%
B	1.80%	2.60%	1.80%	4.40%	0.90%
C	95.60%	93.90%	90.40%	90.40%	89.50%

Figure 13: Best HR images preferred by respondents with the corresponding percentage of votes. 114 respondents have participated in the survey. Images 1 - 5 received above 89% of votes in Method C (our method), which effectively surpasses the bicubic performance (Method A) and the SR method of [DGI17]. It can be observed from the best cases that users prefer clear texts and sharp details.

is considerably faster than the method of [ZC14]. The performance time of the method of [ZC14] is **155.63 seconds** compared to MMISR which is less than **1 second**. The deblurring and denoising stage in the approach of [ZC14] were the most time-consuming. While the technique of [ZC14] is robust and can handle extremely blurred and noisy images, MMISR is considerably faster and more appropriate for mobile devices.

We also compared our results to the work of [KJ14], because they perform a specialized approach in image up-sampling through self-learning. Additionally, we compared our results to the work of [KJ13] and [SLJT08]. The test images in the work of [KJ14] were re-captured from a computer screen, using a mobile device with a 16MP camera. This was performed so that the mobile camera settings such as ISO, exposure and shutter speed, affect the quality of the HR images and provide a fair analysis against the mentioned SR techniques. Figure 17 shows the results. Our proposed technique, produces comparable results as that of related single-image SR works. It can be observed that our technique produces clearer edges among other SR methods.

6 CONCLUSION AND FUTURE WORK

This research presents an improved framework for implementing a mobile multiple-image super-resolution system (MMISR) for mobile devices, by addressing the limitations observed in the implementation of [DGI17]. The system architecture is divided into the following modules: Input Module, Edge Detection Module, Image Selection Module, Image Alignment Module, Alignment Selection Module and Image Fusion Module.

Our results, based from the survey and analysis of the performance time, show a promising direction in MMISR research. Immediate steps needed to further improve and validate our system is to compare it with other state-of-the-art approaches, and apply some of the techniques seen in single-image SR works [TDSVG15,
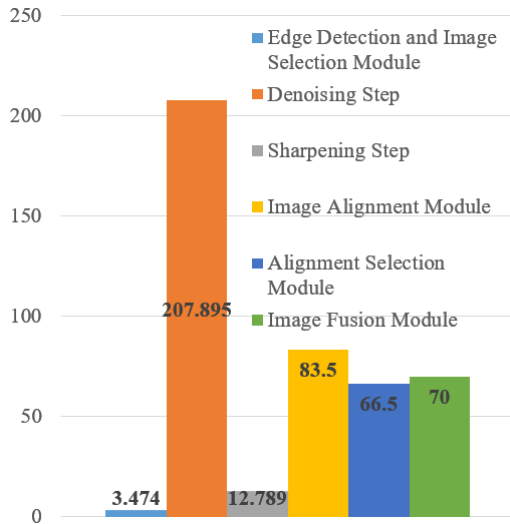
Figure 14: Average performance time of system modules in seconds. Denoising is an optional feature in the actual application due to its very long processing time. Performing image alignment, and image fusion are the heaviest in terms of processing time.
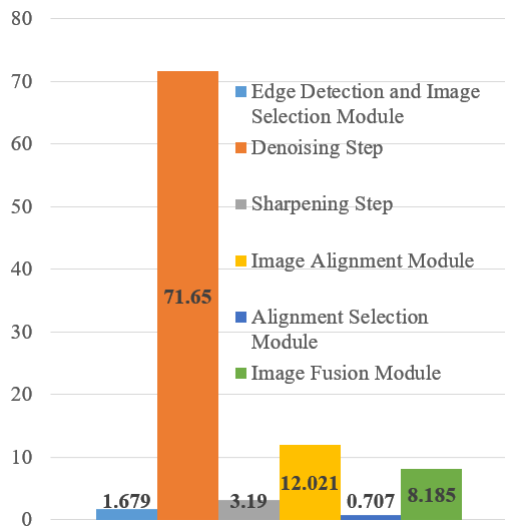

Figure 15: Standard deviation of system modules in seconds. Denoising has the highest standard deviation because the quantity of images selected by the Image Selection Module, varies across test sets.

DLHT16, TAe17]. We also plan to implement specialized approaches in image upsampling [RU90, NTP17].

