INTERNATIONAL JOURNAL OF INNOVATIVE TRENDS IN ENGINEERING (IJITE) ISSN: 2395-2946
ISSUE: 56, VOLUME 36, NUMBER 01, 2017

An Extensive Review on Super Resolution of Real-Life Video Sequences Methods

Ramesh Kumar[1], Prof. Sumit Sharma[2]

[1]*Mtech. Scholar,* [2]*Research Guide*

Department of Computer Science and Engineering, VITS, Bhopal

Abstract - Image superresolution refers to methods that increase spatial resolution by fusing information from a sequence of images, acquired in one or more of several possible ways. The high resolution filtered image is constructed from the aliased (undersampled) noisy and blurred frames with either subpixel shifts or the use of intentional blurring by designing lenses with different point spread functions. Image resolution is determined by two main factors. Blurring, due to optical limits and various other processes while low-sensor density of the imaging device causes aliasing. A video camera is required to deliver a video sequence at desired frame-rate and spatial resolution. Fulfilling this demand is a challenge for some applications due to physical limitations of imaging systems. Obviously, high-resolution, high frame-rate video of a scene is desirable because it contains more recognizable details, and is more pleasant. this work present a survey on super resolution for real life video sequences.

Index Terms- Image superresolution, Video Sequences, Image Processing, Image Blurring, Image Enhancement, blind estimation, SR method.

I. INTRODUCTION

A demand for higher resolution is seen in many fields including bio-medical imaging (for purposes like image-guided surgery and image-assisted medical diagnosis), entertainment (high definition television or HDTV), satellite and astronomical imaging, chemical and biological research (high resolution electron microscopy), military surveillance and remote sensing. Indeed, this demand, in many cases, exceeds the maximum resolution capability of current acquisition systems. In other words, the current state of image sensor technology acts as a limiting factor in acquisition. In various other cases, factors such as cost, physical attributes like size and weight, and quality, instead of technology, constrain the maximum resolution obtainable from the acquisition device. Yet another set of cases exist in which resolution is compromised due to the need for flexibility and robustness with respect to various environmental conditions, among others. In all such cases, the solution to the need for higher resolution necessitates the design of technological devices in the form of digital image processing algorithms to satisfy the demand for high quality and high resolution (HR) images and video.

A trade off exists between frame-rate and resolution, where improving both at the same time is either not possible or leads to an expensive or heavy imaging device, which is not practical for many applications. As an instance, consider a consumer video recorder that uses CMOS image sensors. To increase resolution, one approach is to increase pixel density of the image sensor. This may decreases the Signal to Noise Ratio (SNR) of the sensor output, when the size of the sensor remains the same and thus the area of each pixel on the chip decreases.

Increasing recording time for each frame is equivalent to a lower frame-rate. Although this trade off may be solved by employing cutting edge of semiconductor technology, the associated cost usually renders that unsuitable for consumer applications. As a result, most of digital video cameras produce low-resolution videos at standard frame-rate. What if it need high- resolution video sequence in an application and high-quality imaging devices are not practical? Signal processing approaches may be employed to deliver desired video sequence from a non-ideal imaging system. These approaches may be adopted for an imaging application to keep the cost of the system acceptable, and render low- resolution imaging devices an option for more demanding applications.

II. SYSTEM MODEL

A digital video sequence is in fact the result of 3-dimensional time-space sampling of the scene. Each frame is a 2D representation of the scene, spatially sampled at a time instance. Multiple frames along time axis add the third dimension to the sequence, representing temporal sampling of the scene. Image sensor plays the important role of spatial sampling of the scene. Image sensor is a 2D array of CCD or CMOS light sensor units. Each unit converts the photons detected on its surface into electric signals. The number of light sensor units on the camera image sensor determines the maximum achievable spatial resolution of the camera. Camera electronic or mechanical shutter speed determines the frame-rate and exposure time. These characteristics play an important role in the quality of recorded frames. In addition to these two factors, many other factors such as size, pixel density, and dynamic range of image sensor (which is the ratio of the largest and smallest luminance detectable by the sensor), affect the final quality of recorded frame sequence.

The signal recorded by each light sensor unit is directly related to the number of photons detected on its surface. This number is related to the average light intensity in the corresponding area, the exposure time, and the area of each light sensor. The number of photons received from each point of the scene determines the accuracy of its intensity estimation in presence of different sources of noise. The exposure time must be sufficient to let a minimum number of photons arrive at the image sensor surface to produce an accurate measurement with a reasonable SNR; as a result, the maximum achievable frame-rate, which is inversely related to the exposure time, is limited. This maximum could increase if the light sensor area is increased and/or scene illumination is enhanced, but these are not always practical. Increasing the sensor size means that larger optical devices are required, moreover a larger and heavier camera is not capable of near-field focus. As a trade off, in multi-resolution image sensors, spatial resolution may be sacrificed to achieve higher frame-rates. In this approach, a fraction of maximum spatial resolution of image sensor is achieved at a higher frame-rate by accumulating photons detected over a larger area of the image sensor for each pixel. This method is called “Binning” and is popular for both CCD and CMOS sensors.

