A MULTI-FRAME SUPER-RESOLUTION IMAGE RECONSTRUCTION USING REGULARIZATION FRAMEWORK

BY

MAHMOUD MOUSTAFA ABDELRAHMAN KHATTAB

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy (Computer Science)

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

MAY 2022

ABSTRACT

The global world experiences an enormous advancement in software and hardware technologies within the past decade. One of the primary measurements of image quality is image resolution. High-resolution is generally required and preferred for producing more detailed information inside the digital images; therefore, this leads to improve the pictorial information for human analysis and interpretation and to enhance the automatic machine perception. Unfortunately, the best use of image sensors and optical technologies is usually a high-priced method and is also constrained to increase the image resolution. Therefore, the effective use of image processing techniques for acquiring a high-resolution image generated from low-resolution images is an inexpensive and a powerful solution, which is called multi-frame super-resolution image reconstruction. However, the real imaging systems may introduce some degradation or artifacts in the digital images. These distortions in the images are caused by a variety of factors such as blurring, aliasing, and noise, which may affect the resolution of imaging systems and produce low-resolution images. Numerous strategies like frequency and spatial domain approaches have been proposed in the literature. Spatial domain approaches are classified as one of the most popular approaches and split into interpolation-based approaches and regularization-based approaches. Nevertheless, these techniques still suffer from artifacts. Regularization-based approaches are a challenging in image super-resolution in the last decade. This research intends to enhance the efficiency of multi-frame super-resolution image reconstruction in order to optimize both analysis and human interpretation processes by improving the pictorial information and enhancing the automatic machine perception. As a result, this research proposes new approaches for the image reconstruction of multi-frame super-resolution, so that they are created through the use of the regularization framework. On one hand, an efficient proposed approach is derived from the hybrid of reconstruction models in the image reconstruction stage based on employing adaptive norm and L_p norm in the data-fidelity term, beside adopting bilateral edge preserving and bilateral total variation prior models in the regularization term respectively. On the other hand, an efficient initialization approach is based on estimating the initial high resolution image through the pre-processing stage on the reference low resolution image. The proposed initialization approaches use linear and nonlinear filters including median, mean, Lucy-Richardson, and Wiener filters at the reference low resolution image. The proposed approaches are used to approximate a high-resolution image generated from a sequence of corresponding images with low-resolution to protect significant features of an image such as sharp image edges and texture information while preventing artifacts. In addition, these proposed approaches generate a high-quality image that is used in realistic applications with edges preservation and noise suppression. The experimental results with synthetic data indicate that the proposed approaches have enhanced efficiency visually and quantitatively compared to other existing approaches.