Based from the preliminary survey conducted, results show that respondents prefer our approach. However, a more thorough analysis on the preferences of users must be performed [YMY14]. Using an Unsharp Masking operator to solve **L3** may introduce artificial artifacts in the images. Thus, it is recommendable that a specialized sharpening operation is performed such that it preserves the natural contours and composition of the images, which is an interesting approach in this paper [RIM17].
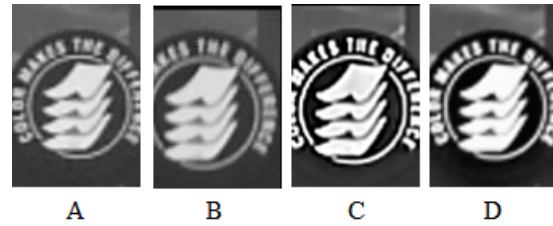

Figure 16: Comparison with related SR work. A: Bicubic image. B: Method of [FREM04]. C: Method of [ZC14]. D: Our method. While Method C contains more high-frequency details than Method D, Method D is considerably faster than Method C, but also has more detail than Method A and B.

REFERENCES

[BCM11]   Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Non-local means denoising. *Image Processing On Line*, 1, 2011.

[Chu13]   Chung-Hua Chu. Super-resolution image reconstruction for mobile devices. *Multimedia Systems*, 19(4):315–337, July 2013.

[DGI17]   Neil Patrick Del Gallego and Joel Ilao. Multiple-image super-resolution on mobile devices: an image warping approach. *EURASIP Journal on Image and Video Processing*, 2017(1):15, 2017.

[DLHT16]   C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, Feb 2016.

[DS15]   M. Delbracio and G. Sapiro. Burst deblurring: Removing camera shake through fourier burst accumulation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2385–2393, June 2015.

[FREM04]   S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, Oct 2004.

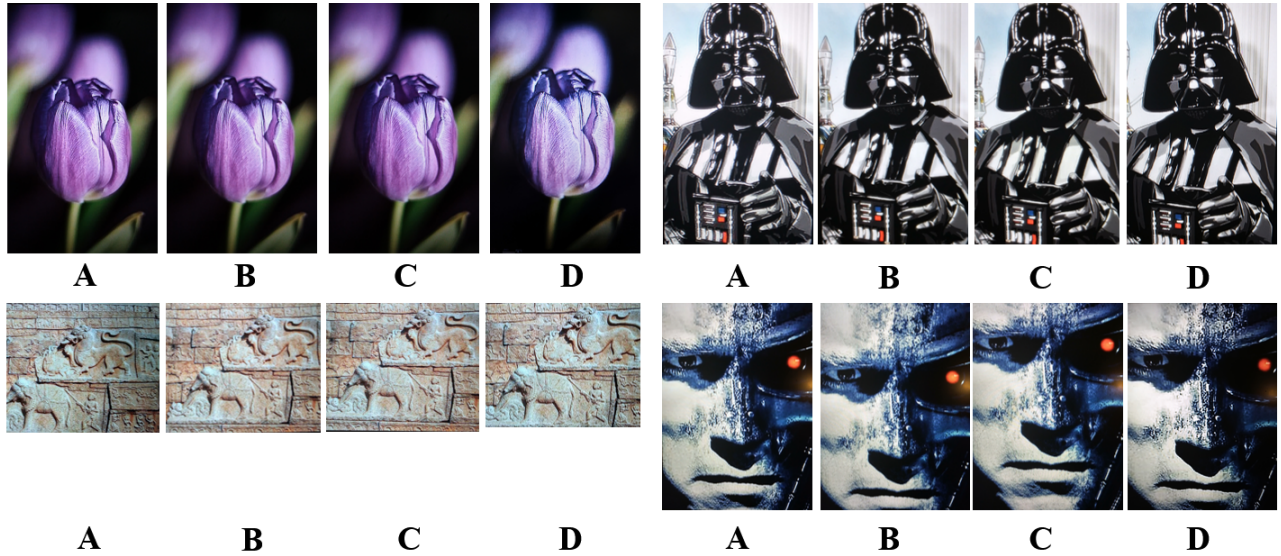[Ho15]   Nghia Ho. Understanding opencv estimate rigid tranform, 2015.

Figure 17: Comparison with related SR work. A: Method by [SLJT08]. B: Method by [KJ13]. C: Method by [KJ14]. D: Our method. Our method is observed to produce clearer edges.

Figure 18: Thumbnails of the best quality HR images (Top 6 - 10), preferred by respondents.