Although it seems to be possible to perform averaging operation off-chip, on-chip binning is dramatically more effective because sources of noise is suppressed directly on the image sensor chip before read- out in the case of on-chip binning. Therefore, image sensor has two or more grids to define area of each pixel. HR grid associates measurement of one light sensor to each output pixel, and LR grid associates many adjacent light sensor to each pixel (See Fig.1).

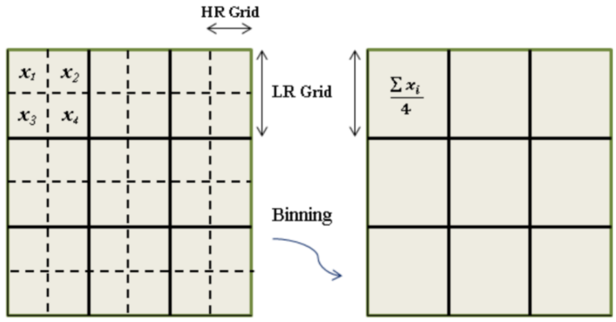

Figure 1.1 Pixel binning, LR and HR grid of image sensor.

III. LITERATURE REVIEW

SR NO.	TITLE	AUTHOR	Year	Approach
1	Blind Super Resolution of Real-Life Video Sequences	E. Faramarzi, D. Rajan, F. C. A. Fernandes and M. P. Christensen,	2016	A novel blind SR method is proposed to improve the spatial resolution of video sequences,
2	Super-resolution reconstruction for license plate image in video surveillance system,	Xiaole Yan, Qiu Shen and Xin Liu,	2015	Combine Fourier-Mellin transform (FMT) and Vandewalle's algorithm into a new technique
3	Video Super-Resolution via Deep Draft-Ensemble Learning,	R. Liao, X. Tao, R. Li, Z. Ma and J. Jia,	2015	SR drafts through the nonlinear process in a deep convolutional neural network (CNN).
4	Super-resolution of face image extracted from a video sequence,	M. Patil and S. D. Ruikar,	2014	Describes a technique to obtain a high resolution image from a given video sequence
5	A Novel Two-Step Approach for the Super-resolution Reconstruction of Video Sequences,	L. He, J. Tan, C. Xie and M. Hu,	2014	A two-step approach for the super-resolution reconstruction of video sequences based on the degraded model
6	Limitations of super resolution image reconstruction and how to overcome them for a single image	S. Gohshi and I. Echizen,	2013	It uses many low-resolution images to reconstruct a high-resolution image
7	A simultaneous method for 3D video super-resolution and high-quality depth estimation,	J. Zhang, Y. Cao and Z. Wang,	2013	A simultaneous method for video super-resolution and high-quality depth estimation of mixed-resolution 3D video

8	Super Resolution results in PANOPTES, an adaptive multi-aperture folded architecture	E. Faramarzi, V. R. Bhakta, D. Rajan and M. P. Christensen,	2013	(DSR) techniques on low resolution data collected using PANOPTES

E. Faramarzi, D. Rajan, F. C. A. Fernandes and M. P. Christensen,[1] Super resolution (SR) for real-life video sequences is a challenging problem due to complex nature of the motion fields. In this research, a novel blind SR method is proposed to improve the spatial resolution of video sequences, while the overall point spread function of the imaging system, motion fields, and noise statistics are unknown. To estimate the blur(s), first, a nonuniform interpolation SR method is utilized to upsample the frames, and then, the blur(s) is(are) estimated through a multiscale process. The blur estimation process is initially performed on a few emphasized edges and gradually on more edges as the iterations continue. Also for faster convergence, the blur is estimated in the filter domain rather than the pixel domain. The high-resolution frames are estimated using a cost function that has the fidelity and regularization terms of type Huber-Markov random field to preserve edges and fine details. The fidelity term is adaptively weighted at each iteration using a masking operation to suppress artifacts due to inaccurate motions. Very promising results are obtained for real-life videos containing detailed structures, complex motions, fast-moving objects, deformable regions, or severe brightness changes. The proposed method outperforms the state of the art in all performed experiments through both subjective and objective evaluations. The results are available online at http://lyle.smu.edu/~rajand/Video_SR/.