إعادة بناء صورة فائقة الدقة متعددة الإطارات باستخدام الإطار التنظيمي

خلاصة البحث

يَشهد العالم العالمي تقدماً هائلاً في تقنيات البرمجيات والأجهزة خلال العقد الماضي. تُعد دقة الصورة أحد القياسات الأساسية لجودة الصورة. يُفضل استخدام الدقة العالية بشكل عام لإنتاج معلومات أكثر تفصيلاً داخل الصور الرقمية، مما يُؤدي إلى تحسين المعلومات التصويرية لتحليل الإنسان وتفسيره وتعزيز الإدراك التلقائي للآلة. لسوء الحظ، فإن استخدام مستشعرات الصور والتقنيات الضوئية عادةً ما تكون أسلوباً باهظ الثمن ومقيد لزيادة دقة الصورة. لذلك، يُعد الاستخدام الفعال لتقنيات معالجة الصور حلاً غير مكلف وكفء للحصول على صورة عالية الدقة والتي تم إنشاؤها عن طريق مجموعة من الصور منخفضة الدقة، وهو ما يسمى إعادة بناء الصورة فائقة الدقة متعددة الإطارات. ومع ذلك، قد تُؤدي أنظمة التصوير الحقيقية إلى حدوث بعض التدهور أو التشوهات في الصور الرقمية. تحدث هذه التشوهات في الصور بسبب مجموعة متنوعة من العوامل مثل الضبابية والتشويش والضوضاء، والتي قد تُؤثر على دقة أنظمة التصوير وتنتج صوراً منخفضة الدقة. تم اقتراح العديد من الاستراتيجيات في الأدبيات مثل أساليب التردد والمجال المكاني. تُصنف أساليب المجال المكاني كواحدة من أكثر الأساليب شيوعاً، وتنقسم إلى أساليب قائمة على الاستيفاء (الإدماج) وأساليب قائمة على التنظيم. ومع ذلك، لا تزال هذه التقنيات تعاني من بعض التشوهات. بينما في العقد الماضي، تُمثل الأساليب القائمة على التنظيم تحدياً كبيراً في دقة الصورة الفائقة. في هذه الأطروحة، نعتزم تعزيز كفاءة إعادة بناء الصورة فائقة الدقة متعددة الإطارات من أجل تحسين عمليات التحليل والتفسير البشري من خلال تحسين المعلومات التصويرية وتعزيز الإدراك التلقائي للآلة. نتيجة لذلك، يقترح هذا البحث طرقًا جديدة لإعادة بناء الصورة ذات الدقة الفائقة متعددة الإطارات، بحيث يتم إنشاؤها من خلال استخدام إطار التنظيم. من ناحية، يتم اشتقاق طريقة مقترحة فعالة من نماذج مختلطة لإعادة الإعمار في مرحلة إعادة بناء الصورة استنادًا إلى استخدام معيار تكيفي ومعيار L_p في مصطلح دقة البيانات، إلى جانب استخدام نماذج المحافظة على الحافة الثنائية والتباين الكلي الثنائي في مصطلح التنظيم على التوالي. من ناحية أخرى، تُستخدم طريقة التهيئة الفعالة لتقدير الصورة الأولية عالية الدقة من خلال مرحلة المعالجة المسبقة على الصورة المرجعية منخفضة الدقة. تُستخدم طرق التهيئة المقترحة المرشحات الخطية وغير الخطية مثل مرشحات الوسيط والمتوسط ولوسي ريتشاردسون ووينر على الصورة المرجعية منخفضة الدقة. تُستخدم الطرق المقترحة لتوليد صورة عالية الدقة يتم إنشاؤها من سلسلة من الصور ذات الدقة المنخفضة لحماية الميزات المهمة للصورة مثل حواف الصورة الحادة ومعلومات الملمس ومنع التشوهات. بالإضافة إلى ذلك، تُولد هذه الطرق المقترحة صورة عالية الجودة تُستخدم في تطبيقات واقعية مع الحفاظ على الحواف وقمع الضوضاء. تُشير النتائج التجريبية إلى أن الطرق المقترحة قد عززت الكفاءة بصرياً وكمياً بالمقارنة مع الأساليب الأخرى.

APPROVAL PAGE

The thesis of Mahmoud Moustafa Abdelrahman Khattab has been approved by the following:

Akram Mohammed Zeki Khedher
Supervisor

Ali A. Alwan Aljuboori
Co-Supervisor

Amelia Ritahani Bt. Ismail
Internal Examiner

Rusli bin Haji Abdullah
External Examiner

Meftah Hrairi
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Mahmoud Moustafa Abdelrahman Khattab

Signature ____________ Date 05/05/2022

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

A MULTI-FRAME SUPER-RESOLUTION IMAGE RECONSTRUCTION USING REGULARIZATION FRAMEWORK

I declare that the copyright holder of this thesis is jointly owned by the student and IIUM.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Mahmoud Moustafa Abdelrahman Khattab.

[signature]	05/05/2022
Signature	Date

It is my utmost pleasure to dedicate this thesis to my late parents (may God have mercy on them, forgive them, and make them spacious in paradise), who granted me the gift of their unwavering belief in my ability to accomplish this goal.

ACKNOWLEDGEMENTS

Firstly, all glory is due to Allah, the Almighty, whose Grace and Mercies have been with me throughout the duration of my programme. Although, it has been tasking, His Mercies and Blessings on me ease the herculean task of completing this thesis.