Method	Image 6	Image 7	Image 8	Image 9	Image 10
A	0.90%	0.0%	1.80%	3.50%	4.40%
B	9.60%	15.80%	18.40%	17.50%	16.70%
C	**89.50%**	**84.20%**	**79.80%**	**78.90%**	**78.90%**

Figure 19: Best HR images (Top 6 - 10) preferred by respondents with the corresponding percentage of votes. 114 respondents have participated in the survey. Method A: Bicubic image. B: SR method of [DGI17]. C: Our method.

[KJ13]   Nilay Khatri and Manjunath V. Joshi. Image super-resolution: Use of self-learning and gabor prior. In Kyoung Mu Lee, Yasuyuki Matsushita, James M. Rehg, and Zhanyi Hu, editors, *Computer Vision – ACCV 2012*, pages 413–424, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[KJ14]   N. Khatri and M.V Joshi. Efficient self-learning for single image upsampling. In *22nd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2014)*, pages 1–8, 2014.

[LHG+10] Xuelong Li, Yanting Hu, Xinbo Gao, Dacheng Tao, and Beijia Ning. A multiframe image super-resolution method. *Signal Processing*, 90(2):405 – 414, 2010.

[MPSC09] Dennis Mitzel, Thomas Pock, Thomas Schoenemann, and Daniel Cremers. Video super resolution using duality based tv-l1 optical flow. *Proceedings of the 31st DAGM Symposium on Pattern Recognition*, pages 432–441, 2009.

[NM14]   Kamal Nasrollahi and Thomas B. Moeslund. Super-resolution: A comprehensive survey. *Machine Vision Applications*, 25(6):1423–1468, August 2014.

[NMG01]   Nhat Nguyen, P. Milanfar, and G. Golub. A computationally efficient superresolution image reconstruction algorithm. *IEEE Transactions on Image Processing*, 10(4):573–583, Apr 2001.

[NTP17]   Mattia Natali, Giulio Tagliafico, and Giuseppe Patan. Local up-sampling and morphological analysis of low-resolution magnetic resonance images. *Neurocomput.*, 265(C):42–56, November 2017.

[PC12]   P. Purkait and B. Chanda. Super resolution image reconstruction through bregman iteration using morphologic regularization. *IEEE Transactions on Image Processing*, 21(9):4029–4039, Sept 2012.

[RIM17]   Y. Romano, J. Isidoro, and P. Milanfar. Raisr: Rapid and accurate image super resolution. *IEEE Transactions on Computational Imaging*, 3(1):110–125, March 2017.

[RU90]   S. P. Raya and J. K. Udupa. Shape-based interpolation of multidimensional objects. *IEEE Transactions on Medical Imaging*, 9(1):32–42, Mar 1990.

[SLJT08] Qi Shan, Zhaorong Li, Jiaya Jia, and Chi-Keung Tang. Fast image/video upsampling. *ACM Transactions on Graphics*,

27(5):153:1–153:7, December 2008.

[Sob68]   Feldman G. Sobel, I. A 3x3 isotropic gradient operator for image processing. *Stanford Artificial Intelligence Project*, 1968.

[TAe17]   R. Timofte, E. Agustsson, and L. V. Gool et.al. Ntire 2017 challenge on single image super-resolution: Methods and results. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1110–1121, July 2017.

[TDSVG15] Radu Timofte, Vincent De Smet, and Luc Van Gool. A+: Adjusted anchored neighborhood regression for fast super-resolution. *Computer Vision – ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part IV*, pages 111–126, 2015.

[War03]   Greg Ward. Fast, robust image registration for compositing high dynamic range photographs from handheld exposures. *Journal of Graphics Tools*, 8:17–30, 2003.

[YMY14]   Chih-Yuan Yang, Chao Ma, and Ming-Hsuan Yang. Single-image super-resolution: A benchmark. In *Proceedings of European Conference on Computer Vision*, 2014.

[YSL+16] Linwei Yue, Huanfeng Shen, Jie Li, Qiangqiang Yuan, Hongyan Zhang, and Liangpei Zhang. Image super-resolution: The techniques, applications, and future. *Signal Processing*, 128:389 – 408, 2016.

[YZS12]   Q. Yuan, L. Zhang, and H. Shen. Multiframe super-resolution employing a spatially weighted total variation model. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(3):379–392, March 2012.

[ZC14]   H. Zhang and L. Carin. Multi-shot imaging: Joint alignment, deblurring, and resolution-enhancement. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2925–2932, June 2014.

[ZWZ13]   H. Zhang, D. Wipf, and Y. Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1051–1058, June 2013.