Xiaole Yan, Qiu Shen and Xin Liu [2] In order to improve the recognition of license plate texts in the real traffic surveillance video, super-resolution reconstruction (SR) method is applied to reconstruct a high-resolution (HR) image from consecutive frames in the video sequence. Current SR methods almost focus on small local translation and rotation, which limits the reconstruction of fast moving or significantly zooming objects, such as license plate images in surveillance video. In this research, combine Fourier-Mellin transform (FMT) and Vandewalle's algorithm into a new technique to improve the registration accuracy of license plate image. FMT is introduced for scaling estimation, while Vandewalle's algorithm is utilized for rotation and translation estimation. Additionally, the image reconstruction is carried out by the projection onto convex sets (POCS) method. The experiments on simulated and real image sequences are carried out respectively, and the results demonstrate that our approach can achieve better performance on reconstructing a HR license plate image.

R. Liao, X. Tao, R. Li, Z. Ma and J. Jia,[3] Propose a new direction for fast video super-resolution (VideoSR) via a SR draft ensemble, which is defined as the set of high-resolution patch candidates before final image deconvolution. Our method contains two main components -- i.e., SR draft ensemble generation and its optimal reconstruction. The first component is to renovate traditional feedforward reconstruction pipeline and greatly enhance its ability to compute different super resolution results considering large motion variation and possible errors arising in this process. Then combine SR drafts through the nonlinear process in a deep convolutional neural network (CNN). Analyze why this framework is proposed and explain its unique advantages compared to previous iterative methods to update different modules in passes. Promising experimental results are shown on natural video sequences.

M. Patil and S. D. Ruikar,[4] This research describes a technique to obtain a high resolution image from a given video sequence. The images obtained from inexpensive cameras are generally of low-quality and low-resolution and feeding those images to facial analysis systems generate undesirable outputs. The approach is to implement learning-based super-resolution algorithm on the low-resolution images to obtain high-resolution output. All the images extracted from the video are not useable in super-resolution algorithm. Therefore face quality assessment using facial feature extraction is utilized to discard the unwanted face images. Based on the quality score, it summarizes the input video sequence into a single best quality frontal face image. The employed super-resolution algorithm is applied on the best image resulting in an improved and enhanced high-quality, high-resolution image.

L. He, J. Tan, C. Xie and M. Hu,[5] In this research, propose a two-step approach for the super-resolution reconstruction of video sequences based on the degraded model. Firstly use the sparse principal component analysis and the linear minimum mean square-error estimation method to remove the noises from the degraded video sequences. Secondly adopt the Newton-Thiele's vector valued rational interpolation which is one of the nonlinear interpolation methods to magnify the results of the previous step. Our method is effective not only for gray video sequences, but also for color video sequences. Experimental results on a series of video sequences demonstrate that our method achieves better visual effects than those presented in [11] and [16], especially in details.

S. Gohshi and I. Echizen,[6] Super resolution image reconstruction (SRR) is a typical super resolution (SR) technology that has been researched with varying results. The SRR algorithm was initially proposed for still images. It uses many low-resolution images to reconstruct a high-resolution image. Unfortunately, in practice, rarely have a sufficient number of low-resolution images for SRR to work. Usually, there is only one (or a few) blurry images. On the other hand, there is a need to improve blurry images in applications ranging from security and photo restoration to zooming functions and countless other examples related to the printing industry. Recently, SRR was extended to video sequences that have many similar frames that can be used as low-resolution images to reconstruct high-resolution frames. In normal SRR, one reconstructs a high-resolution image from low-resolution images sampled from one high-resolution image, but in the video application, the low-resolution video frames are not taken from higher resolution ones. This research proposes a novel resolution improvement method that works without such a high- resolution image. Its algorithm is simple and can be applied to a single image and real-time video systems.

J. Zhang, Y. Cao and Z. Wang,[7] Mixed-resolution approach serves as a feasible solution to 3D video data reduction in limited bandwidth network environments, i.e., mobile networks. In this research, propose a simultaneous method for video super-resolution and high-quality depth estimation of mixed-resolution 3D video. Our method tackles the problem in a joint manner: i). Depth Estimation. calculate the initial depths by stereo matching, and then warp them according to the optical flow field. ii). Video Super-resolution. Under the guidance of the warped depth information, resolve the super-resolution problem by a fusion method, which involves a mapping step and a nonlocal reconstruction step. run the above two steps iteratively until obtaining the stable depth and super-resolution result. The experimental results on public 3D video sequences verify the effectiveness of our proposed method.

E. Faramarzi, V. R. Bhakta, D. Rajan and M. P. Christensen,[8] present experimental results of digital super resolution (DSR) techniques on low resolution data collected using PANOPTES, a multi-aperture miniature folded imaging architecture. The flat form factor of PANOPTES architecture results in an optical system that is heavily blurred with space variant PSF which makes super resolution challenging. also introduce a new DSR method called SRUM (Super-Resolution with Unsharpenning Mask) which can efficiently highlight edges by embedding an unsharpenning mask to the cost function. This has much better effect than just applying the mask after all iterations as a post-processing step.