I am most indebted to by supervisor, Prof. Dr. Akram Mohammed Zeki Khedher, whose enduring disposition, kindness, promptitude, thoroughness and friendship have facilitated the successful completion of my work. I put on record and appreciate his detailed comments, useful suggestions and inspiring queries which have considerably improved this thesis. His brilliant grasp of the aim and content of this work led to his insightful comments, suggestions and queries which helped me a great deal. Despite his commitments, he took time to listen and attend to me whenever requested. The moral support he extended to me is in no doubt a boost that helped in building and writing the draft of this research work. I am also grateful to my co-supervisor, Assoc. Prof. Dr. Ali A. Alwan Aljuboori, whose support and cooperation contributed to the outcome of this work.

I would also like to take this opportunity to thank my viva examiners, for their very helpful comments, suggestions, and constructive criticisms, in addition for letting my defense be an enjoyable moment.

My gratitude extends to the Kulliyyah of Information and Communication Technology (KICT), International Islamic University Malaysia (IIUM), for giving me the opportunity to carry out my doctoral research and undertake my studies to the fullest. In addition, my personal thanks go to all lecturers and staff members at the Kulliyyah as well as the university for all their considerate guidance and whose wisdom have shaped my way of looking at things one way or another, because they have the ability for creativity, dedication, and hard work.

Finally, my gratitude goes to my beloved wife and lovely children; for their prayers, understanding, and endurance while away. Furthermore, I must acknowledge my work to my late parents, without their support, concern, and love, it was impossible for me to complete my Ph.D. studies.

Once again, we glorify Allah for His endless mercy on us one of which is enabling us to successfully round off the efforts of writing this thesis. Alhamdulillah.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADMM	Alternating Direction Method of Multipliers
ADSF	Anisotropic Diffusion Shock Filter
AM	Alternating Minimization
ATV-SR	Adaptive Total Variation Super-Resolution
BEP	Bilateral Edge Preserving
BEP-BTV	Bilateral Edge Preserving - Bilateral Total Variation
BTV	Bilateral Total Variation
CCD	Charge Coupled Device
CG	Conjugate Gradient
CMOS	Complementary Metal Oxide Semiconductor
CT	Computer Tomography
CWP	Channel Weight Parameters
dB	Decibel
DCT	Discrete Cosines Transform
DDE	Direct Difference Error
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
EGI	Edge Guided Interpolation
EPS	Edge Preserving Smoothing
GF	Gradient Field
GIF	Guided Image Filter
GVF	Gradient Vector Flow
GVFHF	Gradient Vector Flow Hybrid Field
HFET	High Frequency Energy Term
HMRF	Huber Markov Random Field

HR	High Resolution
HVS	Human Visual System
IBP	Iterative Back Projection
IRLS	Iteratively Reweighted Least Squares
K-L	Kullback–Leibler
LARSR	Local Adaptive Regularized Super Resolution
LF	Light Field
LOR	Lorentzian
LR	Low Resolution
LRTV	Low Rank Total Variation
LWAR	Locally Weighted Anisotropy Regularization
MAP	Maximum A Posteriori
MAP-MRF	Maximum A Posteriori Markov Random Field
MBD	Multichannel Blind Deconvolution
ML	Maximum Likelihood
MLRW-BTV	Median, Lucy-Richardson, Wiener - Bilateral Total Variation
MM	Majorization Minimization
MMW-BTV	Median, Mean, Wiener - Bilateral Total Variation
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NCCR	Normalized Cross Correlation Ratio
NEDI	New Edge Directed Interpolation
NL	Non Local
NLTV	Non Local Total Variation
PCA	Principal Component Analysis
PDA	Personal Digital Assistant
POCS	Projection Onto Convex Set

PSF	Point Spread Function
PSNR	Peak Signal to Noise Ratio
PSO	Particle Swarm Optimization
RAWC	Regional Adaptive Weight Coefficients
RMSE	Root Mean Square Error
ROI	Region of Interest
RWP	Residual Weight Parameters
SAR	Simultaneous Autoregressive
SR	Super Resolution
SRT	Structural Regularization Term
SSIM	Structural Similarity
SWDTV	Stroke Width-Based Directional Total Variation
SWTV	Spatial Weighted Total Variation
TV	Total Variation

CHAPTER ONE

INTRODUCTION

1.1 RESEARCH BACKGROUND

Over the last two decades, the world has experienced an enormous advancement in software and hardware technologies. Industrial sectors have made the best use of modern technology to generate electronic devices such as computer systems, cellular mobile phones, personal digital assistant (PDA), and innumerable devices at inexpensive costs (Yue et al., 2016). Moreover, the manufacturing methods of camera sensors have been highly developed to generate high-quality digital cameras. Many applications of computer vision such as medical imaging, satellite imaging, pattern recognition, surveillance and forensic, astronomical imaging, and target detection are still in urgent need of high-resolution (HR) image which frequently exceeds the abilities of the HR digital cameras (Park, Park, & Kang, 2003; Hou, Wang, & Wang, 2016).

In many real-life imaging systems, there are some artifacts such as blurring, aliasing, and noise that affect the image resolution as shown in Figure 1.1 (M. Kumar & Diwakar, 2016; Yue et al., 2016; Huangpeng et al., 2017). The blurring effects can appear within the image during the shooting process based on some factors such as scene movements, incorrect focusing, atmospheric confusion, and optical point spread function. Accordingly, it is much easier to remove the blur effects from the image accurately, if the shooting conditions at the time of getting the image are known. Also, noise can be caused by a wide range of factors such as differences in detector sensitivity, visual defects, and environmental changes. There is no relationship between the pixels and the noise because the noise is not spatially connected to the image. Also, down-sampling is a result of an inadequate spatial sampling that led to overlapping between high and low-frequency components (Park et al., 2003; Begin & Ferrie, 2006; Macwan, Patel, Prajapati, & Chavda, 2014; Yue et al., 2016; Huangpeng et al., 2017). These distortions in images may affect the resolution of imaging systems and produce low-resolution (LR) images in which the pixel density within the image is small. Accordingly, LR images have a serious shortage of the stored information, which influences the quality of human interpretation and automatic machine perception.

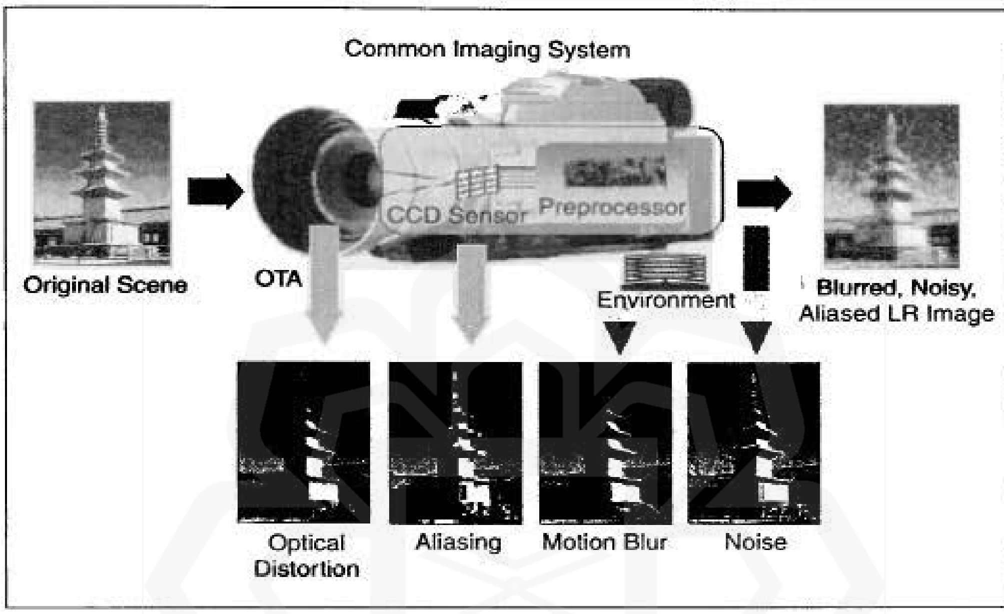

Figure 1.1 The common image acquisition system (Macwan et al., 2014)