IV. PROBLEM STATEMENT

In most imaging applications there is a tradeoff between the resolution and other parameters. For instance, to increase the resolution the pixel size is reduced in the sensor, but this reduces the amount of light available. It also generates shot noise that degrades the image quality severely. Another example is the processing and storage limitations that appear when taking a video with a mobile phone, which make the manufacturer limit the video resolution for these devices. Sensor technology is also important: while charged-coupled devices (CCD) and CMOS image sensors for current cameras offer sufficient resolution for most applications, forward looking infrared (FLIR) sensors have still many limitations. Besides of increasing the resolution, the idea of reducing the noise, deblur or minimize other degradations of an image emerges naturally when having a sequence of images that include the same region of interest in all of them.

V. CONCLUSION

This research brief the extensive survey on super resolution. Work presents survey of new computationally efficient algorithms for superresolution which offer better performance than other current methods. Typically, the superresolution module is followed by a separate denoising module. But in the case of the superresolution algorithms are presented in this research, this module is rendered unnecessary since denoising is achieved simultaneously with superresolution, and is hence implicit to the superresolution module. Some related work in the field of superresolution and its application in daily life video application has been discussed based on the work reported in their respective article. This research play a important role to implement new algorithm to improves on the performance of a popular and effective blind deconvolution algorithms which extends directly to the multi-frame case.

REFERENCES

[1] E. Faramarzi, D. Rajan, F. C. A. Fernandes and M. P. Christensen, "Blind Super Resolution of Real-Life Video Sequences," in IEEE Transactions on Image Processing, vol. 25, no. 4, pp. 1544-1555, April 2016.

[2] Xiaole Yan, Qiu Shen and Xin Liu, "Super-resolution reconstruction for license plate image in video surveillance system," 2015 10th International Conference on Communications and Networking in China (ChinaCom), Shanghai, 2015, pp. 643-647.

[3] R. Liao, X. Tao, R. Li, Z. Ma and J. Jia, "Video Super-Resolution via Deep Draft-Ensemble Learning," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 531-539.

[4] M. Patil and S. D. Ruikar, "Super-resolution of face image extracted from a video sequence," 2014 International Conference on Communication and Signal Processing, Melmaruvathur, 2014, pp. 1620-1624.

[5] L. He, J. Tan, C. Xie and M. Hu, "A Novel Two-Step Approach for the Super-resolution Reconstruction of Video Sequences," 2014 5th International Conference on Digital Home, Guangzhou, 2014, pp. 85-90.

[6] S. Gohshi and I. Echizen, "Limitations of super resolution image reconstruction and how to overcome them for a single image," 2013 International Conference on Signal Processing and Multimedia Applications (SIGMAP), Reykjavik, 2013, pp. 71-78.

[7] J. Zhang, Y. Cao and Z. Wang, "A simultaneous method for 3D video super-resolution and high-quality depth estimation," 2013 IEEE International Conference on Image Processing, Melbourne, VIC, 2013, pp. 1346-1350.

[8] E. Faramarzi, D. Rajan, and M. P. Christensen, “Unified blind method for multi-image super-resolution and single/multi-image blur deconvo- lution,” IEEE Trans. Image Process., vol. 22, no. 6, pp. 2101–2114, Jun. 2013.

[9] E. Faramarzi, V. R. Bhakta, D. Rajan, and M. P. Christensen, “Super resolution results in PANOPTES, an adaptive multi-aperture folded architecture,” in Proc. 17th IEEE Int. Conf. Image Process. (ICIP), Sep. 2010, pp. 2833–2836.

[10] S. Borman and R. L. Stevenson, “Spatial resolution enhancement of low- resolution image sequences: A comprehensive review with directions for future research,” Dept. Elect. Eng., Univ. Notre Dame, Notre Dame, IN, USA, Tech. Rep., Jul. 1998.

[11] S. Borman and R. L. Stevenson, “Super-resolution from image sequences A review,” in Proc. Midwest Symp. Circuits Syst., Notre Dame, IN, USA, Aug. 1998, pp. 374–378.

[12] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical overview,” IEEE Signal Process. Mag., vol. 20, no. 3, pp. 21–36, May 2003.

[13] R. R. Schultz, L. Meng, and R. L. Stevenson, “Subpixel motion estimation for super-resolution image sequence enhancement,” J. Vis. Commun. Image Represent., vol. 9, no. 1, pp. 38–50, Mar. 1998.

[14] A. M. Tekalp, Digital Video Processing (Prentice Hall Signal Processing Series). Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.

[15] Y. Caspi and M. Irani, “Spatio-temporal alignment of sequences,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 11, pp. 1409–1424, Nov. 2002.

[16] O. Shahar, A. Faktor, and M. Irani, “Space-time super-resolution from a single video,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2011, pp. 3353–3360.