Optical resolution is certainly a method of measuring the capability of the camera system or an element of the camera system utilized for explaining the image details. Accordingly, there are two primary methods of raising the spatial image resolution: firstly, the technical strategy method which is related to hardware solutions. Secondly, the analytical strategy method which is associated with software solutions (Yue et al., 2016). With respect to the technical strategy concept, it identifies either the improvement of the registration device or the replacement with a higher resolution device. However, the use of a highly qualified camera is often limited by its high price, large size, or sensor manufacturing limitations (C. Kumar, 2011). Concerning the analytical strategy concept, it is usually low-priced and more flexible in comparison to the hardware solutions. The class of resolution improvement methods has been named super-resolution (SR) image reconstruction (Park et al., 2003; Patanavijit, 2009; Protter, Elad, Takeda, & Milanfar, 2009; Yue et al., 2016).

On one hand, SR image reconstruction usually represents a great enhancing and challenging method of digital imaging. The reason is that it attempts to rebuild HR images by combining the partial information presented inside several LR images of a particular scene through the image reconstruction process as shown in Figure 1.2 (Long et al., 2017). On the other hand, SR incorporates up-sampling of LR images. Then, it eliminates distortions such as noise and blurring. In comparison to different image improvement techniques, SR not only increases the quality of LR images by improving their particular spatial resolution but also tries to eliminate distortions (Park et al., 2003; Yue et al., 2016).

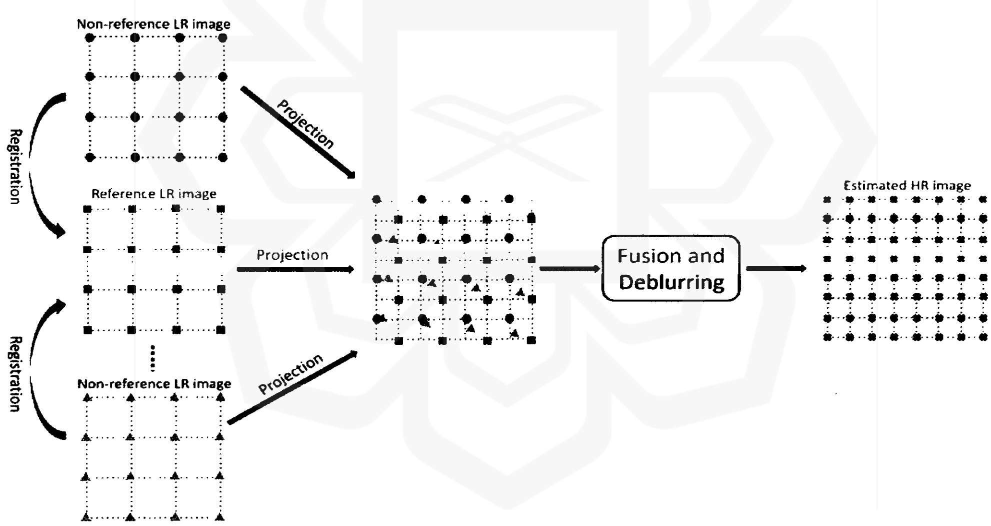

Figure 1.2 The concept of multi-frame SR (Long et al., 2017)

This research intends to enhance the efficiency of multi-frame SR image reconstruction in order to optimize both analysis and human interpretation processes by improving the pictorial information and enhancing the automatic machine perception. As a result, the main aim of this research work is to develop new SR image reconstruction approaches capable of generating an HR image to be used in the real-life contemporary applications. The proposed approaches are used to preserve the important image characteristics as much as possible including borders and corners, preserve the sharp image edges, and overcome artifacts. Furthermore, they increase the high-frequency components, eliminate degradations in the image capturing systems, and achieve a good balance for edge protection and noise reduction.

1.2 PROBLEM STATEMENT

In many visual applications, the HR image is generally required and preferred for producing more detailed information inside the digital images, therefore, this leads to the improvement of pictorial information for human analysis and interpretation and also for automatic machine perception (Köhler et al., 2016; Huang et al., 2018).

Unfortunately, the real imaging systems may introduce several degradations or artifacts in digital images. These distortions in images are caused due to a variety of reasons such as blurring, aliasing, and noise, which may affect the resolution of imaging systems and produce LR images. Accordingly, LR images have a serious shortage of the stored information, which influences the quality of human interpretation and automatic machine perception. The best utilization of image sensors and optical technologies to increase the image pixel density is usually restrictive and overpriced. Therefore, the effective use of image processing techniques for acquiring a high-resolution image generated from low-resolution images is an inexpensive and powerful solution. As a result of these reasons, many researchers develop various methods for producing a high-quality image based on SR image reconstruction approaches (Laghrib, Ghazdali, Hakim, & Raghay, 2016; Nayak & Patra, 2016; Shen, Peng, Yue, Yuan, & Zhang, 2016; Huang et al., 2018; Laghrib, Hadri, Hakim, & Raghay, 2019). However, most of the SR approaches that have been developed in the literature remain suffering from an imbalance between the edges preservation and the noise suppression inside the reconstructed HR image (L. Wang, Lin, Deng, & An, 2017a; Hakim, Ghazdali, & Laghrib, 2020). In which, if the noise is completely eliminated from the reconstructed HR image, this leads to smoothness in the edges. On the other side, if the edges are preserved well in the reconstructed HR image, this leads to high image noise (El Mourabit et al., 2017; Kiani & Drummond, 2017; Long et al., 2017; Mohan, 2017; L. Wang et al., 2017a; Huang et al., 2018; Laghrib et al., 2019).

Thus, to resolve this issue, an efficient image reconstruction approach is needed to protect image texture information while eliminating noise and improving the edges of the estimated image from a series of corresponding LR images. The approach should take into consideration increasing the high frequency components and eliminating the degradations generated through the imaging systems, and achieving a good balance between preserving the edges and suppressing the noise.

1.3 RESEARCH QUESTIONS

This section describes the research questions to be answered in this research work. The research questions are as follows:

1. What are the most crucial causes that affect the resolution of the images?
2. What are the current SR approaches that are commonly used to construct the HR image?
3. Why are existing approaches inefficient to reconstruct the HR image?
4. How to treat LR images to be HR images?
5. How to verify the efficiency and the quality of the reconstructed HR image?

1.4 RESEARCH OBJECTIVES

The main aim of this research work is to develop a new SR image reconstruction approaches capable of generating an HR image for the real-life contemporary applications. Therefore, to achieve this aim, the following objectives are set:

1. To identify the main factors that affect the resolution of the images.

2. To investigate the various current well-known SR approaches which are commonly and efficiently used for constructing the HR image.
3. To identify the main reasons for the ineffectiveness of existing approaches in reconstructing the HR image.
4. To propose efficient SR approaches that are capable of treating the unexpected artifacts, restoring the missing high frequency details, and achieving a reasonable balance between preserving the edges and suppressing the noise in the reconstructed HR image.
5. To evaluate the efficiency and the effectiveness of the proposed SR approaches compared to other existing approaches taking into consideration the visual efficiency and the peak signal to noise ratio (PSNR) and structural similarity (SSIM) values quantitatively.

1.5 RESEARCH SIGNIFICANCE

As explained in the aforementioned section, the process of reconstructing HR images is one of the hottest research areas in recent years in which a wide range of useful details are acquired from images. SR approaches are used in different domains to analyze and extract the essential information from the images (Yue et al., 2016). SR technologies are used in a wide range of applications to achieve the HR image and maybe distinct in different applications (Köhler et al., 2016; Hakim et al., 2020). The HR image is generally required and preferred for producing more detailed information inside the digital images, therefore, this leads to the improvement of pictorial information for human analysis and interpretation and also for automatic machine perception.

SR images are widely used in a variety of applications and it can be highly beneficial if multiple images with the same scene are simply produced. Many applications of computer vision such as medical imaging, satellite imaging, pattern recognition, surveillance and forensic, astronomical imaging, and target detection are still in urgent need for HR images due to: