DISSERTATIONS IN
FORESTRY AND
NATURAL SCIENCES

Zijian Zhu

High Quality High Dynamic Range Imaging

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences

UNIVERSITY OF
EASTERN FINLAND

ZIJIAN ZHU

High Quality High Dynamic Range Imaging

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences
No 113

Academic Dissertation
To be presented by permission of the Faculty of Science and Forestry for public examination in the Louhela auditorium at the University of Eastern Finland, Joensuu, on August, 7, 2013, at 12 o'clock noon.

School of Computing

Kopijyvä Oy
Joensuu, 2013
Editor: Prof. Pertti Pasanen

Distribution:
Eastern Finland University Library / Sales of publications
P.O.Box 107, FI-80101 Joensuu, Finland
tel. +358-50-3058396
http://www.uef.fi/kirjasto

ISBN: 978-952-61-1171-1 (Printed)
ISSNL: 1798-5668
ISSN: 1798-5668
ISBN: 978-952-61-1172-8 (PDF)
ISSN: 1798-5676 (PDF)

Author's address: University of Eastern Finland
School of Computing
P.O.Box 111
80101 JOENSUU
FINLAND
email: zhuzj@i2r.a-star.edu.sg

Supervisors: Professor Pasi Fränti, Ph.D.
University of Eastern Finland
School of Computing
P.O.Box 111
80101 JOENSUU
FINLAND
email: pasi.franti@uef.fi

Professor Susanto Rahardja, Ph.D.
Institute for Infocomm Research
Department of Signal Processing
1 Fusionopolis Way, #21-01 Connexis
138632 SINGAPORE
SINGAPORE
email: rsusanto@i2r.a-star.edu.sg

Reviewers: Professor Rafal Mantiuk, Ph.D
Bangor University
School of Computer Science
Dean Street, BANGOR, LL57 1UT
UNITED KINGDOM
email: mantiuk@bangor.ac.uk

Dr. Radu Ciprian Bilcu, Ph.D
Nokia Research Center
Multimodal Sensing and Context
Visiokatu 1
33720 TAMPERE
FINLAND
email: radu.bilcu@nokia.com

Opponent: Dr. Alessandro Foi, Ph.D
Academy of Finland Research Fellow
Department of Signal Processing
Tampere University of Technology
33101 TAMPERE
FINLAND
email: alessandro.foi@tut.fi

ABSTRACT

A *high dynamic range* (HDR) image uses a large bit depth up to 32-bit per pixel per color channel. Due to hardware limitation, it can neither be captured by conventional camera in a single photo, nor be displayed on a conventional monitor. In this thesis, a software solution of HDR imaging is introduced from the synthesis to the display of high quality HDR images.

Contrary to conventional imaging method, HDR imaging uses multiple exposures. Therefore, the variance of the same pixel in different images is non-linear and difficult to measure. To solve this, a new *inter-pixel relationship function* (IRF) is proposed using both spatial and temporal correlations. It is widely used in many HDR imaging components.

Multiple input images produce a new ghosting artifact due to moving object. A real time de-ghosting method is first proposed using bi-directional comparison and IRF based moving object detection and patching. It is lightweight in terms of both time complexity and physical memory consumption, which makes it suitable for mobile devices. We further extend it by merging the IRF with a histogram intensity mapping and adopting a new threshold model based on statistical study. The extended de-ghosting scheme is more robust in the scenarios where the moving object occupies larger areas.

Noise in the input images will be kept in the synthesized HDR image. A 2D de-noising factor is proposed for the synthesis of 32-bit HDR image, and a noise reduced tone-mapping is proposed to map the 32-bit HDR image back to 8-bit for the display based on *information content weighting* (ICW). Both methods have low time complexity.

It is also important to measure if the input images are from the same scene, as huge artifacts may appear if a wrong image is involved in the synthesis process. We proposed a structure similarly based metric which is robust to the images with different exposures. We also proposed a metric for the measurement of two 32-bit HDR images based on their histograms of radiance maps.

AMS Classification: 68U10; 94A08
Universal Decimal Classification: 004.932
Library of Congress Subject Headings: High dynamic range imaging; Computational photography; Image enhancement - Methods; Image systems - Image quality

Yleinen suomalainen asiasanasto: digitaalikuvaus; kuvankäsittely

Acknowledgements

The efforts of my PhD studies span from Singapore to Finland crossing more than 9000 kilometers. It would not have been possible without the kind support and help of many people. And I would like to extend my sincere thanks to all of them.

I would like to express my great appreciation to Prof. Pasi Fränti, my supervisor, for his support and guidance over the years. His willingness to give his time and constructive suggestions so generously has been very much appreciated.

I would like to express my deep gratitude to Prof. Susanto Rahardja, my co-supervisor, for his strategic suggestions during the planning and development of my research work. He is also the one who provided me a connection from Singapore to Finland.

I would like to thank Prof. Mantiuk and Dr. Bilcu, the reviewers of the thesis, for their helpful comments, and Dr. Foi, my opponent, for the valuable time.

Besides, I wish to thank my colleagues in Image Processing Group in Institute for Infocomm Research for co-authoring the papers. Particularly, I would like to thank to Dr. Zhengguo Li for his great suggestions, inspirations and critiques. Nevertheless, my grateful thanks are extended to the Speech and Image Processing Unit in the University of Eastern Finland for their help in offering me the resources during my trips in Finland.

Thanks and appreciation to the support from the University of Eastern Finland, Institute for Infocomm Research and Nokia Foundation.

Last but not least, I would like to express my special thanks to my family for their understanding and support.

Singapore Jun 25th 2013 Zijian Zhu

LIST OF ABBREVIATIONS

CCD	Charge-Coupled Device
CRF	Camera Response Function
CMOS	Complementary Metal–Oxide–Semiconductor
DR	Dynamic Range
EV	Exposure Value
HDR	High Dynamic Range
HVS	Human Visual System
ICW	Information Content Weighting
IFC	Information Fidelity Criterion
IMF	Intensity Mapping Function
IRF	Inter-pixel Relationship Function
ISO	Light Sensitivity Setting
LDR	Low Dynamic Range
PSNR	Peak Signal-to-Noise Ratio
SNR	Signal-to-Noise Ratio
WLS	Weighted Least Squares

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following selection of the author's publications in the field of high dynamic range imaging:

[P1] **Z. J. Zhu**, Z. G. Li, S. Rahardja, and P. Fränti, "Recover real world scene: high quality image inpainting using multi-exposed references", *Electronics Letters*, pp. 1310-1312, Vol. 45, Issue 25, Dec 2009.

[P2] Z. G. Li, **Z. J. Zhu**, and S. Rahardja, "Structural similarity indices for high dynamic range imaging", *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1145-1148, Prague, Czech Republic, May 2011.

[P3] **Z. J. Zhu**, S. Rahardja, Z. G. Li, P. Fränti, "Detecting and composing near-identical HDR images without exposure information", *IEEE International Conference on Image Processing*, pp. 3337-3340, Hong Kong, Sep 2010.

[P4] **Z. J. Zhu**, S. Rahardja, S.Q. Wu, and P. Fränti, "Real-time ghost removal for composing high dynamic range images", *IEEE Conference on Industrial Electronics and Applications*, pp. 1627-1631, Taichung, Taiwan, Jun 2010.

[P5] **Z. J. Zhu**, Z. G. Li, S. Rahardja, and P. Fränti, "De-ghosting of HDR images with double-credit intensity mapping", *IEEE International Conference on Image Processing*, pp. 1361-1364, Brussels, Belgium, Sep 2011.

[P6] Z. G. Li, **Z. J. Zhu**, and S. Rahardja, "Fast movement detection for high dynamic range imaging", *IEEE International Conference on Image Processing*, pp. 373-376, Brussels, Belgium, Sep 2011.

[P7] **Z. J. Zhu,** Z. G. Li, S. Q. Wu, J. H. Zheng, W. Yao, P. Fränti, and S. Rahardja, "Noise reduced high dynamic range tone-mapping based on information content weighting", Manuscript

[P8] **Z. J. Zhu,** Z. G. Li, S. Rahardja, and P. Fränti, "2D denoising factor for high dynamic range imaging", ACM SIGGRAPH Posters, pp. 46:1-46:1, Los Angeles, US, 2012.

Throughout the overview, these papers will be referred to as [P1]-[P8]. The papers have been included in this thesis with permission of their copyright holders.

AUTHOR'S CONTRIBUTION

In [P1, P3-P5], the ideas, implementation and experiments are original from the author. The papers were written by the author too. In [P2], author contributed the idea of how to use the intensity mapping for balancing the different dynamic range. In [P6], author contributed to the idea of how to improve the speed for moving object detection, as well as the C implementation on both PC and mobile phone. The HDR de-noising [P7-P8] was inspired by Dr. Li Zhengguo's early work on weighted guided filter and weighted frame average, while the author proposed the methods to solve the problems. At the same time, author was responsible for the implementation, experiments and composing the papers.

Contents

1 Introduction

When a scene with high light contrast is captured by a camera, either the dark area or the bright area will be saturated in the output image, as shown in Fig. 1.1. This is due to the limitation of the camera sensor, and has been existed since the first camera was invented. To solve this, *high dynamic range* (HDR) imaging has been invented by recovering the real world scene using multiple conventional *low dynamic range* (LDR) images [1, 2, 3, 4, 5, 6].

Fig. 1.1 Problem in capturing high dynamic range scene, where (Top) is the real world scene perceived by human eyes, and (Bottom) are the images captured by camera using different exposure times (Δt). Both over-exposed saturation (red square) and under-exposed saturation (blue square) are present, even in the 'best exposed' image (Bottom Middle). The 'real world scene' image was synthesized through HDR imaging by using the bottom 3 images.

1.1 DYNAMIC RANGE

The *dynamic range* (DR) of a scene is defined as the range from the lowest light intensity to the highest light intensity in the scene. It is named as *scene contrast* too. In real world, the full range of light intensities that can be perceived by human vision system spans from as low as star light to as high as the sun shine. The human vision system is capable of perceiving light over a range of 5 orders of magnitude simultaneously. On top of that, the pupil adapts very fast to the changes in lighting conditions, which result to a concurrently perception of nearly 10 orders of magnitude [1, 7]. In comparison to this biological sensor, modern cameras can only capture a small dynamic range of about two orders of magnitude in a single exposure, as shown in Fig. 1.2. Thus, information is lost in the very bright and very dark areas, named *saturated* areas.

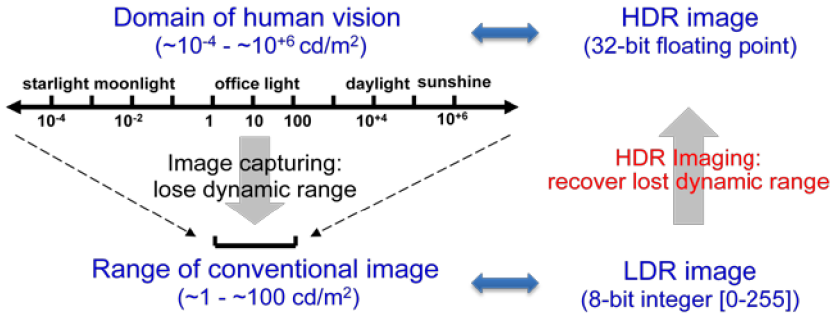

Fig. 1.2 The domain of human vision compare to the range of conventional LDR images. The image capturing process lose information due to sensor limitation, while HDR imaging recover the lost dynamic range through reverse engineering by merging multiple shots with different exposures.

1.2 HIGH DYNAMIC RANGE ACQUISITION

In order to recover the real world scene, various HDR synthesis methods have been proposed with different acquisition methods, as listed in Table 1.1.

It is possible to capture an HDR image using a single shot with the modification of camera hardware. Capturing image

gradients rather than actual pixel intensities was shown to increase the dynamic range at the cost of computationally expensive *Poisson solver* [8]. As far as I know, the gradient camera is still a theoretical solution. In *assorted pixels* method, multiple neighboring pixels with fixed pattern filter are used to capture at different exposures [9, 10]. Each pixel in the final HDR image was reconstructed by 2-4 neighboring pixels, which results a loss of camera resolution. To overcome this problem, an alternative design uses an aligned spatial light modulator with programming imaging [11, 12]. However, such a design is difficult to implement with extra hardware expenses.

Table 1.1 Typical acquisition approaches for capturing high dynamic range images.

Acquisition Methods	**HDR Synthesis Methods**
Single image capture	Gradient camera Assorted pixels Dynamic range adaptive imaging
Single device multiple image capture	Exposure bracketing (HDR video) Generalized mosaics
Multi-device capture	Split aperture imaging Optical splitting trees

The second acquisition approach uses single camera with multiple shots. Among all the methods, exposure bracketing [1, 2, 13, 14, 15, 16] is most widely used as no hardware modification is required. Within a short period of time, multiple shots with different exposures are captured to minimize the change of the lighting condition as well as object movement. Then, intensity mapping is estimated using image correlation and an HDR radiance map is recovered by using the best exposure information in different input images. Similar approach was adopted to generate HDR video too [17, 18, 19, 20]. In addition to exposure bracketing, a static filter with varying transitivity, named *generalized mosaicing* [21], was proposed to be mount in front of the camera sensor to mimic the exposure bracketing effect.

Third acquisition approach uses multiple devices to capture the same scene with the help of prisms [22] or beam-splitters [23,

24, 25]. While these approaches allow dynamic content to be recorded, the additional optical elements split the emitting light and make it difficult to capture the low light scenes in their best exposures.

1.3 HIGH DYNAMIC RANGE DISPLAY

Similar to image capturing system, the same hardware limitation exists in display system, where the HDR contents cannot be displayed on the conventional LDR display devices, such as monitors, printers and projectors.

Nowadays, most HDR images are converted back into LDR images using HDR tone-mapping before being presented on conventional display devices. There are two major categories of HDR tone-mapping algorithms, global operators [26, 27, 28] and local operators [29, 30, 31, 32]. Global operators are optimized on the whole image without local adjustment based on neighboring information. It is fast with no intensity inversing, meaning the brighter pixels in the HDR image remains brighter or equal in the tone-mapping result. A good application is HDR video [33]. The tradeoffs of the global operators are the degradation of visual quality, such as lack of local contrast and lose of small details. In order to improve the visual quality, local operators use the neighboring information for every pixel to preserve the fine details. These operators are usually computationally expensive, and come with halo artifacts [29, 34] in some scenarios.

Besides software solution, research based on hardware solutions has been carried out to display the original HDR content. A dual modulation display system [35] was invented by adding dynamic backlight using light-emitting diode (LED) to the conventional liquid crystal display (LCD), where the LCD provides color, resolution and normal contrast, while the programmable LED provides additional rear-illumination. Similar ideas have also been applied to increase the dynamic range of printer [36] and microscopy [37]. Another approach

focused on HDR projection by reallocating the light inside a conventional projector using an analog micro-mirror array in the optical path [38]. The light reallocation does not only increase the dynamic range of the projector display, but also reduces the device heat due to dumpling excessive background light out instead of physically enclosed in the device.

1.4 WORKFLOW OF SOFTWARE SOLUTION OF HDR IMAGING

In this thesis, we focus on the software solution of HDR imaging. Therefore, exposure bracketing is used for HDR acquisition, and HDR tone-mapping is adopted for displaying. This solution is most widely used nowadays in consumer industry due to their capabilities of directly applying on existing hardware, which reduces the total cost.

As shown in Fig. 1.3, a typical software solution of HDR imaging is not only one or two, but series of technologies that transforms the captured LDR images (usually 8-bit per color channel) into HDR images (usually 32-bit per color channel), and then display them on existing conventional display devices. The modules in amber color in Fig. 1.3 are the research topics covered in this thesis. The modules in green color are technologies used in our software solution of HDR imaging system, which are not covered in this thesis.

Nevertheless, most consumer-level smart phones, such as iPhone 5, Samsung Galaxy S3, and HTC 1X[1], embedded HDR acquisition using exposure bracketing and display using *exposure fusion* [39, 40, 41, 42, 43, 44, 45], a combined fast process of HDR synthesis and HDR tone-mapping. We use exposure fusion for our mobile HDR solution, which is not covered in this thesis.

---

[1] Information and datasheet can be found at
iPhone 5: http://www.apple.com/iphone/
Galaxy S3: http://www.samsung.com/global/galaxys3/
1X: http://www.htc.com/www/smartphones/htc-one-x/

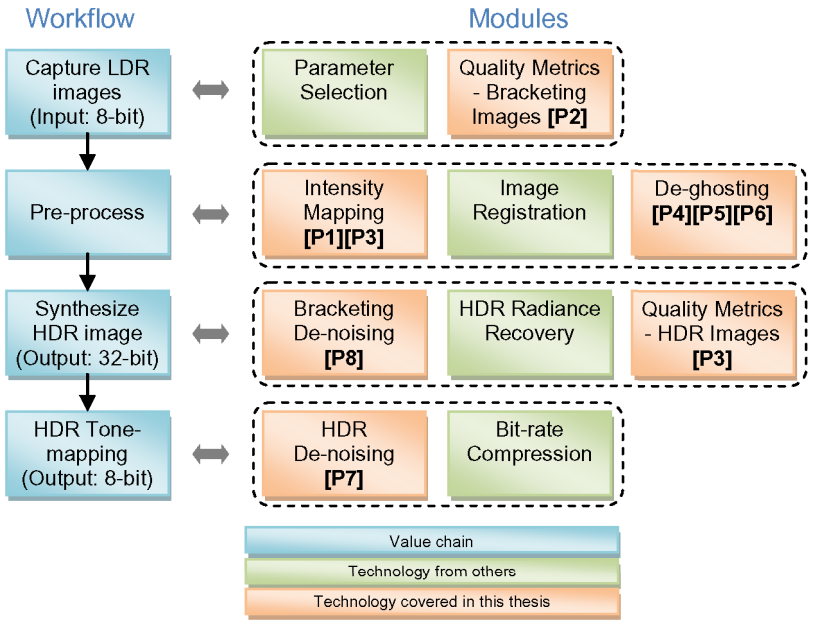

Fig. 1.3 A typical workflow of software solution for high dynamic range imaging.

1.5 STRUCTURE OF THE THESIS

The rest of the thesis is organised as follows. Intensity mapping is discussed in Chapter 2. HDR de-ghosting is studied in Chapter 3. Bracketing image de-noising and tone-mapping de-noising are summarized in Chapter 4. Image difference metrics for the exposure bracketing images and the HDR images are presented in Chapter 5. A summary of the contributions is given in Chapter 6 and conclusions drawn in Chapter 7.

2 *Intensity Mapping*

Exposure bracketing is a photography technique to capture the same scene with multiple shots using various exposure settings. It allows high dynamic range contents to be stored at different exposure levels [46, 47, 48, 49].

Fig. 2.1 A sequence of exposure bracketing images from the highest exposure (top left) to the lowest exposure (bottom right). The image of high exposure captured the detail of dark area with the saturation in the bright area, while the image of low exposure captured the bright area with the saturation in the dark area.

HDR images are usually synthesized by recovering the radiance map using the correlations among these bracketing images. Among all the correlations, intensity mapping is the most important feature defined as

$$\psi_{i,j}(Z_i(p)) = Z_j(p) \qquad (2.1)$$

where Z_i and Z_j are the intensity value of the same co-located pixel p at different exposures of i and j, and $\psi_{i,j}$ is the intensity mapping from image i to image j. It is noted that the intensity mapping is a directional function ($\psi_{i,j} \neq \psi_{j,i}$). The *intensity mapping function* (IMF) is not only hardware dependant, but also

will change according to different capturing settings. And in most cases, it is not a linear function.

The intensity mapping is the key to many HDR applications, such as image in-painting with exposure bracketing images, HDR image synthesis, and HDR de-ghosting.

2.1 BACKGROUND

When taking a photo, each pixel in the digitized output image will contain a pixel value known as *intensity*. However, this value is rarely true measurement of the real radiance of that pixel [14]. Instead, it is the output of a mapping function, known as *camera response function* (CRF) [1, 2, 50], which maps the real world radiance to a pixel value within the display range.

Intuitively, CRF is a good candidate for intensity mapping. It is a characteristic of the camera and the exposures and, thus, is not scene dependant and is commonly estimated from the exposure bracketing images. A popular CRF estimation is proposed in [2] by minimizing the quadratic objective function

$$\mathcal{O} = \sum_{p=1}^{P}\sum_{i=1}^{N}[g(Z_{i,p}) - lnE_p - ln\Delta t_i]^2 + \lambda \sum_{z=Z_{min}+1}^{z=Z_{max}-1} g^{"}(z)^2 \qquad (2.\ 2)$$

where g is the log reverse CRF function, Z_{ij} is the known pixel value at pixel location p (total P pixels) in the input image i (total N images), E denotes the radiance value, Δt denotes the exposure time, and λ represents a weight for the smoothness term. The CRF estimation is computationally expensive, and not very robust to noise, moving object or camera movement. Using CRF, the intensity mapping can be represented as

$$\psi_{i,j}(Z_i(p)) = g^{-1}(g(Z_i(p)) - ln\ \Delta t_i + ln\ \Delta t_j) \qquad (2.\ 3)$$

where the exposure time must be known in advance.

Another important approach uses a histogram based IMF[51], which determines the intensity mapping by using cumulative histogram. An IMF (ψ) is defined as

$$\psi_{i,j}(Z(p)) = H_j^{-1}(H_i(Z(p)) \qquad (2.\ 4)$$

where $H(z) = \int_0^z h(x)dx$ is the cumulative histogram of the original histogram h. Compared to the CRF-based approach, the histogram based intensity mapping does not require exposure time information, and does not require hard image alignment. A drawback is that it only makes use of the pixel statistic information, but not any spatial correlation.

2.2 INTER-PIXEL RELATIONSHIP FUNCTION

The *inter-pixel relationship function* (IRF) is proposed in **[P1]**, and the idea is illustrated in Fig. 2.2. If the intensity at A′ and B′ are the same at one exposure time in the reference, they shall be the same at the other exposure times (not for saturated pixels). This is according to photography reciprocity [52]: when exposure time changes, the pixel values change correspondingly. However, during the image capturing process, sensor noise, sampling noise and compression noise are commonly generated. Thus, it is more accurate to find all the pixels with the same intensity in the reference image and calculate their co-location values in the test image using mean average. The IRF is defined as

$$\psi_{i,j}(Z_i(p)) = \frac{\sum_{x \in \Omega(Z_i(p))} Z_j(x)}{|\Omega(Z_i(p))|} \qquad (2.\ 5)$$

where $|\Omega(Z_i(p))|$ is the cardinality of the spatially co-located pixels set $\Omega\big(Z_i(p)\big)$: $\{x|Z_i(x) = Z_i(p)\}$.

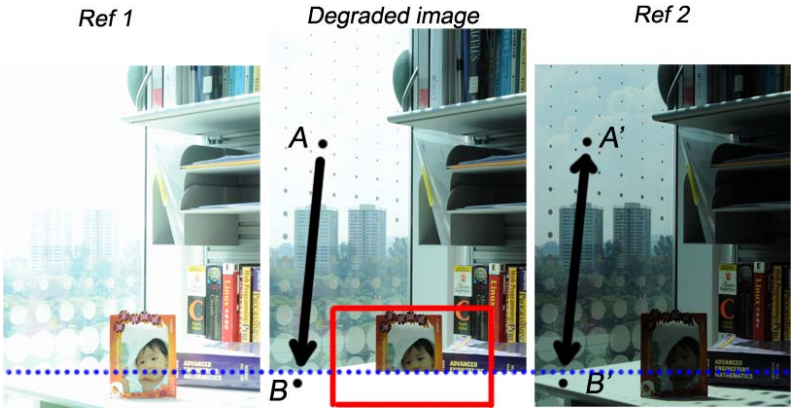

Fig. 2.2 Three differently exposed images with degraded image due to camera shake. The pixel value of B can be copied from A as their co-location A′ and B′ have the same intensity in the reference image.

The IRF has three useful characteristics inherited from the physical camera response. (1) The IRF is a monotonically increasing function. (2) The pixels located at left end (dark pixels), as shown in Fig. 2.3, and right end (bright pixels) are highly compressed due to dynamic range limit. (3) When choosing different reference images, shorter exposure time leads to smaller slope at the left end and bigger slope at the right end.

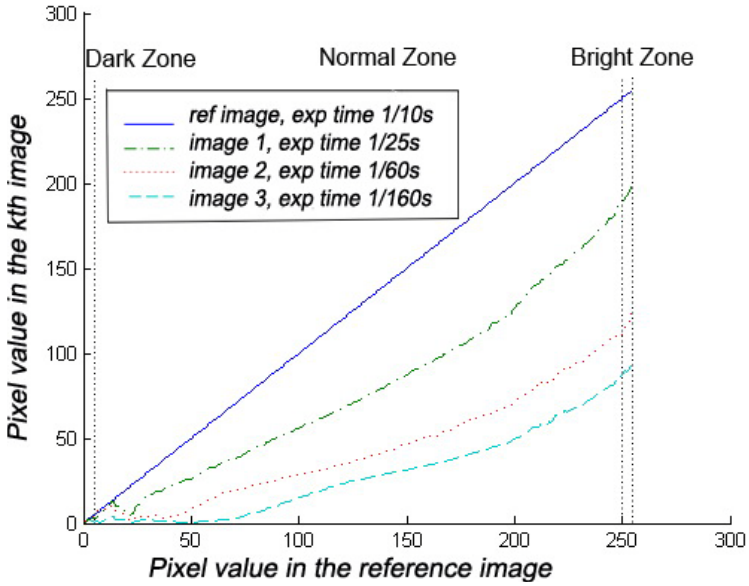

Fig. 2.3 An example of IRF function. Multiple IRF curves are presented with different images (captured at different exposure time) to the same reference image.

2.3 INPAINTING WITH EXPOSURE BRACKETING IMAGES

Images of same scene can be captured with different exposures and combined with computing power to synthesize image that overcomes limitation of conventional cameras. However, useful data can be lost due to camera shake, especially when capturing by a hand held device, which generates noticeable artifact in the synthesized image. In other words, different from the traditional image in-painting [53, 54, 55, 56] and scene completion [57], which generate only photorealistic patches, the degraded image in-painting in digital photography requires true luminance value of real world scene. Therefore, the challenge of patching is to find useful relations between missing pixels and the remaining pixels.

In **[P1]**, an IRF is proposed to in-paint the missing area with a dual patching, where dual reference images are used. In order to increase the accuracy, the reference image is selected to have the smallest exposure difference with the degraded image. However, the dynamic range lost caused by the characteristic 2, described in chapter 2.2, is still inevitable.

If the reference image has shorter exposure time than the degraded image, then as can be seen from the characteristic 3, the dark pixels in the reference image are mapped from a large contrast to a small one. In other words, it is a compressing process mapping multiple values to one, which in turn makes the IRF in this area more reliable. On the contrary, a highly compressed bright pixel in the reference image is mapped into multiple values in the degraded image, which causes inaccuracy due to the dynamic range lost. Thus, multiple reference images, with longer and shorter exposure time respectively, can be adopted to recover the lost dynamic range and enhance the patching accuracy. The missing pixel intensity is calculated by

$$Z_t(p) = \frac{\psi_{i,t}(Z_i(p)) \cdot \omega_{i,t}(Z_i(p)) + \psi_{j,t}(Z_j(p)) \cdot \omega_{j,t}\left(Z_j(p)\right)}{\omega_{i,t}(Z_i(p)) + \omega_{j,t}\left(Z_j(p)\right)} \qquad (2.\ 6)$$

where image i and j are the two reference images, image t is the degraded image to be patched, and ω is the weighting function defined as

$$\omega_{i,t}(z) = \begin{cases} \log(z+1), & \Delta t_{\mathrm{i}} \geq \Delta t_{\mathrm{t}} \\ \log(256-z), & \Delta t_{\mathrm{i}} < \Delta t_{\mathrm{t}} \end{cases}, z \in [0,255] \qquad (2.7)$$

which is plotted in Fig. 2.4.

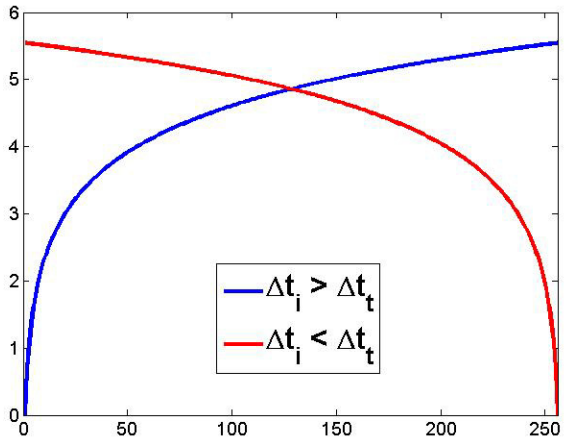

Fig. 2.4 Plot of the weighing function ω.

The missing area in an exposure bracketing image set can be recovered using the proposed intensity mapping. Compare with other methods, the exemplar-based in-painting [54] works well on simple texture, such as the table top. However, obvious errors can be seen at complex contents, like the baby's face and the title of the book as shown in Fig. 2.5 (b), due to short of reference. The CRF-based method [58] recovers the contents by the luminance shift from the reference pixels, where obvious artifact can be seen at the border. The method proposed in **[P1]** recovers the original scene effectively in terms of speed and quality. We tabulate the *peak signal-to-noise ratio* (PSNR) of each method in Table 2.1. The test was set up similar to Fig. 2.5 (a), where part of the image is missing due to camera movement, but can be patched using two references (a brighter reference

and a darker reference). This scenario is common in HDR imaging. The result was averaged from ten test sets.

Table 2.1 Typical acquisition approaches for capturing high dynamic range images.

Method	Inpainting [54]	CRF based [58]	IRF dual patching **[P1]**
PSNR (dB)	11.75	20.54	33.66

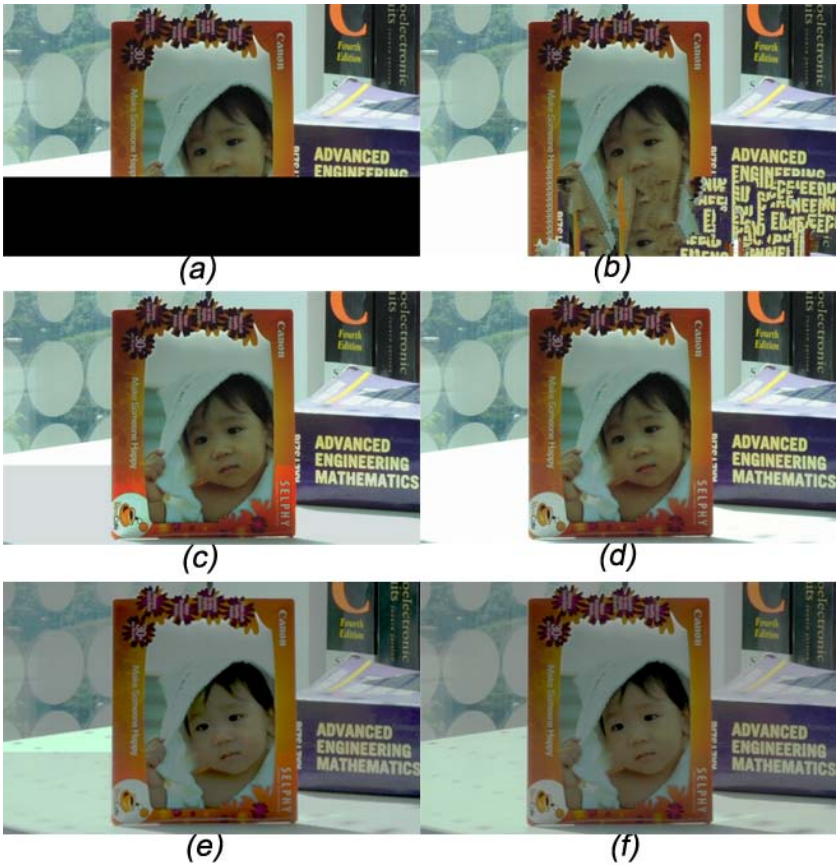

Fig. 2.5 Patching results of a degraded image from an exposure bracketing set shown in Fig. 2.2. (a) Original degraded image, (b) fixed image using exemplar-based in-painting [54], (c) fixed image using CRF [58], (d) fixed image using the proposed method, (e) HDR image synthesized with degraded image, and (f) HDR image synthesized with patched image ***[P1]****.*

The structure of the proposed method, intensity mapping plus dual patching, is extended to patch-based in-painting in

our later research works [59, 60]. It is proven to be an important tool for HDR de-ghosting [61, 62, 63].

2.4 BLIND EXPOSURE VALUE ESTIMATION

The exposure time is an important feature in synthesizing HDR images [1, 2, 50]. However, this information can be easily lost during copy or editing. Some methods have been introduced to recover the missing exposures. A radiometric model [14] was introduced to estimate the exposure ratio with an initial value set by user and an iterative process to fine-tune it. The results vary due to different initial settings and the order of the polynomial model used in the estimation. Another rough approximation was obtained by using IMF [51] between two images. However, as realized by the authors themselves, there are a lot of restrictions and assumptions involved. Thus, they used the approximated exposure ratio as the initial value to the radiometric model [14] in their experiments.

In **[P3]**, we present an accurate algorithm to recover the rate in between the different exposure ratios by using IRF **[P1]** and CRF [2]. For any pair of differently exposed images, each co-located pixels represent the same intensity. Thus, if they are not saturated, their reverse CRF mapping difference shall be a constant that represents the radiance difference of their exposure ratio. The exposure ratio is defined using *exposure value* (EV) interval as $EV_i = log_2(\Delta t_i / \Delta t_{i+1})$. For any three images in an exposure set $(i, i+1, i+2)$, we select one image (i) as reference. The function CRF (g) is calculated using the reference image and the first test image, while the pixel correlations are calculated using IRF between the reference image and the second image. As such, the EV unit in between the second test image and the reference image calculated by

$$\delta_{i+2,i+1,i} = \frac{\sum_{z=\acute{Z}_{i,i+2}}^{\grave{Z}_{i,i+2}} [g_{i,i+1}(z) - g_{i,i+1}\left(\varphi_{i,i+2}(z)\right)]}{\grave{Z}_{i,i+2} - \acute{Z}_{i,i+2} + 1} \quad (2.\ 8)$$

where $\grave{Z}$ and $\acute{Z}$ are two constants that defines the valid region boundaries as

$$\begin{cases} \acute{Z}_{i,j} = \max(\alpha, \varphi_{i,j}^{-1}(\alpha)) \\ \grave{Z}_{i,j} = \min(\beta, \varphi_{i,j}^{-1}(\beta)) \end{cases} \quad (2.9)$$

where α and β are two constants defines the saturation value. When the first and second test images are the same, a base EV unit is estimated. The unknown exposure ratio is then calculated by

$$k_{i,i+2} = \left(\frac{\delta_{i+2,i+1,i}}{\delta_{i+1,i+1,i}}\right) \cdot k_{i,i+1} \quad (2.10)$$

If there is no prior exposure knowledge, assumptions of the exposure ratio between the first two LDR images are given, as shown in the experiments in Table 2.2. The estimated EV internals have small errors of 5-10%, which will not generate obvious distortion in the final HDR image. If the initial EV was not given correctly, as shown in EXP3, all the estimated EV will have the same proportion to the given initial EV, which results in synthesizing a near-identical HDR image (Chapter 5.2).

Table 2.2 EV correction with given initial exposure ratio

		EV1	**EV2**	**EV3**	**EV4**
EXP1	Original	1	1	1	1
	[P3]	1 (given)	0.92	0.94	0.88
EXP2	Original	1.66	1.66	1.66	1.66
	[P3]	1.66 (given)	1.63	1.62	1.58
EXP3	Original	1	1	1	1
	[P3]	1.58 (given)	1.53	1.54	1.51

3 *De-ghosting*

When synthesizing HDR image using exposure bracketing images, ghosting artifacts usually appear due to moving object [64, 65, 66, 67], such as moving people and trees waving in the wind. To solve this, de-ghosting is applied before the HDR synthesis.

There are three major challenges in the de-ghosting process. Firstly, LDR images are taken with different exposures and cannot be compared directly. It is difficult to find a good criterion to bring different exposed images to the same comparable scale. Secondly, moving object pixels are classified by a threshold, but other factors, such as capturing parameters (ISO, shutter speed), capturing hardware and light conditions, affect the pixel co-relationship and militate against the selection of a static threshold. The third challenge is how to recover the missing dynamic range information.

3.1 IMAGE REGISTRATION

Before jump into de-ghosting, there is another important pre-processing step: *image registration* [68, 69, 70]. It is required if the images are captured using hand held devices. Otherwise, blur will appear in the synthesized HDR image. Luckily, when capturing multiple images in burst mode, misalignment are commonly seen as rotation and translation only. Thus, in most cases, only image alignment is required in HDR imaging.

Fig. 3.1 gives a simple comparison between image alignment and de-ghosting, which both cause blur in the synthesized HDR image. Blur due to mis-alignment is a global effect, while ghosting artifact due to moving object is content based. Thus, it is more difficult to do de-ghosting than alignment.

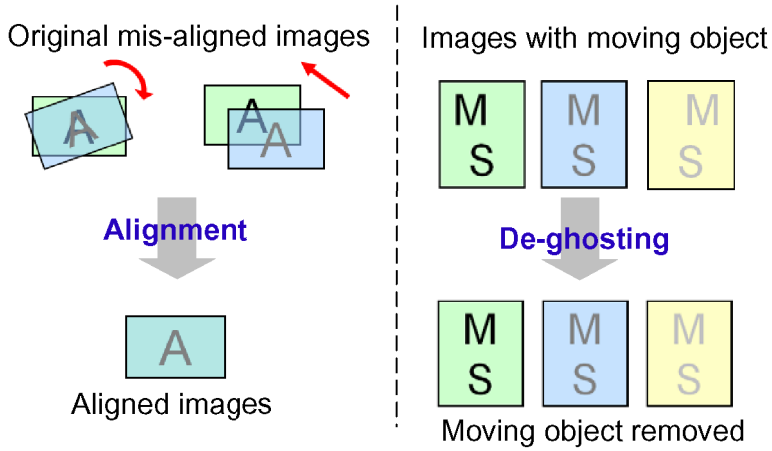

Fig. 3.1 Comparison with alignment (left) and de-ghosting (right).

We have made some study on image alignment methods for HDR imaging, as shown in Table 3.1. Some of the ideas are used in the final HDR imaging application.

Table 3.1 Comparison of image alignment methods

Method	[71]	[17]	[72]	[58]	[73]	[74]	[75]	[76]
Using feature	Median value	Pixel	Pixel	Median value	Pixel	Key points	Corner	Histogram
Support extreme DR		√	√					
No CRF			√	√		√	√	√
Fast feature matching	√					√	√	√

3.2 REVIEW OF DE-GHOSTING METHODS

Different from traditional motion detection[77, 78, 79, 80] in video processing, the input images (frames) are not at the same exposure level. Thus, special processes are needed.

An iterative approach was proposed in [81] by detecting the probability of a pixel belonging to a moving object or background, which minimized the assumption on the input image sequence. However, it is difficult to remove the moving object completely, and the algorithm is computationally expensive. Local entropy was proposed in [82] by matching the structure of the distribution of the local pixel value. It is a relative fast process, as there is no need for response curve estimation. However, as realized by the authors themselves, it cannot detect even large intensity differences when two co-located areas have similar local structure. Grosch [58] proposed a method using *camera response function* (CRF). It is suitable for the scenario when CRF is known, and can therefore perform a very fast detection process. Otherwise, with moving object, it is very difficult to estimate CRF accurately. A pre-classification process is adopted to solve this problem in [83]. Gallo et al. [84] used a similar idea with an assumption that log exposure can be mapped to a straight line, which in practice may not be the case due to different capturing hardware.

3.3 REALTIME DE-GHOSTING FRAMEWORK

A real-time de-ghosting method is proposed in **[P4]**, which includes detecting of moving object pixels and patching them using only the background pixels. Its major benefit is that it is fast and requires little computational resources, including processing power and physical memory. In all cases, the de-ghosting process involves only two images: a test image and the reference image. Because of this, the de-ghosting process is capable of working concurrently with the image capturing process, as shown in Fig. 3.2. The main function modules are the IRF estimation **[P1]**, bi-directional comparison, moving object classification and moving object correction.

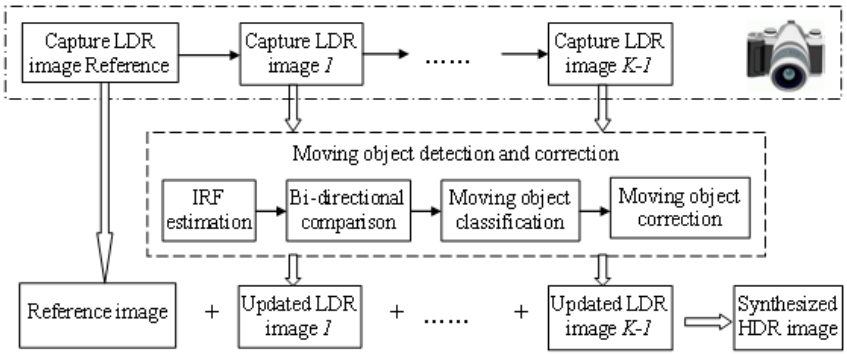

Fig. 3.2 Framework of the proposed real-time de-ghosting method.

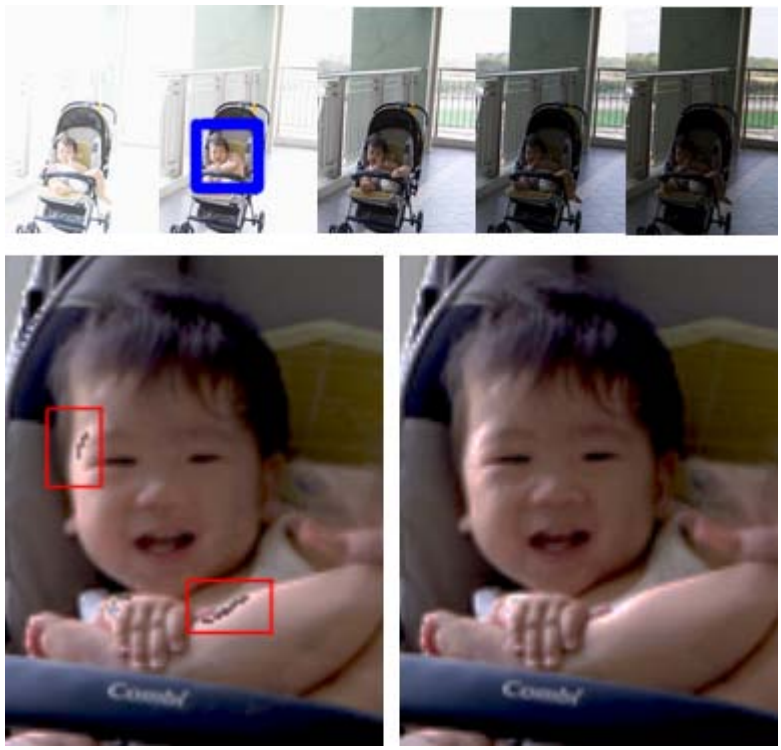

Fig. 3.3 Original input LDR images (top), and the zoom-in HDR images composed using uni-direction method (bottom left) and bi-directional method ***[P4]*** *(bottom right).*

The reciprocity law fails at underexposed and saturated pixels [52], as they are cut-off values due to the limitation of the dynamic range of the camera. Thus, the IRF values of those regions are unreliable. To solve this, bi-directional comparison is

proposed to mapping only from the more reliable pixel to the less reliable pixel as

$$\Gamma_{i,j}(p) = \begin{cases} \left|Z_j(p) - \psi_{i,j}(Z_i(p))\right|, & \omega(Z_i(p)) \geq \omega(Z_j(p)) \\ \left|\psi_{j,i}(Z_j(p)) - Z_i(p)\right|, & \omega(Z_i(p)) < \omega(Z_j(p)) \end{cases} \quad (3.\ 1)$$

where ω is a triangle weighting function [2], and Γ(p) is the error estimator at pixel p. The error estimator simulates the distance between the two co-located pixels despite the saturation area in the reference image. An example is given in Fig. 3.3.

Fig. 3.4 Original LDR images (top), and the zoom-in HDR images synthesized without outline re-evaluation (bottom left) and with outline re-evaluation ***[P4]*** *(bottom right).*

It is worth noticing that a simple but powerful outline re-evaluation filter can be adapted to the system before patching using IRF. The outline artifacts usually appear at the boundaries due to a similar value between moving object and the background, which makes it difficult to detect. The idea of the re-evaluation filter is to use a tight threshold and re-detect the

boundary pixels. An example of the outline re-evaluation filter is shown in Fig. 3.4.

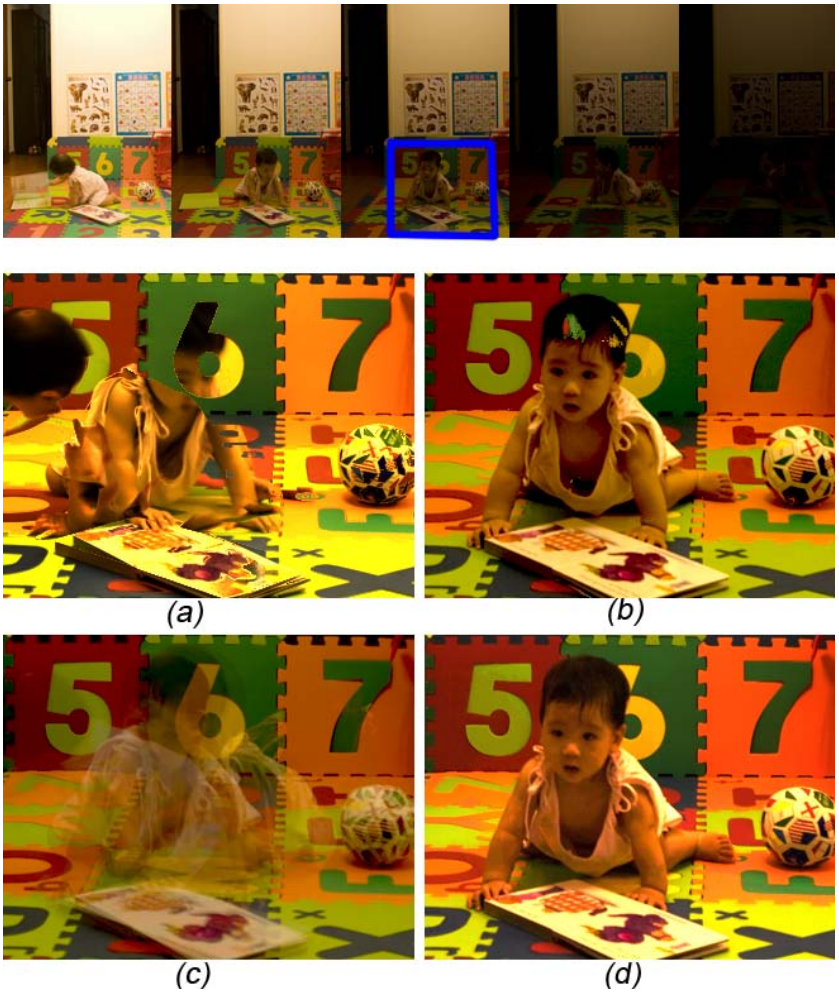

Fig. 3.5 HDR image synthesized from (top) five exposure bracketing images, using (a) FDRTools[1], (b) Photomatix[2], (c) Qtpfsgui[3], and (d) ***[P4]****.*

Similar to [58], once the error estimator is calculated, the moving object pixels are classified using empirical fixed threshold and patched using IRF. Experiments show a

[1] FDRTools is available at http://www.fdrtools.com/

[2] Photomatix is available at http://www.hdrsoft.com/

[3] Qtpfsgui is available at http://qtpfsgui.sourceforge.net/

significant improvement in the visual quality comparing to the off-the-shelf applications, such as Photoshop.

Our real-time de-ghosting framework has been tested on smart phones, where computation resources are limited. Combine with sub-sampling in Chapter 3.5, the processing time of the whole de-ghosting process for three 8MB (resolution 3264x2488) input images is only 700ms on Samsung Galaxy S3.

Fig. 3.6 An example of the failed case due to extreme large intensity changes, where (top) are the input image sequence, and (bottom) is the synthesized HDR image.

The tradeoff of this fast processing is the fixed threshold, which fails in some extreme scenarios, such as an example of a night scene shown in Fig. 3.6.

3.4 ROBUST DE-GHOSTING

The real-time de-ghosting method introduced in Chapter 3.3 is a useful solution for mobile applications, where processing power and memory are the most important concerns. On the other hand, when quality is the key requirement, such as an offline HDR synthesis using PC, a more robust de-ghosting method should be used. In **[P5]**, an improved de-ghosting method is proposed using double-credit intensity mapping which suits for more scenarios and gives better visual quality.

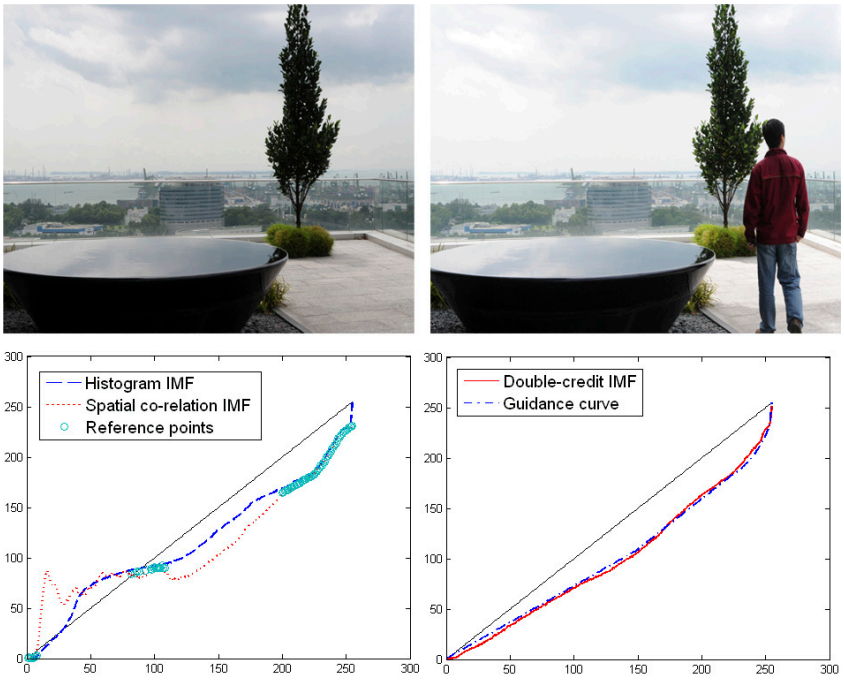

Fig. 3.7 An example of double-credit IMF estimation. Two input images with different exposures (Top); extraction of reference points (Bottom-left); guidance curve and the fine-tuned double-credit IMF (Bottom-right).

Histogram-based intensity mapping [51] uses the statistic information of the intensity distribution. However, such a process loses position information. On the other hand, the proposed IRF **[P1]** uses spatial correlation which preserves the position information, but can be distorted by large moving objects. Thus, a double-credit de-ghosting method is used,

which combines the advantages of both statistic information and spatial correlation, and generates HDR image with high visual quality.

The final double-credit IMF is estimated using the following steps. First, reference points are detected as the points with the same value in the histogram-based IMF as well as the spatial correlated IRF. A guidance curve is an approximation by extending the reference points with Bezier curve approximation [84, 85]. The guidance curve is a monotonic increasing function that cut across the identified reference points. It provides the initial value for the final IMF, and is useful to filter out large intensity changes due to moving object. The final double-credit IMF is then fine-tuned by calculating the spatial correlation using only the reliable pixels as

$$\psi(z) = \frac{\sum_{p\in\Omega(z)} Z(p) + \psi_G(z)\cdot\beta}{|\Omega(z)| + \beta}, z = 0, \dots, 255 \qquad (3.2)$$

where $\Omega(z)$ is the set of pixels with the same intensity of z in the reference image, $|\Omega(z)|$ represents the cardinality of this set, Z denotes the pixel of the test image, $\psi_G(z)$ is the guidance curve and β is a predefined weight to the guidance curve.

In order to robustly detect moving objects in different lighting conditions, an empirical fixed threshold is no longer feasible. In **[P5]**, a data driven training is conducted based on images captured from tripod cameras with no moving object in the scene. The training is aimed to find a threshold model suitable for different scenarios. A threshold vector is calculated based on pixel intensities (from 0 to 255) to classify 95% of the pixels as background. An example is shown at the bottom graph of Fig. 3.8. The image database includes more than 300 images from different cameras including Nikon D3, D300, Canon EOS-1, IXUS850, IXUS900, and images downloaded from Internet. The images have been captured using different capturing parameters at different lighting conditions to ensure extensive coverage. The adaptive threshold model is concluded as

$$T(z,M)=\begin{cases} a, & 0\le z\le d \\ a+\dfrac{b-a}{M-d}\cdot(z-d), & d<z\le M \\ b+\dfrac{c-b}{255-M}\cdot(z-M), & M<z\le 255 \end{cases} \quad (3.3)$$

where z represents the original RGB value, M represents the optimum bi-directional comparison central point, and a, b, c, d are the key parameters derived from the variance vector (V) generated during the fine-tuning of the double-credit IMF estimations. We use the following:

$$\begin{cases} a=0.8\cdot\sqrt{\dfrac{1}{d}\sum_{z=1}^{d}V(z)} \\ b=0.8\cdot\sqrt{\dfrac{1}{d}\sum_{z=M-d/2}^{M+d/2}V(z)} \\ c=5 \\ d=30 \end{cases} \quad (3.4)$$

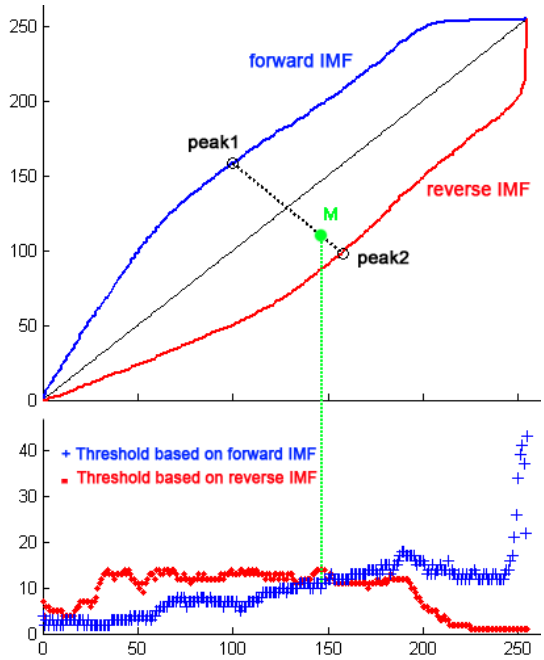

Fig. 3.8 The top graph plots the forward IMF and the reverse IMF of two static images. The bottom graph plots the thresholds, with which 95% of the pixels are classified as background. The forward IMF gives better performance on the left side of the central point M, and the reverse IMF minimizes the threshold value on the right.

Table 3.2 SNR comparison of different intensity mapping functions

SNR	Histogram	Spatial	Double-credit **[P5]**
SET1	20.32 dB	21.79 dB	27.50 dB
SET2	28.49 dB	18.39 dB	33.05 dB

The proposed double-credit IMF is compared with the histogram based mapping [51] and spatial correlation based mapping **[P1]** using two sets of images, of which each contains 30 bracketing image pairs. In SET1, the moving object (small size) appears only in one image. In SET2, the moving object appears in both images at different positions. An average improvement of 7 dB is achieved, see Table 3.2, and no ghosting artifacts appears in the visual quality test when comparing with the leading commercial software in Fig. 3.9 and Fig. 3.10.

There is an extreme case where the proposed method will fail. When the moving object is in a saturated area in the reference image, there is no information to patch during the synthesis. In this case, the result of the proposed de-ghosting algorithm may contain grey color patches. A hybrid patching algorithm was proposed in [59, 60] to handle these saturated pixels using block patching.

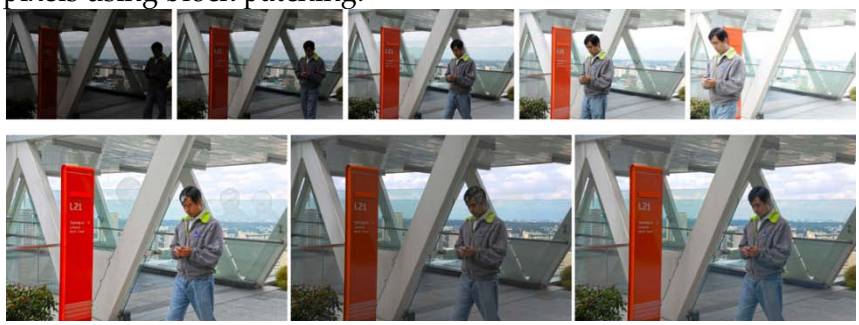

*Fig. 3.9 De-ghosting visual comparison in a day scene with (top) input images, using (bottom-left) Photomatix, (bottom-middle) Photoshop[1], and (bottom-right) **[P5]**.*

[1] Photoshop is one of the leading image processing software, and can be found at http://www.photoshop.com/.

Fig. 3.10 De-ghosting visual comparison in a night scene with (top) input images, using (bottom-left) Photomatix, (bottom-middle) Photoshop and (bottom-right) ***[P5]****.*

3.5 MOTION DETECTION FOR SUB-SAMPLED IMAGES

Modern smart phones, such as iPhone 5 and Galaxy S3, are capable of capturing images with resolution up to 8MB. Thus, sub-sampling based fast movement detection is proposed in **[P6]** to save computational resources for the de-ghosting algorithm to be running on the mobile devices.

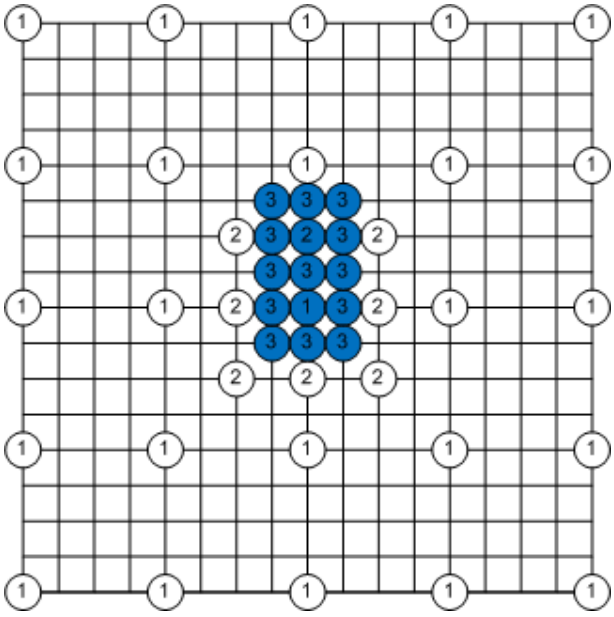

Fig. 3.11 A sub-sampling based movement detection scheme

An example with a sub-sampling factor of 4 is selected to illustrate the proposed fast movement detection scheme in Fig.

3.11. The pixels that are labeled with 1 are checked at the first round, and they are marked by the blue color if they are detected as moving object pixels. The neighboring pixels of a blue pixel, labeled with 1, are then checked by using a sub-sampling method. In Fig. 3.11, all neighboring pixels that are labeled by 2 are checked in the second round. They are also marked by the blur color if they are moving object pixels. Finally, the eight neighboring pixels of a blue pixel labeled by 2, i.e., those pixels are labeled by 3, are checked in the final round.

Only 1/16 of all pixels are detected in the first round. Since only a small portion of pixels in an LDR image belongs to moving objects, the second and third rounds of detections are only conducted for a small amount of pixels in the LDR image. As such, the complexity of the improved movement detection scheme is significantly reduced. Meanwhile, since many pixels are in the neighborhoods of two pixels, a flag is attached to each pixel to indicate whether it has been detected. With the flag, each pixel will only be processed once.

Experiments show that the overall number of processed pixels is reduced by up to 58.61%.

4 HDR De-noising

4.1 DE-NOISING OF LDR IMAGES

Image noise is usually caused by the capturing device. Common sources for noise are photon shot noise, dark current noise, and readout noise[87]. Noise may also come from special hardware, such as *fixed pattern noise* in *charge-coupled device* (CCD) camera [88, 89] and *thermal noise* and *flicker noise* in *complementary metal–oxide–semiconductor* (CMOS) device [90].

Lots of research studies have been conducted to tackle the single image de-noising problem. Simple local filters, such as *Gaussian mean filter* and *median filter*[91] bring in blur in the de-noised image. Edge preserved filters, such as bilateral filter[92], are developed to keep the sharp edges. Studies have been carried out on the *anisotropic diffusion* based techniques[93] and wavelet-based techniques[94]. Patch-based image de-noising find similar patch information inside the original image and stack them together to remove the noise while keep the edges, such as non-local mean [95, 96], BM3D[97], PLOW [98, 99].

When multiple images of the same scene are available, it is possible to merge the information from the multiple inputs. The simplest approach is the *frame average*[100], which theoretically reduces the noise variance by N (number of input images). However, when the multiple images are not captured using the same exposure, frame average fails due to un-aligned exposures. To solve this, noise-reduced HDR synthesis is studied.

4.2 NOISE-REDUCED HDR SYNTHESIS USING 2D-DENOISING FACTOR

The presence of noise in an HDR synthesis poses a serious degradation to the HDR image especially when the input

images are captured at low light condition or with high sensitivity settings. Various methods have been proposed to tackle this problem. We further classify them into three categories: 1) weighted frame average based preprocessing method; 2) intensity weighting based direct HDR synthesis method; 3) exposure time based HDR composition method.

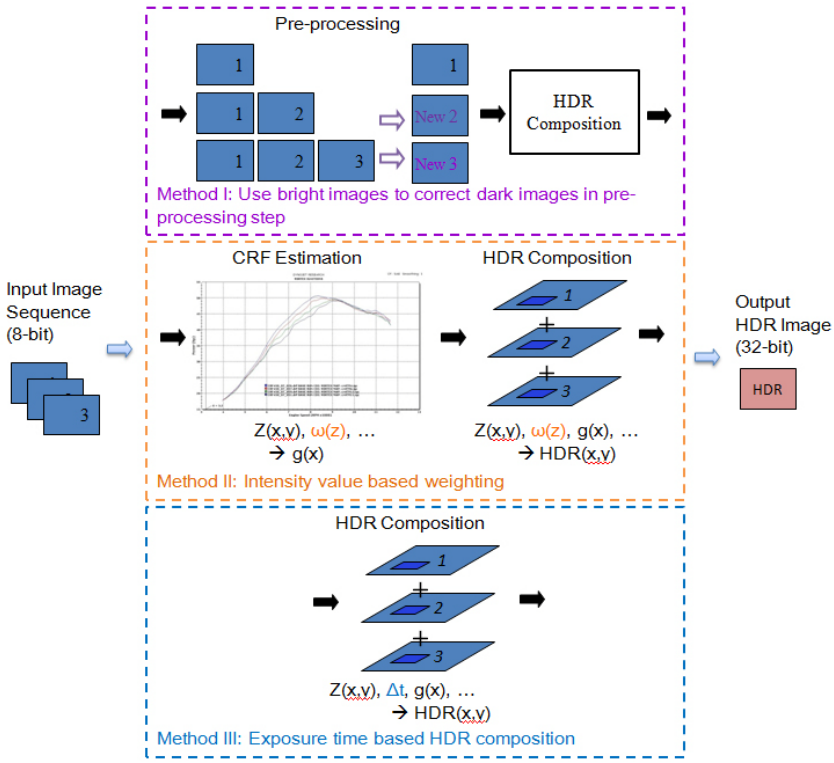

Fig. 4.1 Different approaches in noise-reduced HDR synthesis: (I) spatial-based edge preservation, (II) temporal based weighted frame average, (III) temporal-based intensity weighting, (IV) temporal-based exposure time weighting, and (proposed) temporal-based adaptive exposedness de-noising factor [P8]. The input image sequence can have more than 3 images.

The weighted frame average based methods, as shown in Fig. 4.1 (II), are based on an observation that brighter images contain less noise than the darker ones, as the brighter images have stronger signal and result in higher signal-to-noise ratios (SNRs). In order to average different exposed images, a calibration of the co-located pixels is required. CRF is used in [87] to recover the full radiance map for weighted averaging. IMF is used in [101]

to correct the dark image on LDR domain. These methods are capable of reducing noise in dark images at the cost of extra memory and computation time.

To reduce those costs, the intensity weighting based methods are designed to work directly in the HDR synthesis, as shown in Fig. 4.1 (III). A well-exposed pixel is assigned higher weight than an over/under-exposed pixel in both CRF estimation and HDR composition [87, 102, 103]. These methods are computational efficient and work well at extreme bright and dark area. But they are not as effective in the areas in between the result of weighted frame average methods.

Another approach attempts to use exposure times as additional weight [104] during the HDR composition, as shown in Fig. 4.1 (IV). However, the link between the estimation of the CRF and the proposed additional weight is missing. Color shift is another problem due to high weight at saturated pixels.

In **[P8]**, we propose a two-dimensional (2D) denoising factor to assign higher weight to pixel with less noise based on both pixel luminance and image exposure. It is controlled by two coefficients. It preserves edges and fine details without blurring artifact. In addition, both memory and computation time are significantly reduced compare to other denoising methods.

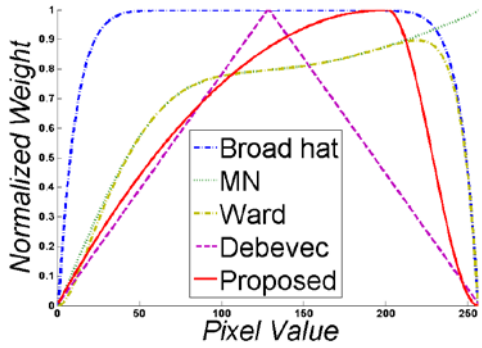

Fig. 4.2 The proposed luminance based weighting compares to other weighting functions: broad hat [1], MN [14], Ward [1] and Debevec [2].

The first dimension of the proposed denoising factor assigns a high weight to a pixel with a large luminance. Several methods have been proposed using different normalized

weighting functions, as shown in Fig. 4.2. Given an assumption that the noise distribution is independent of the measurement pixel value z, it is argued in [14] that a luminance based weighting function (*MN weight*) $\omega = f(z)/f'(z)$ achieves the best signal to noise ratio, where $f(z)$ is the radiometric response function. In [87], the MN weight is modified by replacing pixel value with luminance value. A broad hat function $h(z) = 1 - (\frac{2z}{255} - 1)^{12}$ was used to restrict the saturated pixels that may cause color cast.

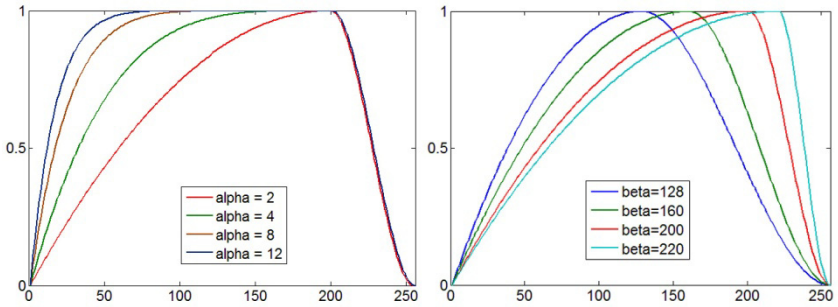

Fig. 4.3 A family of weightings in the first dimension calculated by using (left) different α with β=200, and (right) different β with α=2.

Since the radiometric response function is usually monotonic increasing, we approximate the luminance based weight by a controllable hat function and a Hermite interpolation. Thus, we can significantly reduce processing time on response function recovery by defining a new weighting factor as

$$\omega(z) = \begin{cases} 1 - \left|\frac{z}{\beta} - 1\right|^{\alpha}, & 0 \le z < \beta \\ 1 - 3\left(1 - \frac{255 - z}{255 - \beta}\right)^2 + 2\left(1 - \frac{255 - z}{255 - \beta}\right)^3, & \beta \le z < 255 \end{cases} \quad (4.1)$$

where the two key coefficient are the denoising strength coefficient α and the saturation control coefficient β. The smaller the value of α, the hat function will become steeper and this will result in better denoising effect. A large α gives high weights to pixels with small value (luminance), which remains noise in the

synthesized HDR image. A family of the proposed weighting function is plotted in Fig. 4.3. In our experiments, $\alpha = 2$ is chosen. Our tests indicate that $\alpha = 12$ generetes similar results as [87]. The saturation control coefficient limits the near saturated pixels to avoid color cast due to gamut limitations (an empirical value $\beta = 200$ is used here).

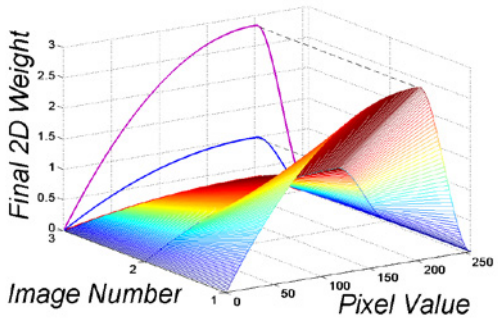

Fig. 4.4 An example of the proposed 2D denoising factor with 3 input images.

The second component of the proposed denoising factor is based on exposure time. More photons reach the camera sensor with a longer exposure time (Δt), which results in a more accurate reading. Thus, the proposed 2D weighting factor is designed to multiply the geometrically normalized exposure times with the luminance based denoising factor as

$$W(z, \Delta t_j) = \sqrt{\frac{\Delta t_j}{\sqrt[P]{\prod_{p=1}^{P} \Delta t_p}}} \cdot \omega(z) \qquad (4.2)$$

where j denotes the jth image among the P input images. The geometrical normalization avoids overwhelming big weights cause by images with very large exposure time. Thus, a modified objective function for estimating the CRF function is derived from [2] as

$$\sum_{i=1}^{N}\sum_{j=1}^{P}\{W(z,\Delta t_j)\cdot[g(Z_{i,p})-lnE_p-ln\Delta t_i]\}^2 + \lambda\sum_{z=Z_{min}+1}^{Z_{max}-1}[W(z,\max(\Delta t_1,\ldots,\Delta t_P))\cdot g^{"}(z)]^2 \quad (4.3)$$

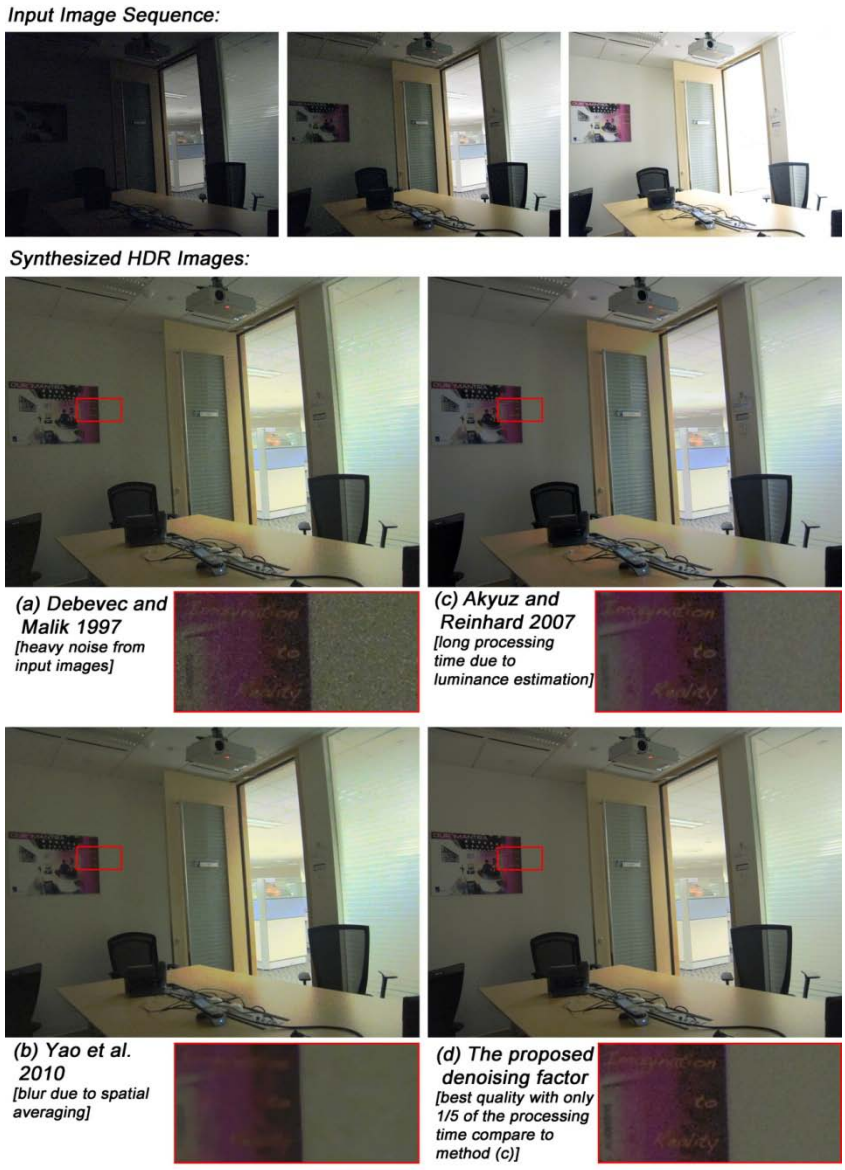

Fig. 4.5 Visual quality comparison for 2D de-noising factor

The proposed denoising method is tested by comparing it with three HDR synthesis methods, as shown in Fig. 4.5. The noise is significantly reduced compared to [2] with the same processing time. No blur artifact is generated as in [105] due to pure spatial averaging. Comparing with [87], the quality of the proposed is about the same. It is achieved with only 1/5 of the processing time, since no intermediate steps for approximating response function is required.

4.3 NOISE REDUCED TONE-MAPPING

Tone-mapping is usually the last step in the software solution of HDR imaging to display the HDR content on LDR display devices. In general, the tone-mapping is designed by using either global operator [26, 27, 28] or local operator [29, 30, 31, 32, 106]. The benefits of global operator are fast, easy to implement and no lighting inverse. Although some studies [107, 108] show that global operators can produce good results in some scenarios, many research studies are conducted based on sophisticated local operators for better local contrast and fine details.

A popular local tone mapping approach is based on the *Retinex theory* [109] that an image (I) is regarded as a product of two components ($I = L \cdot R$): an *illuminance component* (L) which contains large luminance variance, and a *reflectance component* (R) which contains *intrinsic* information. *Guassian filtering* method for decomposition was proposed in [110], and was soon be replace by *bilateral filtering* [29, 111] due to its better edge preservation. However, *halo artifact* is usually seen in the compressed image due to edges leaking to the reflectance component. Thus, better edge-preserving methods are proposed. A *weighted least square* (WLS) framework is proposed in [34] with progressive detail layer decomposition. Other methods focused on the smoothness constraint [32, 112] and the localized data term [113] for better visual quality and fast processing.

However, despite these improvements the main problem still remains: when the original HDR image contains noise, it remains in the tone-mapping result. There are two ways of de-noising for HDR tone-mapping. One is to direct apply a single image de-noising method [95, 96] after the tone-mapping algorithm. This approach is more time costly. The other approach is to do noise reduction during the tone-mapping process, as proposed in **[P7]**.

The process of the tone-mapping in **[P7]** is illustrated in Fig. 4.6, where the input HDR image is decomposed into a base layer and a detail layer [34]. The base layer contains all the high dynamic range content, and the detail layer contains the local contrast information. Noise is easily classified into detail layer as they share the same attributes as the small details, such as introducing small gradient and slightly increasing of local variance near the noise pixels.

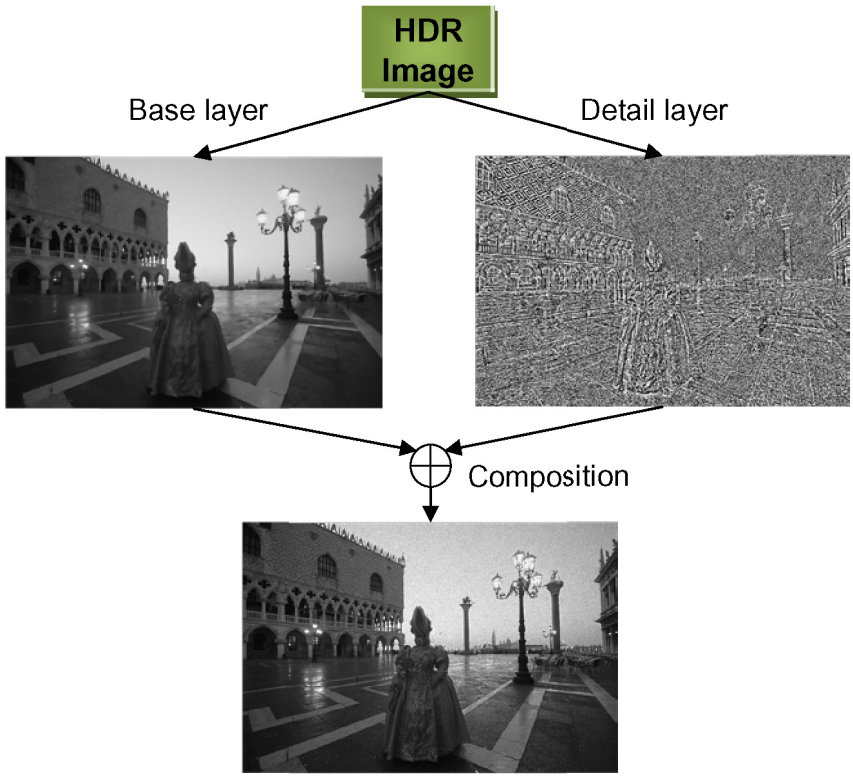

*Fig. 4.6 Proposed HDR tone mapping process in **[P7]***

In **[P7]**, a base layer *information content weighting* (ICW) and a detail layer ICW are proposed to reduce the noise in two steps. In edge-preserving decomposition, the data term defines the fidelity between the based layer and the original luminance. It is proposed in an L2 norm for easy calculation. First, the base layer ICW is introduced to be used on the data term as

$$\min_{\mathrm{B(p)}}\left\{\omega_b\big(I(p)\big)\cdot\sum_p\big(B(p)-I(p)\big)^2 + \lambda\sum_p\left(\Phi\left(\frac{\partial B(p)}{\partial x},a_x(p)\right)+\Phi\left(\frac{\partial B(p)}{\partial y},a_y(p)\right)\right)\right\} \tag{4. 4}$$

where ω_b is the spatial varying ICW based on perceptual of HVS, a_x and a_y are the smoothness coefficient, I denotes the log HDR luminance, B denotes the target base layer. Without the base layer ICW, the above equation is exactly the tone mapping using WLS edge-stopping filter [34]. Inspired by the information theory [114, 115, 116], the base layer ICW is defined as

$$\omega_b(p)=\frac{1}{2}\cdot\log\left(1+\frac{\sigma^2(p)}{\sigma_c^2}\right) \tag{4. 5}$$

where $\sigma(p)$ is the local variance at pixel p, and σ_c^2represents the channel noising power, which is determined by the saliency coefficient (c) as

$$\sigma_c^2=\vec{\Gamma}(\mathrm{c}\cdot\mathrm{P}) \tag{4. 6}$$

where $\Gamma(\mathrm{p})=\{\sigma^2(\mathrm{p}),\mathrm{p}\epsilon\mathrm{P}\}$, and $\vec{\Gamma}$ denotes the ascending sort of Γ. For example, $\mathrm{c}=0.5$ indicates the median value of Γ.

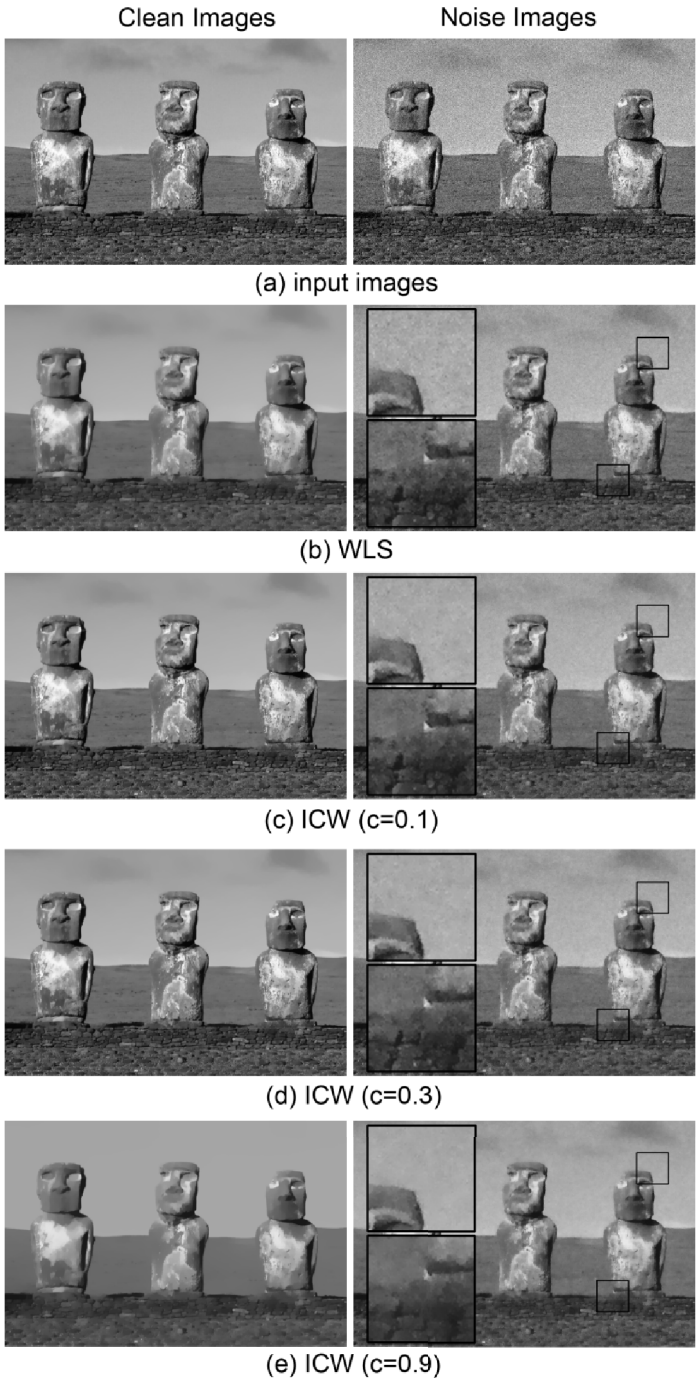

Fig. 4.7 Visual comparison of different saliency coefficients on (left column) clean image and (right column) noise image.

The saliency coefficient is an important factor that controls the saliency of the base layer ICW. A small saliency coefficient indicates a small channel noise, and results a high fidelity between base layer and the original image. A large saliency coefficient indicates a large channel noise, and results a smooth base layer. If the saliency coefficient is too big, the base layer can be over smoothed. A visual comparison of different saliency coefficient is illustrated in Fig. 4.7. In our implementation, $c = 0.3$ is chosen as a balanced value. However, if the noise level of the input HDR image is very high, a bigger saliency coefficient generates better denoising result.

During the decomposition process, the ICW helps to push the out of the base layer. Therefore, when the detail layer is subtracted from the clean base layer, more noise appears. Thus, a detail layer ICW is adopted as

$$D(p) = \omega_d(p) \cdot \big(I(p) - B(p)\big) \tag{4. 7}$$

where ω_d is the detail layer ICW derived from information fidelity criteria [114, 115, 116] as

$$\omega_d(p) = \frac{1}{\alpha} \cdot \log_2(1 + \beta \cdot (\frac{\sigma_{B \cdot I}(p)}{\sigma_B^2(p)})^\gamma) \tag{4. 8}$$

where σ_B^2 denotes the local variance of the base layer in a small window, $\sigma_{B \cdot I}$ denotes the covariance between the base layer and the original luminance, β and γ are two constants control the effectiveness of the weighting function, and α is a normalization factor which full fill the constraints of $\omega_d(p) \epsilon [0,1]$. If $\omega_d(p)$ is bigger than 1, the detail layer is amplified. It is commonly used in detail enhancement, with the risk of amplifying the noise at the same time. An example of the effectiveness of the detail layer ICW is given in Fig. 4.8.

The de-noising level is controlled by the different selection of two smooth coefficients β and γ, as shown in Fig. 4.8. It is worth noticing that the required details may be hidden by the heavy noise. In that case, when removing the noise, the details are

smoothed out too. Thus, balanced smooth coefficients are carefully chosen.

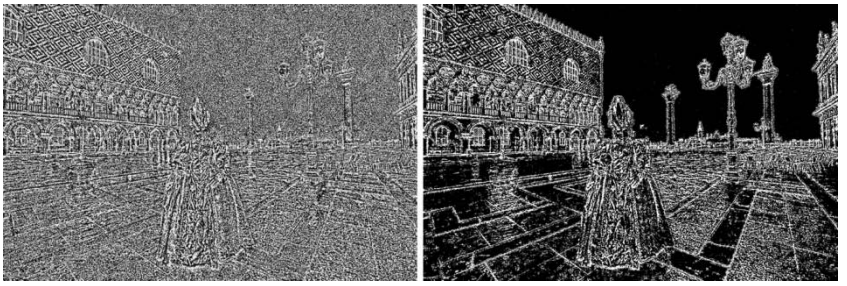

Fig. 4.8 Detail layer with (left) normal decomposition and (right) ICW-based de-noising factor

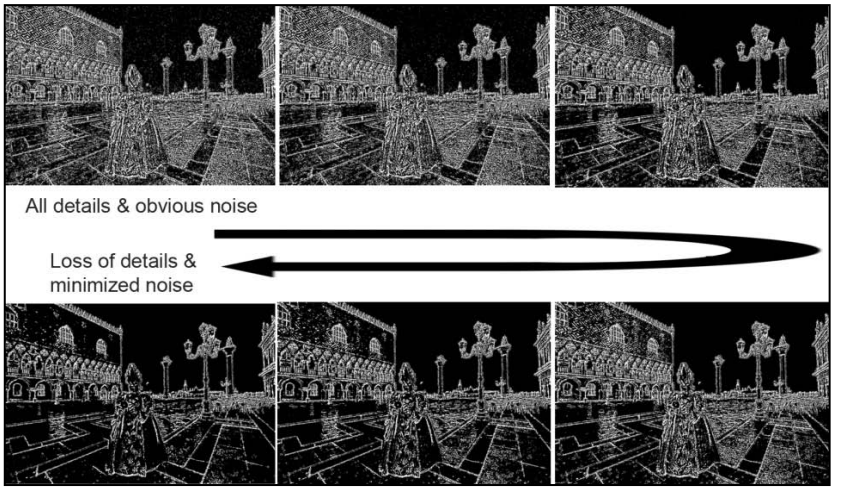

Fig. 4.9 Detail layer with different de-noising levels. The de-noising effect is increased from left to right, and from top to bottom.

We compared our solution with the approach of applying single image de-noising after tone mapping. We choose BM3D [97], which is one of the leading de-noising algorithms. The implementation was downloaded from the website provided by original authors[1]. The interface is a Matlab function, but the core is implemented in C. In order to test against the full capability of

[1] Matlab code of BM3D is downloaded from http://www.cs.tut.fi/~foi/GCF-BM3D/

BM3D, the noise level sigma is varied from 15 to 70, in order to find the best PSNR for BM3D.

Gaussian white noise is added to the original clean HDR image with different signal-to-noise ratio vary from 10dB to 25dB. Three tone mapping solutions are tested using WLS, WLS+BM3D, and the proposed ICW. As shown in Table 4.1, on average, the result of the proposed ICW is 1dB less than the result of WLS+BM3D solution, but the processing time is significantly reduced. The big difference occurs when the noise level of the input image is high. The BM3D method is able to use the information from the neighboring blocks at the tradeoff of the processing time.

In some experiments, the proposed ICW-based solution outperforms the combination of WLS + BM3D, when the noise level of the input image is not high. A possible explanation is that the proposed ICW works directly on the full radiance map of HDR image, while BM3D works on the result of WLS, which contains less information due to compression.

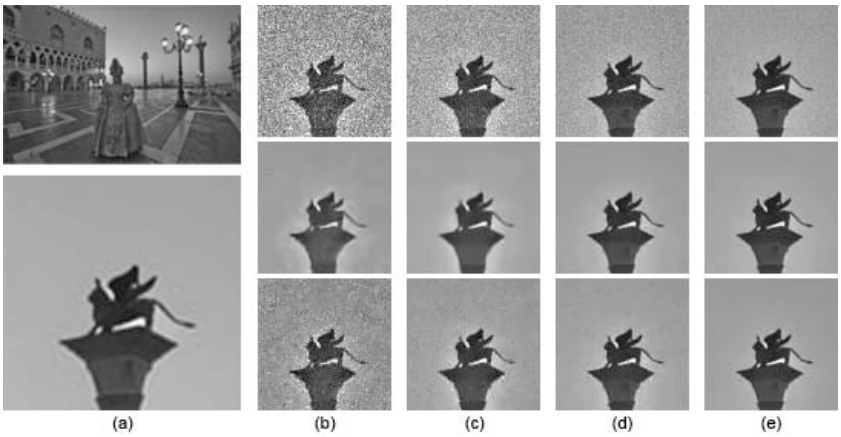

Fig. 4.10 Visual comparison with WLS+BM3D. (a) Clean image and a zoom-in object. The input image of (b-e) are added with different noise level at (b) SNR=10dB, (c) SNR=15dB, (d) SNR=20dB, and (e) SNR=25dB. In (b-e), the (top row) is WLS, the (middle row) is WLS+BM3D, and the (bottom row) is ICW proposed in ***[P7]****.*

Visual comparison was conducted with five leading tone mapping algorithms based on different approaches. We selected the state-of-the-art global operator (2008) [108], and used the

implementation in Luminance HDR [1]. The other four local operators were selected from a scale-decomposition-based solution (2005) [117], a bilateral-based decomposition (2007) [111], a slide-window-based direct luminance compression (2010) [118], and an edge-preserved decomposition (2008) [34]. All implementations were downloaded from the website of original authors. Visual comparison show good improvement in terms of noise reduction.

Table 4.1 Comparison with WLS+BM3D using quality metrics.

HDR image noise level	Method	WLS	WLS+BM3D	[P7]
SNR=10	PSNR	24.65	29.94	28.08
	SSIM	0.3949	0.8994	0.6869
SNR=15	PSNR	28.82	32.95	32.60
	SSIM	0.6194	0.9209	0.8913
SNR=20	PSNR	33.13	36.25	35.47
	SSIM	0.8230	0.9643	0.9456
SNR=25	PSNR	36.82	38.53	37.88
	SSIM	0.9330	0.9804	0.9710
Average Execution Time (Resolution 1200x800)		3 s	35 s	7 s

For simplicity, we applied the same method directly on the R, G, B color channels respectively, and therefore, color shift can be seen in some tone mapping result. Future work will focus on how to find in a better HDR color space for the least color distortions. Pioneer work about HDR color space is proposed in [119].

[1] Luminance HDR 2.3.1 is available at http://qtpfsgui.sourceforge.net/. It was previously named Qtpfsgui.

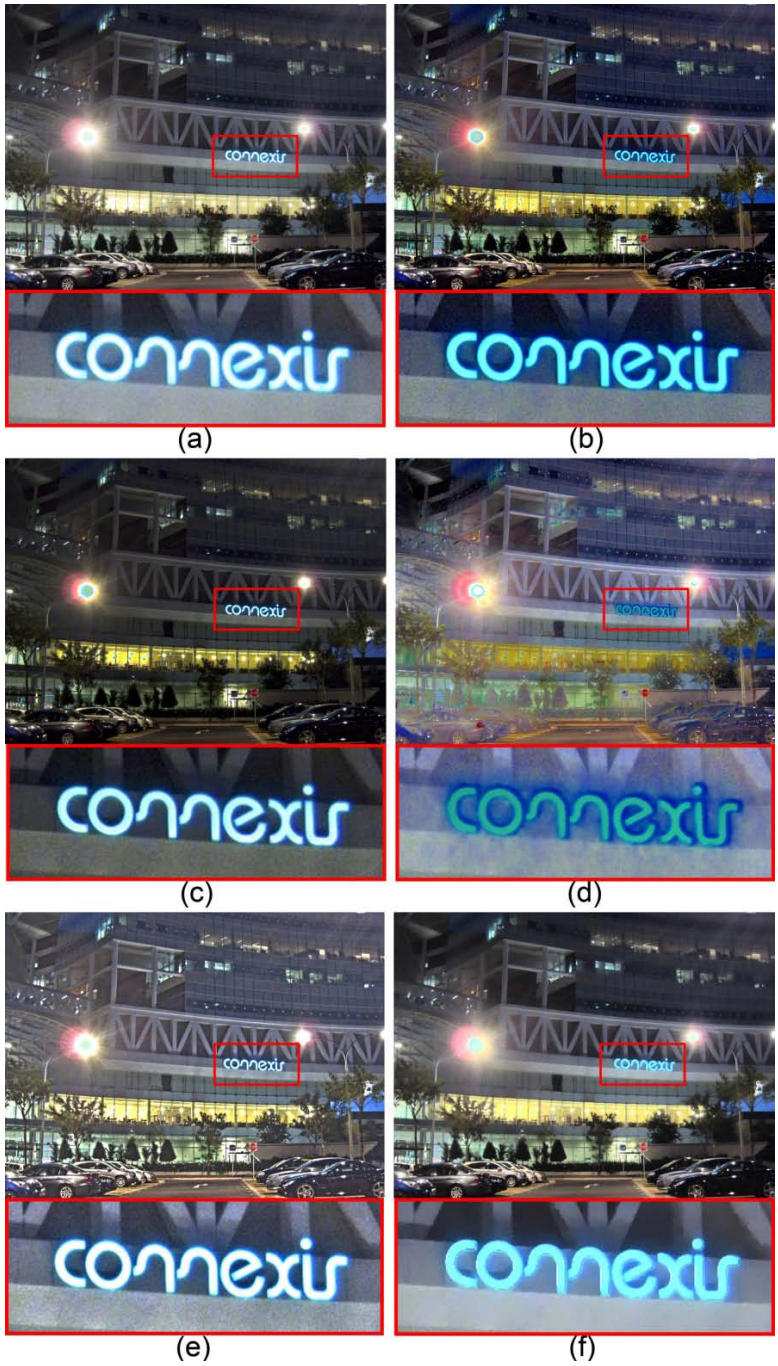

Fig. 4.11 Tone mapping visual comparison: (a) global tone curve [108], (b) scale-decomposition [117], (c) bilateral filtering decomposition [111], (d) direct luminance compression [118], (e) edge-preserved decomposition[34], and (f) noise reduced tone mapping based on ICW in ***[P7]****.*

5 Image Difference Metrics

5.1 IMAGE DIFFERENCE METRICS FOR EXPOSURE BRACKETING IMAGES

Quality metrics have been well studied for both image and video processing in the LDR domain. In image quality assessment, when a pair of images is given, the measurement of their similarity is classified as *full-reference* measurement. Many intensity-based indices [120, 121] have been proposed for their simplicity and fast processing. The most straightforward method is the *peak signal to noise ratio* (PSNR) calculated from the *mean square error* (MSE). However, PSNR is not accurate in many cases as it does not consider *human visual system* (HVS) [120]. An HVS-based structural similarity (SSIM) index [120] was proposed and was applied to video coding[122].

Fig. 5.1 Images captured with moving people and vehicle using different exposures settings: (a) +1 EV, (b) 0 EV, and (c) -1 EV. Although they are captured at the same scene within a short period of time, the similarity index (SSIM) between each pair of images appears to be low at SSIM(a,b)=0.6685 and SSIM(a,c)=0.4216.

Although these indices are designed to focus on cases where the dynamic ranges and intensities of the two images are almost the same, in practical situation where processing of high dynamic range imaging requires different level of exposures [2], such indices would not serve the scenarios, as shown in Fig. 5.1. In fact, it is known that the dynamic range can also be changed

due to images or videos with different bit-depth[123]. Moreover, the images being compared are assumed to be well registered by the intensity-based indices. However, it is unavoidable that there are rotations and translations in a set of differently exposed LDR images, especially when they are captured for outdoor scenes. Thus, in order to test if an image sequence is captured at the same scene and can be used for an HDR synthesis, it is desirable to provide simple similarity indices that are robust to translation, rotation, intensity change and dynamic range between two input images.

In **[P2]**, a structural similarity index for two LDR image by using intensity mapping and Richards' curve [124] is proposed. Due to the limitation of image un-alignment, the histogram based IMF [51] is used to unify the dynamic ranges or intensities of two images from the same scene[125]. On the other hand, when there is neither intensity nor dynamic range change between two images from the same scene, the accumulated histograms of two similar images are almost the same, and the structure of the corresponding image is kept even though it is mapped by the IMFs. The Richards' curve is adopted to design a switch on the deduction of the mean values from these two images. When two images are from the same scene, the similarity index value is calculated without the deduction of mean values from the corresponding images. As a result, the proposed index is robust to small translation and rotation between two images. When two images are from different scenes, the mean value is deducted from the corresponding images before the similarity index value is computed. As such, the proposed index is sensitive to two "similar" images from different scenes.

Different from the traditional quality metrics, the proposed image difference metrics does not contain any *mean-opinion-scores*. As such, it is designed to test if two images are from the same scene instead of measuring the image quality.

The proposed index is then extended by dividing the whole image into local windows as in[120]. The similarity of two images is detected by checking all pairs of local windows. Such

a local similarity index is insensitive to large intensity changes, dynamic range changes, translation and small rotations between images from the same scene.

The proposed image difference metrics start with mapping of two input images. Similar to the bi-directional comparison in **[P4]**, the bi-directional mapping is written as

$$\dot{Z}_1(p) = \begin{cases} \Lambda_{1,2}(Z_1(p)), & \omega(Z_1(p)) \geq \omega(Z_2(p)) \\ Z_1(p), & \omega(Z_1(p)) < \omega(Z_2(p)) \end{cases} \quad (5.\ 1)$$

$$\dot{Z}_2(p) = \begin{cases} \Lambda_{2,1}(Z_2(p)), & \omega(Z_2(p)) \geq \omega(Z_1(p)) \\ Z_2(p), & \omega(Z_2(p)) < \omega(Z_1(p)) \end{cases} \quad (5.\ 2)$$

where ω is a triangle weighting function, $Z_1(p)$ and $Z_2(p)$ are the intensity value of the two input images at pixel p, Λ denotes the histogram based IMF [51], and $\dot{Z}(p)$ represents an intensity unified pixel. This intensity unified pixel is then further transformed using Richards' curve [124] as

$$\ddot{Z}_1(p) = \dot{Z}_1(p) - \left(1 - \Gamma\left(\beta\left(\dot{Z}_1, \dot{Z}_2\right)\right)\right) * \mu_{\dot{Z}_1} \quad (5.\ 3)$$

$$\ddot{Z}_2(p) = \dot{Z}_2(p) - \left(1 - \Gamma\left(\beta\left(\dot{Z}_1, \dot{Z}_2\right)\right)\right) * \mu_{\dot{Z}_2} \quad (5.\ 4)$$

where $\mu_{\dot{Z}_1}$ and $\mu_{\dot{Z}_2}$ are the mean values of $\dot{Z}_1$ and $\dot{Z}_2$, respectively, $\beta\left(\dot{Z}_1, \dot{Z}_2\right)$ is the cosine of the angle between the two lexicographic order vectors of $\dot{Z}_1$ and $\dot{Z}_2$, and $\Gamma(z)$ is the Richards' curve defined as

$$\Gamma(z) = \frac{1}{1 + \exp\left(\varsigma_1(\varsigma_2 - z)\right)} \quad (5.\ 5)$$

where ς_1 and ς_2 are the growth rate and the threshold of maximum growth. Both Γ and β are used to detect the similarity of two images. On one hand, when two images are from the same scene, the existence of intensity changes, small translation and rotation between them would appear in the value of Γ

wherein the value will be almost equal to 1. On the other hand, when two images are from different scenes, the value of Γ approaches 0.

The similarity index is defined as

$$\mathrm{S_s}(Z_1, Z_2) = \frac{2\mu_{\ddot{Z}_1}\mu_{\ddot{Z}_2} + \varepsilon_1}{\mu_{\ddot{Z}_1}^2 + \mu_{\ddot{Z}_2}^2 + \varepsilon_1} \cdot \frac{\sum_p 2\ddot{Z}_1(p)\ddot{Z}_2(p) + \varepsilon_2}{\sum_p \ddot{Z}_1^2(p) + \ddot{Z}_2^2(p) + \varepsilon_2} \quad (5.6)$$

where ε_1 and ε_2 are two small positive constants and they are adopted to improve the robustness of the proposed index when the local signal to noise ratio is very low. Similar to SSIM index [120], the proposed similarity index is applied locally rather than globally, and the final index is the summation of all the local indices as

$$\mathrm{S_s^b}(Z_1, Z_2) = \frac{1}{p}\sum_{i=1}^{p} S_s(Z_1, Z_2). \quad (5.7)$$

Experiments show similar result when the input images are of the same exposure level, while the proposed similarity index has great improvement when dealing with exposure bracketing images. Table 5.1 shows the metrics comparison based on the images in Fig. 5.1.

Table 5.1 Metrics comparison using images in Fig. 5.1. The similarity indices mark from 0 (least similar) to 1 (identical).

Image Pair	(1,2)	(1,3)
SSIM [120]	0.6685	0.4216
S_s **[P2]**	0.9828	0.9780
S_s^b **[P2]**	0.9681	0.9504

Table 5.2 Metrics comparison using images in Fig. 5.2. The similarity indices mark from 0 (least similar) to 1 (identical).

Image Pair	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
SSIM [120]	0.5064	0.4750	0.4791	0.3803	0.3048
S_s **[P2]**	0.9787	0.9357	0.9879	0.9674	0.9845
S_s^b **[P2]**	0.9220	0.8547	0.9414	0.8875	0.8706

Another example is given with the input images captured using hand held camera, shown in Fig. 5.2. The image sequence contains camera movement, moving object, and exposure changes. The propose image difference metrics show the robustness in Table 5.2.

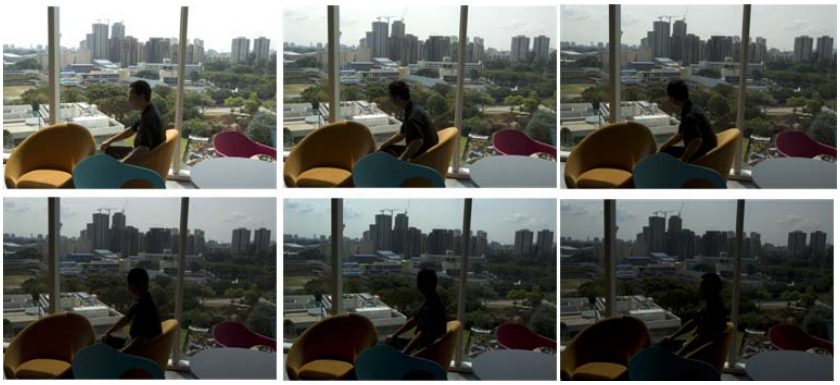

Fig. 5.2 Sequence of images contain camera movement, moving object, and exposure difference of EV interval of 2/3. Images are marked as 1 to 6 from left to right, top to bottom.

5.2 IMAGE DIFFERENCE METRICS FOR HDR RADIANCE MAP

Due to lack of strict link between physical light radiance and image radiance map, HDR images with the same scene can be very different in terms of their direct peak signal-to-noise ratio (PSNR), as shown in Fig. 5.3.

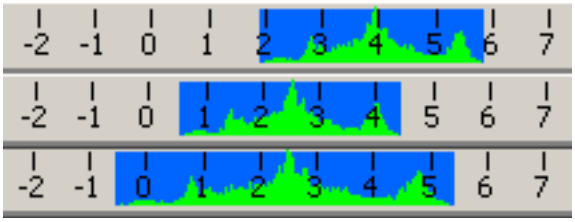

Fig. 5.3 Histogram in log radiance domain of (top) original HDR image, (mid) shifted HDR image, and (bottom) scale-invariant HDR image.

Near-identical HDR images are defined in **[P3]** as the images that carry the same scene information, but are shifted or have

different scale in the log radiance domain. These images are classified as

$$\mathrm{H} = \left\{\mathrm{HDRI_q}: \mathrm{R_q} \middle| \log\left(\mathrm{R_q}\right) = \mathrm{a_q} \cdot \log(\mathrm{R_0}) + \mathrm{b_q}\right\} \qquad (5.\ 8)$$

where R is the radiance of the HDR image, 0 represents the reference image defined as original, q represents the test image, $\mathrm{a_q}$ is the scale factor, and $\mathrm{b_q}$ is the shift factor. If the HDR image is composed from an LDR image set and the exposure info is known, the scale factor can be estimated as

$$\mathrm{a_q} = \frac{\log\left(k_{q,1,2}\right)}{\log\left(k_{0,1,2}\right)} \qquad (5.\ 9)$$

where $\mathrm{k_{1,2}}$ represents the exposure ratio between the first two images in the bracketing LDR image set. If the scale factor equals to 1, we call this HDR image shift-identical to the reference HDR image, and the shift factor is calculated as

$$\mathrm{b_q} = \log\left(\Delta t_{0,1}\right) - \log(\Delta t_{\mathrm{q},1}) \qquad (5.\ 10)$$

where Δt denotes the exposure time of the first LDR image in the sequence. If the scale factor does not equal to 1, we name it scale-identical, where the value of the shift factor is related to the scale factor and the image radiance.

When comparing the similarity of two HDR images, they are assumed to be near-identical. Thus, the scale factor is estimated using

$$\mathrm{a_q} = \frac{\max(h|\phi_0(h) > \varepsilon) - \min(h|\phi_0(h) > \varepsilon)}{\max\left(h\middle|\phi_q(h) > \varepsilon\right) - \min\left(h\middle|\phi_q(h) > \varepsilon\right)} \qquad (5.\ 11)$$

where ϕ is the histogram function and ε is a small number for the robustness to noise. The shift factor is calculated by correlating two HDR radiance maps using a sliding window. The biggest coefficient represents the best fit as

$$\mathrm{b_q} = \frac{\kappa(h) - H}{H} \cdot (M - m) \qquad (5.\ 12)$$

where H is the total number of bins in plotting the histogram, M and m are the maximum and minimum of the input radiance map, $\kappa(h) = \{h | \sigma(h) = \max(\sigma(i)), i = 1,2,\dots,2H\}$, and σ is the correlation coefficient function corresponding to the starting bin number of the sliding window. Finally, the proposed HDR metrics is calculated as

$$\mathrm{PSNR_{HDR}} = 10 \cdot log_{10}\left(\frac{P \cdot \Phi^2}{\sum_{p=1}^{P}[\mathrm{a_q} \cdot \log(R_{q,p}) + \mathrm{b_q} - \log(R_{0,p})]^2}\right) \qquad (5.\ 13)$$

where Φ is a constant.

Table 5.3 Compose HDR images with the following parameters in six experiments.

	$\Delta t_1(s)$	ΔEV_1	ΔEV_2	ΔEV_3	ΔEV_4
Original	0.02	1	1	1	1
EXP1	0.50	1.00	0.98	0.99	0.96
EXP2	0.50	1.58	1.53	1.54	1.51
EXP3	0.50	0.32	0.93	0.49	1.58
EXP4	0.02	1.32	1.42	0.58	0.74
EXP5	0.50	0.15	0.17	3.22	3.60
EXP6	0.03	0.26	0.32	1.00	3.32

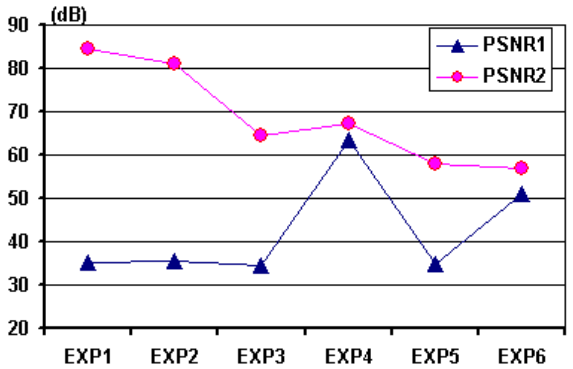

Fig. 5.4 Chart of PSNR from EXP1-6 from Table 5.1, where PSNR1 is the result of direct pixel-by-pixel comparison and PSNR2 is the result of ***[P3]****.*

An example is shown with HDR images synthesized with different parameters, as shown in Table 5.3. They are then compared with original HDR image using direct comparison and the proposed metrics. Results are illustrated in Fig. 5.4. In EXP1 and EXP2, despite the difference of initial exposure time (Δt_1) and small distortion of EV intervals ($\Delta EV_{\mathrm{i}} = \log_2(\Delta t_{\mathrm{i}+1}/\Delta t_{\mathrm{i}})$), the relative relations among ΔEV_1 to ΔEV_4 remains constant. The composed HDR images are verified to have high similarity in comparison to the original image, while the direct comparison shows a very low PSNR value in Fig. 5.4. In EXP3-6, random numbers are given for all parameters. The proposed metrics find the distortion and shown a drop of PSNR at average of 20dB. On the contrary, the direct comparison method was able to recognize the distortions.

6 Summary of Contributions

6.1 CONTRIBUTION OF THE THESIS

[P1]: A new intensity mapping, named inter-pixel relationship function (IRF), was proposed to map in between two images with different exposures by using their spatial correlations. The IRF has a variety of applications related to HDR imaging. In this paper, a dual patching is demonstrated using IRF with 10dB improvement compared to single directional patching using camera response function (CRF).

[P2]: In this paper, a structural similarity index is proposed for identifying if two images are captured from the same scene and can be used for HDR synthesis. Different from traditional quality metrics, the proposed method is robust to different dynamic range, small moving object, as well as possibly small rotation and translation. Therefore, it is useful to find exposure bracketing sequence from a large image database for batch processing.

[P3]: In this paper, we introduce the concept of near-identical HDR images, which carry the same scene information with different radiance representations. The same set of near-identical HDR images can be transformed to each other with linear functions in log radiance domain, which is useful in displaying HDR images and HDR tone mapping. By using the log histogram of the radiance map, we manage to derive the scale and shift parameters for the near-identical HDR images, which will help to get an accurate assessment using PSNR. A simple application of how to compose near-identical HDR images using IRF [P1] is also presented in the paper.

[P4]: A real-time framework to detect and remove the ghost artifact is proposed based on IRF and bi-directional comparison. The algorithm delivers robust ghost removal in O(n) time, which makes it a good candidate for mobile application.

[P5]: A robust de-ghosting algorithm is proposed to solve an extreme scenario where a moving object occupies a large area and distort the spatial correlation. The distorted spatial correlation is corrected by using pixel intensity distribution. A statistical threshold model is trained from the image database, and the key parameters are determined on the fly. Experiments show that the proposed algorithm achieves good visual quality in both day-time and night-time lighting conditions.

[P6]: In this paper, a sub-sampling based moving object detection is proposed to shorten the de-ghosting processing time for images with large resolution. Experiments show that the detection speed can be increased by 50%. Combining with this sub-sampling scheme, the real-time de-ghosting framework proposed in [P4] is capable of running on mobile device to de-ghost on three 8MB image within 700ms.

[P7]: A noise reduced tone-mapping is proposed to suppress the noise from the input HDR image. The de-noising process is embedded in the tone-mapping process based on the information content weighting. Therefore, it requires very little additional processing time. The experiments show a significant improvement in the visual quality of final tone-mapping image in terms of noise reduction.

[P8]: A 2D de-noising factor is proposed to generate noise reduced HDR image from a set of noise LDR images. Contrary to existing solutions that operates in the HDR radiance domain, the proposed de-noising factor works directly on the pixels of input LDR images, which reduces the processing time to 20% of the radiance-based method.

6.2 SUMMARY OF RESULTS

The publications **[P1-P8]** have covered a large portion of the HDR imaging value chain, as shown in Fig. 1.3.

Technologies described in **[P1, P4-6]** are implemented in our PC application and mobile application. An example of three input images with moving object is shown in Fig. 6.1. Results from two state-of-the-art algorithms, the entropy-based de-ghosting algorithm [82] and the patch-based iterative de-ghosting algorithm [126], are shown in Fig. 6.2 and Fig. 6.3 respectively. The former is good in the background sky but fails in the foreground people, while the latter performs the opposite way.

We also compare with two leading image processing applications, Photoshop and Photomatix. The results are illustrated in Fig. 6.4 and Fig. 6.5, where the ghosting artifacts are clearly visible. Our proposed method, as shown in Fig. 6.6, generates clear image without ghosting artifacts.

Expect the entropy-based algorithm [82], which requires to be running on MacOS, all the rest four are tested on the same laptop. Our proposed method uses the shortest processing time of 1.5 seconds. The others vary from 2.5 seconds to 350 seconds.

Fig. 6.1 Three input images captured with moving object at different exposure levels with 2EV appart from each other. The image resolution is 2144x1424. No camera movement is involved.

Fig. 6.2 HDR image synthesized using entropy-based de-ghosting algorithm proposed in [82]. Obvious ghosting artifacts are seen in red and green blocks. The implementation was downloaded from original author's website in the binary of PhotoSphere[1]. The processing time was around 5 seconds on a Mac Pro desktop. An HDR image was synthesized from the binary app, and it was then been compressed using Photomatix through tone mapping.

[1] PhotoSphere is a MacOS based binary implemented by the original authors. It is downloaded from http://www.anyhere.com/.

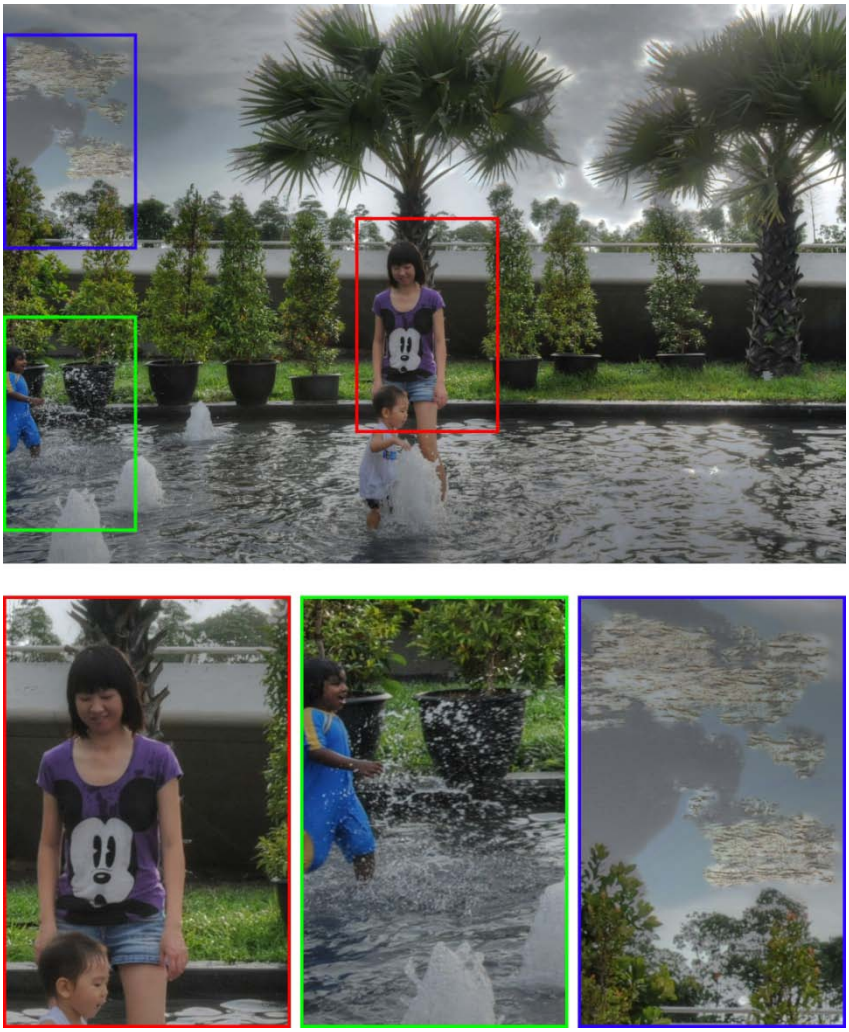

Fig. 6.3 HDR image synthesized using patch-based iterative approach in [126]. The de-ghosting performs well at foreground object in the red and green blocks. However, it fails at sky region. Obvious halo artifacts are seen around the palm tree leaves, and ghosting artifacts are seen in the blue block. The implementation was downloaded from original author's website[1] in the format of Matlab interface calling C routines. The whole processing cost 350 seconds.

[1] Source code was downloaded from author's project webpage at http://ece.ucsb.edu/~psen/hdr.

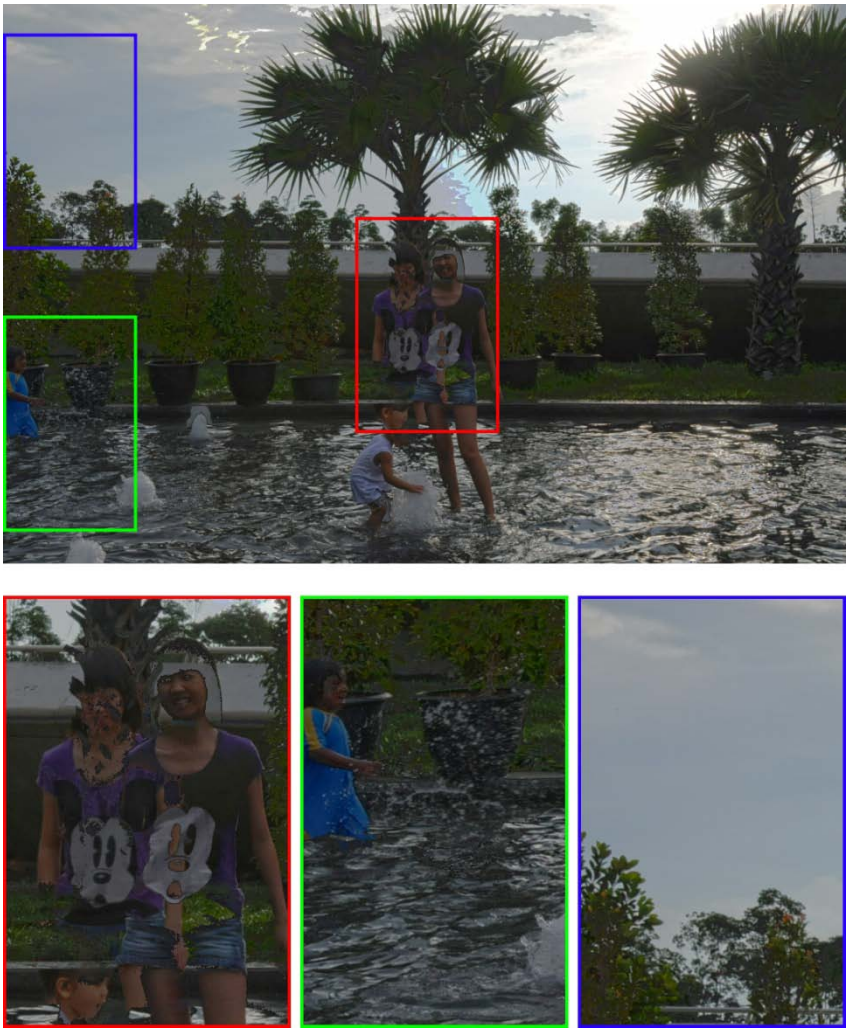

Fig. 6.4 HDR image synthesized using Photoshop CS5. The sky in the blue block is well kept, while some under the palm tree are distorted. The foreground people are heavily distorted as seen in the red and green block. The processing cost around 7 seconds on a Dell Precision M6700 laptop.

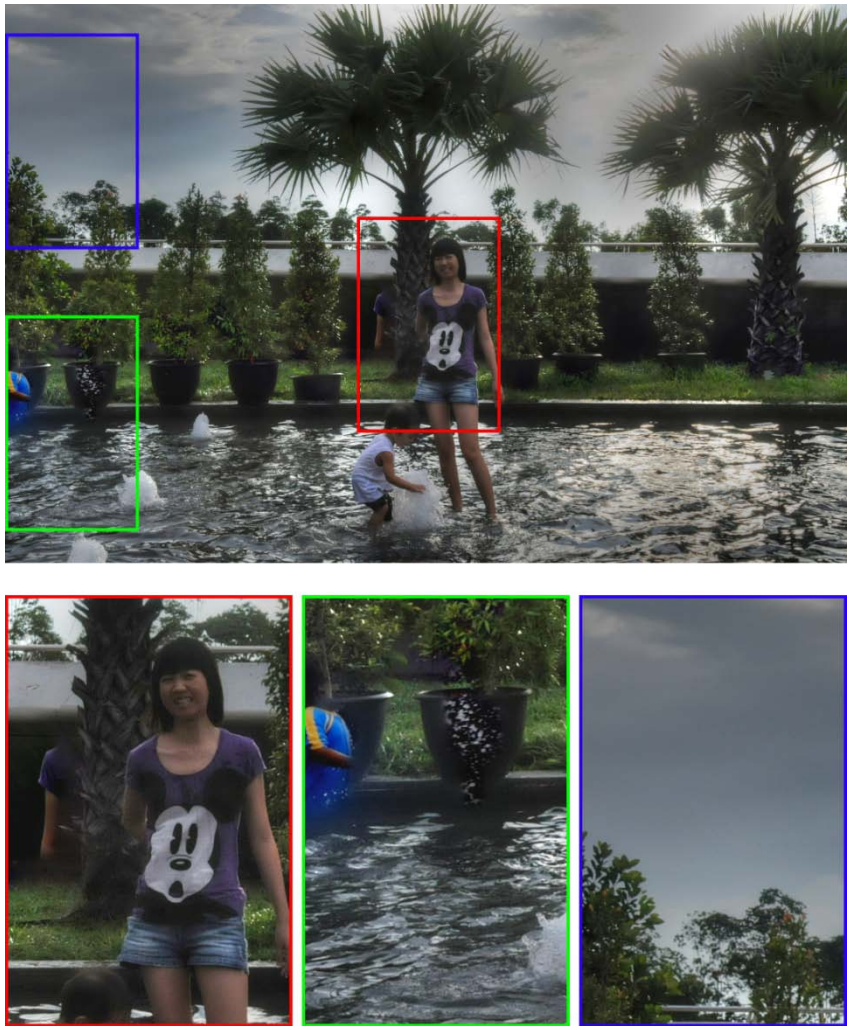

Fig. 6.5 HDR image synthesized using Photomatix Pro 4.2. The sky in the blue block performs well, while the foreground people suffer from ghosting artifacts, as seen in the red and green blocks. The processing time is around 2.5 seconds on a Dell Precision M6700 laptop.

Fig. 6.6 HDR image synthesized using our algorithm. A clear and sharp image is presented for both foreground people and background sky. The processing time is 1.5 seconds on a Dell Precision M6700 laptop.

State-of-the-art tone mapping algorithms usually focus on detail preservation and local contrast enhancement. We make use of information content weightings and proposed a noise-reduced tone mapping in **[P7]**. White Gaussian noise (SNR=20dB) is added in the original HDR image.

Fig. 6.7 Tone mapping using tone curve [108]. Implementation is found in Luminance HDR.

Fig. 6.8 Tone mapping using subbands decomposition [117]. Implementation is found from authors' website[1].

[1] Matlab source code is provided at
http://www.mit.edu/~yzli/hdr_companding.htm

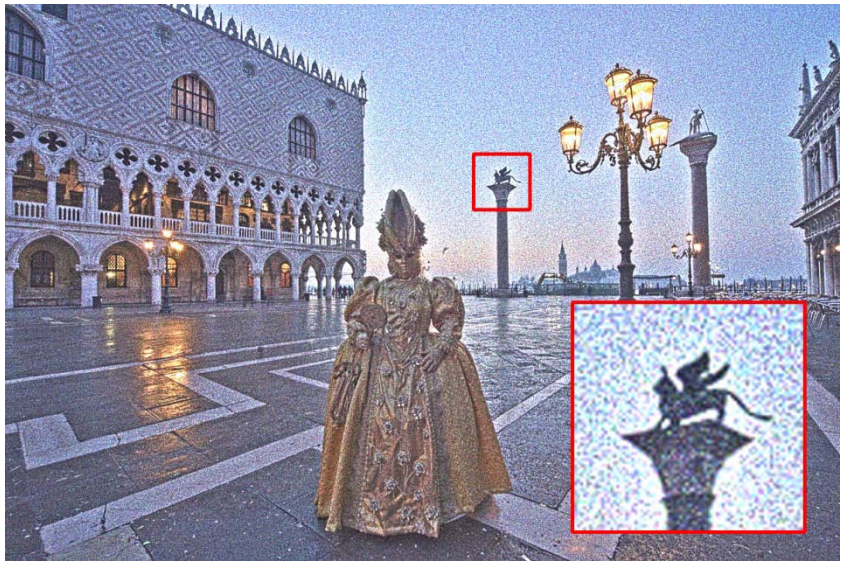

Fig. 6.9 Tone mapping using edge-preserved decomposition [34]. Implementation is found from authors' website[1].

Fig. 6.10 Noise reduced tone mapping proposed in ***[P7]****.*

We compare with three leading tone mapping algorithms: global tone curve [108] as shown in Fig. 6.7, subbands decomposition [117] as shown in Fig. 6.8, and edge-preserved

[1] Matlab source code is provided at
http://www.cs.huji.ac.il/~danix/epd/

decomposition [34] as shown in Fig. 6.9. Result of the proposed method is shown in Fig. 6.10.

It is important to test if a sequence of images is from the same scene, and can be used for HDR synthesis. A useful application is the batch processing for generating HDR images from a large database. Fig. 6.11 shows five pair of images captured from the same scene with different exposures. The proposed method from **[P2]** demonstrates the robustness to the exposure change.

Fig. 6.11 Five image pairs captured at the same scene with different exposures. The exposure value differences for the five pair of images are (1) 2/3EV, (2) 4/3EV, (3) 2EV, (4) 8/3EV, and (5) 10/3EV.

Table 6.1 Comparison of similarity indices using image pairs from Fig. 6.11

Image Pair	(1)	(2)	(3)	(4)	(5)
SSIM [120]	0.8638	0.5502	0.3229	0.1985	0.1067
S_s [P2]	0.9999	0.9999	0.9999	0.9999	0.9999
S_s^b [P2]	0.9994	0.9972	0.9917	0.9845	0.9855

Another five pairs of images are shown in Fig. 6.12, where pair 1 shows two images with moving objects and slightly exposure difference, and pair 2 and 3 are captured using hand-held camera with camera movement. The proposed image differnce metrics **[P2]** is robust to all these changes. Pair 4 and 5 are from two different scenes but looks similar. The proposed indices are able to identify the differences, as shown in Table 6.2.

Fig. 6.12 Five image pairs of different scenarios: (1) moving object in the image; (2) camera movement; (3) camera movement; (4) similar scene; and (5) similar scene. Pairs of (1)-(3) are suitable for HDR synthesis, while (4)-(5) are not.

Table 6.2 Comparison of similarity indices using image pairs from Fig. 6.12

Image Pair	(1)	(2)	(3)	(4)	(5)
SSIM [120]	0.4750	0.5688	0.6454	0.4730	0.4725
S_s [P2]	0.9357	0.9199	0.9634	0.6466	0.7193
S_s^b [P2]	0.8547	0.9153	0.8932	0.4802	0.5030

7 Conclusion

In this thesis, we have studied a high quality software solution for HDR imaging, including intensity mapping, de-ghosting, noise-reduced HDR synthesis, noise-reduced tone mapping and simple image difference metrics.

De-ghosting is an important step to remove artifacts caused by moving object. A real-time solution is proposed for the application of mobile devices. An extended robust solution is proposed for handling moving object covering large area. Both solutions require moving object detection and correction by using intensity mapping. In the future, a model of how to detect moving object without hard threshold shall be studied.

We have proposed to incorporate de-noising during the process of HDR synthesis and tone mapping respectively. The major advantage is computational efficiency. A 2D de-noising factor based on signal-to-noise ratio is proposed in the noise-reduced HDR synthesis. An information content based de-noising factor is proposed based on human visual system in the noise-reduced tone mapping. Both two de-noising steps operate in the RGB domain. In future work, better HDR luminance color space [119] shall be considered.

An image difference metrics for measuring whether a pair of images is from the same scene is proposed to identify correct image sequence from a large image database for the synthesis of HDR images. It is robust to limited distortions, including small image un-alignment, small moving object and large luminance change due to exposure difference. Another metrics measures the similarity of two HDR images using histogram-based transform. This is still a very preliminary work for difference measures. In the future more sophisticated dynamic range independent quality metrics [127, 128] can be studied, which can measure images with different bit depth.

8 Reference

[1] E. Reinhard, G. Ward, S. Pattanaik and P. Debevec, High dynamic range imaging: Acquisition, display and image-based lighting, San Mateo, CA: Morgan Kaufmann, 2005.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of Annual Conference on Computer Graphics and Interactive Techniques*, 1997.

[3] A. R. Varkonyi-Koczy, A. Rovid and T. Hashimoto, "Gradient-based synthesized multiple exposure time color HDR image," *IEEE Transactions on Instrumentation and Measurement,* vol. 57, no. 8, pp. 1779-1785, 2008.

[4] F. Banterle, A. Artusi, K. Debattista and A. Chalmers, Advanced high dynamic range imaging: theory and practice, CRC Press, 2011.

[5] J. J. McCann and A. Rizzi, The art and science of HDR imaging, John Wiley & Sons, 2011.

[6] K. Myszkowski, R. Mantiuk and G. Krawczyk, High dynamic range video, Morgan & Claypool Publishers, 2008.

[7] C. Bloch, The HDRI handbook: high dynamic range imaging for photographers and CG artists, Rocky Nook, 2007.

[8] J. Tumblin, A. Agrawal and R. Raskar, "Why I want a gradient camera," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[9] S. K. Nayar and T. Mitsunaga, "High dynamic range iImaging: spatially varying pixel exposures," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2000.

[10] G. Wetzstein, I. Ihrke and W. Heidrich, "Sensor saturation in fourier multiplexed imaging," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[11] S. K. Nayar, V. Branzoi and T. E. Boult, "Programmable imaging using a digital micromirror array," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[12] S. K. Nayar, V. Branzoi and T. E. Boult, "Programmable imaging: towards a flexible camera.," *International Journal of Computer Vision,* vol. 70, no. 1, pp. 7-22, 2006.

[13] S. Mann and R. W. Picard, "On being `undigital' with digital cameras: extending dynamic range by combining differently exposed pictures," in *Proceedings of IS&T*, 1995.

[14] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Fort Collins, CO, 1999.

[15] M. A. Robertson, S. Borman and R. L. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," *Journal of Electronic Imaging,* vol. 12, no. 2, pp. 219-228, 2003.

[16] M. Granados, B. Ajdin, M. Wand and C. Theobat, "Optimal HDR reconstruction with linear digital cameras," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.

[17] S. B. Kang, M. Uyttendaele, S. Winder and R. Szeliski, "High dynamic range video," *ACM Transactions on Graphics (TOG),* vol. 22, no. 3, p. 319–325, 2003.

[18] G. Impoco, S. Marsi and G. Ramponi, "Adaptive reduction of the dynamics of HDR video sequences," in *IEEE International Conference on Image Processing (ICIP)*, 2005.

[19] S. Mangiat and J. Gibson, "Spatially adaptive filtering for registration artifact removal in HDR video," in *IEEE International Conference on Image Processing (ICIP)*, 2011.

[20] D. Li, P. Tao and J. Wen, "High dynamic range video with synthesized gain control," in *IEEE International Conference on Emerging Signal Processing Applications (ESPA)*, 2012.

[21] Y. Schechner and S. K. Nayar, "Generalized mosaicing: high dynamic range in a wide field of view," *International Journal of Computer Vision,* vol. 53, no. 3, p. 245–267, 2003.

[22] M. Aggarwal and N. Ahuja, "Split aperture imaging for high dynamic range," *International Journal of Computer Vision,* vol. 58, no. 1, pp. 7-17, 2004.

[23] M. Mcguire, W. Matusik, H. Pfister, B. Chen, J. F. Hughes and S. K. Nayar, "Optical splitting trees for high-precision monocular imaging," *IEEE Computer Graphics and Applications,* vol. 27, no. 2, pp. 32-42, 2007.

[24] M. Tocci, C. Kiser, N. Tocci and P. Sen, "A versatile HDR video production system," *ACM Transaction on Graphics,* vol. 30, no. 4, p. 41, 2011.

[25] J. Kronander, S. Gustavson, G. Bonnet and J. Unger, "Unified HDR reconstruction from raw CFA data," in *Computational Photography (ICCP)*, 2013.

[26] A. Pardo and G. Sapiro, "Visualization of high dynamic range images," *IEEE Transactions on Image Processing,* vol. 12, no. 6, pp. 639-647, 2003.

[27] J. Duan and G. P. Qiu, "Fast tone mapping for high dynamic range images," in *International Conference on Pattern Recognition*, 2004.

[28] C. H. Liu, "Image characteristic oriented tone mapping for high dynamic range images," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2008.

[29] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Transactions on Graphics (TOG),* vol. 21, no. 3, pp. 257-266, 2002.

[30] R. Fattal, D. Lischinski and M. Werman, "Gradient domain high dynamic range compression," *ACM Transactions on Graphics (TOG),* vol. 21, no. 3, pp. 249-256, 2002.

[31] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics (TOG),* vol. 21, no. 3, pp. 267-276, 2002.

[32] Z. G. Li, S. Rahardja, S. S. Yao, J. H. Zheng and W. Yao, "High dynamic range compression by half quadratic regularization," in *IEEE International Conference on Image Processing (ICIP)*, 2009.

[33] T. Jinno, K. Mouri and M. Okuda, "HDR video tone mapping

based on gamma blending," in *IEEE International Conference on Image Processing (ICIP)*, 2010.

[34] Z. Farbman, R. Fattal, D. Lischinski and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Transactions on Graphics (TOG),* vol. 27, no. 3, p. 67, 2008.

[35] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh and A. Vorozcovs, "High dynamic range display systems," *ACM Transactions on Graphics (TOG),* vol. 23, no. 3, pp. 760-768, 2004.

[36] O. Bimber and D. Iwai, "Superimposing dynamic range," *ACM Transactions on Graphics (TOG),* vol. 27, no. 5, pp. 1-8, 2008.

[37] O. Bimber, T. Amano, A. Grundhofer, D. Kurz, D. Klock, S. Thiele and F. Hantsch, "Closed-loop leedback illumination for optical inverse tone-mapping in light microscopy," *IEEE Transactions on Visualization and Computer Graphics,* vol. 17, no. 6, pp. 857-870, 2010.

[38] R. Hoskinson, B. Stoeber, W. Heidrich and S. Fels, "Light reallocation for high contrast projection using an analog micromirror array," *ACM Transactions on Graphics (TOG),* vol. 29, no. 6, pp. 165:1-10, 2010.

[39] A. A. Goshtasby, "Fusion of multi-exposure images," *Image and Vision Computing,* vol. 23, no. 6, pp. 611-618, 2005.

[40] R. C. Bilcu, A. Burian, A. Knuutila and M. Vehvilainen, "High dynamic range imaging on mobile devices," in *IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2008.

[41] T. Mertens, J. Kautz and F. V. Reeth, "Exposure fusion," in *Pacific Conference on Computer Graphics and Applications*, 2007.

[42] T. Jinno and M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," *IEEE Transcations on Image Processing,* vol. 21, no. 1, pp. 358-365, 2012.

[43] M. Tico, N. Gelfand and K. Pulli, "Motion-blur-free exposure fusion," in *IEEE International Conference on Image Processing (ICIP)*, 2010.

[44] T. Kartalov, Z. Ivanovski and L. Panovski, "Fully automated exposure fusion algorithm for mobile platforms," in *IEEE International Conference on Image Processing (ICIP)*, 2011.

[45] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *IEEE International Conference on Computer Vision (ICCV)*, 1993.

[46] M. D. Grossberg and S. K. Nayar, "High dynamic range from multiple images: Which exposures to combine?," in *ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV)*, 2003.

[47] N. Barakat, A. N. Hone and T. E. Darcie, "Minimal-bracketing sets for high dynamic range image capture," *IEEE Transactions on Image Processing,* vol. 17, no. 10, pp. 1864-1875, 2008.

[48] K. Hirakawa and P. J. Wolfe, "Optimal exposure control for high dynamic range imaging," in *IEEE International Conference on Image Processing (ICIP)*, 2010.

[49] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H. P. Seidel and H. Lensch, "Optimal HDR reconstruction with linear digital cameras," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[50] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Transaction on Image Processing,* vol. 9, no. 8, pp. 1389-1406, 2000.

[51] M. D. Grossberg and S. K. Nayar, "Determining the camera response from images: What is knowable?," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),* vol. 25, no. 11, p. 1455–1467, 2003.

[52] "Wikipedia Reciprocity," Wikipedia, [Online]. Available: http://en.wikipedia.org/wiki/Reciprocity_(photography).

[53] J. C. Hung, C. H. Hwang, Y. C. Liao, N. C. Tang and T. J. Chen, "Exemplar-based image inpainting base on structure construction," *Journal of Software,* vol. 3, no. 8, pp. 57-64, 2008.

[54] A. Criminisi, E. Perez and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE*

Transactions on Image Processing, vol. 13, no. 9, pp. 1200-1212, 2004.

[55] A. Bugeau, M. Bertalmio, V. Caselles and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Transactions on Image Processing,* vol. 19, no. 10, pp. 2634-2645, 2010.

[56] Z. B. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Transactions on Image Processing,* vol. 19, no. 5, pp. 1153-1165, 2010.

[57] H. James and A. E. Alexei, "Scene completion using millions of photographs," *ACM Transactions on Graphics (TOG),* vol. 26, no. 3, p. 4, 2007.

[58] T. Grosch, "Fast and robust high dynamic range image generation with camera and object movement," in *Proceedings of Vision Modeling and Visualization*, 2006.

[59] J. H. Zheng, Z. G. Li, Z. J. Zhu and S. Rahardja, "A hybrid patching scheme for high dynamic range imaging," in *Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, 2011.

[60] J. H. Zheng, Z. G. Li, Z. J. Zhu, S. Q. Wu and S. Rahardja, "Patching of moving objects for ghosting-free HDR synthesis," in *ACM SIGGRAPH Poster*, 2012.

[61] Z. G. Li, Z. J. Zhu, S. Q. Wu, J. H. Zheng and S. Rahardja, "Fast correction of moving regions for differently exposed images with moving objects," in *IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2012.

[62] Z. G. Li, Z. J. Zhu, S. Q. Wu and S. Rahardja, "Fast patching of moving regions for high dynamic range imaging," in *ACM SIGGRAPH Asia 2010 Posters*, 2010.

[63] Z. G. Li, S. Q. Wu, Z. J. Zhu, S. L. Xie and S. Rahardja, "Anti-ghost of Differently Exposed Images with Moving Objects," in *IEEE International Conference on Image Processing (ICIP)*, 2012.

[64] H. L. Eng, J. X. Wang, A. H. K. S. Wah and W. Y. Yau, "Robust human detection within a highly dynamic aquatic environment in real time," *IEEE Transactions on Image*

Processing, vol. 15, no. 6, pp. 1583-1599, 2006.

[65] L. Y. Li, W. M. Huang, I. Y. H. Gu and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," *IEEE Transactions on Image Processing,* vol. 13, no. 11, pp. 1459-1472, 2004.

[66] A. Monnet, A. Mittal, N. Paragios and V. Ramesh, "Background modeling and subtraction of dynamic scenes," in *IEEE International Conference on Computer Vision (ICCV)*, 2003.

[67] K. K. Hadziabdic, J. H. Telalovic and R. Mantiuk, "Comparison of deghosting algorithms for multi-exposure high dynamic range imaging," in *Proc. of Spring Conference on Computer Graphics*, 2013.

[68] K. Simonson, S. Drescher and F. Tanner, "A statistics based approach to binary image registration with uncertainty analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),* vol. 29, no. 1, pp. 112-125, 2007.

[69] D. G. Lowe, "Object recognition from local scale-invariant features," in *IEEE International Conference on Computer Vision (ICCV)*, 1999.

[70] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2004.

[71] G. Ward, "Fast robust image registration for compositing high dynamic range photographs from hand-held exposures," *Journal of Graphics Tools,* vol. 8, no. 2, pp. 17-30, 2003.

[72] F. M. Candocia, "On the featureless registration of differently exposed images," in *Proceedings of International Conference on Imaging Science, Systems and Technology*, 2003.

[73] L. Cerman and V. Hlavac, "Exposure time estimation for high dynamic range imaging with hand held camera," in *Computer Vision Winter Workshop*, 2006.

[74] A. Tomaszewska and R. Mantiuk, "Image registration for multi-exposure high dynamic range image acquisition," in *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, 2007.

[75] M. Gevrekci, "On geometric and photometric registration of images," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2007.

[76] B. Guthier, S. Kopf and W. Effelsberg, "Histogram-based image registration for real-time high dynamic range videos," in *IEEE International Conference on Image Processing (ICIP)*, 2010.

[77] A. Mittal and N. K. Paragios, "Motion-based background subtraction using adaptive kernel density," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[78] Y. Ren, C. S. Chua and Y. K. K. Ho, "Motion detection with non-stationary background," in *International Conference on Image Analysis and Processing*, 2001.

[79] T. Mitsui and H. Fujiyoshi, "Object detection by joint features based on two-stage boosting," in *IEEE International Conference on Computer Vison (ICCV)*, 2009.

[80] Y. K. Chen, T. Y. Cheng and S. T. Chiu, "Motion detection with using theory of entropy," in *IEEE International Symposium on Industrial Electronics*, 2009.

[81] E. A. Khan, A. O. Akyuz and E. Reinhard, "Ghost removal in high dynamic range images," in *IEEE International Conference on Image Processing (ICIP)*, 2006.

[82] K. Jacobs, C. Loscos and G. Ward, "Automatic high dynamic range image generation for dynamic scenes," *IEEE Computer Graphics and Applications,* vol. 128, no. 2, pp. 84-93, 2008.

[83] Z. G. Li, Z. J. Zhu, S. L. Xie, S. Q. Wu and S. Rahardja, "Robust generation of high dynamic range images," in *IEEE Int. Conf. on Acoustics, Speech and Signal Procesing*, 2010.

[84] O. Gallo, N. Gelfand, W. C. Chen, M. Tieo and K. Pulli, "Artifact-free high dynamic range imaging," in *IEEE International Conference on Computational Photography*, 2009.

[85] "Wolfram Math World," [Online]. Available: http://mathworld.wolfram.com/BezierCurve.html.

[86] "Wikipedia Bazier Curve," Wikipedia, [Online]. Available: http://en.wikipedia.org/wiki/B%C3%A9zier_curve.

[87] A. O. Akyuz and E. Reinhard, "Noise reduction in high

dynamic range imaging," *Journal of Visual Communication and Image Representation,* vol. 18, no. 5, pp. 366-376, 2007.

[88] C. Liu, W. T. Freeman, R. Szeliski and S. B. Kang, "Noise estimation from a single image," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006.

[89] G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),* vol. 16, no. 3, pp. 267-276, 1994.

[90] K. H. Lundberg, "Noise sources in bulk CMOS," 2002.

[91] N. C. Gallagher and G. L. Wise, "A theoretical analysis of the properties of median filters," *IEEE Transactions on Acoustics, Speech and Signal Processing,* vol. 29, no. 6, p. 1136–1141, 1981.

[92] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *IEEE International Conference on Computer Vision (ICCV)*, 1998.

[93] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),* vol. 12, no. 7, pp. 629-639, 1990.

[94] J. Portilla, V. Strela, M. J. Wainwright and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transactions on Image Processing,* vol. 12, no. 11, p. 1338–1351, 2003.

[95] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Transactions on Image Processing,* vol. 15, no. 10, p. 2866–2878, 2006.

[96] C. Kervrann and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image denoising and representation," *International Journal of Computer Vision,* vol. 79, no. 1, p. 45–69, 2008.

[97] K. Dabov, A. Foi, V. Katkovnik and K. O. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing,* vol. 16, no. 8, p. 2080–2095, 2007.

[98] P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE*

Transactions on Image Processing, vol. 19, no. 4, p. 895–911, 2010.

[99] P. Chatterjee and P. Milanfar, "Patch-based near-optimal image denoising," *IEEE Transactions on Image Processing,* vol. 21, no. 4, pp. 1635-1649, 2012.

[100] A. C. Bovik, Handbook of image and video processing, Academic Press, 2010.

[101] W. Yao, Z. G. Li and S. Rahardja, "Intensity mapping function based weighted frame averaging for high dynamic range imaging," in *IEEE International Conference on Industrial Electronics and Applications (ICIEA)*, 2011.

[102] A. A. Bell, C. Seiler, J. N. Kaftan and T. Aach, "Noise in high dynamic range imaging," in *IEEE International Conference on Image Processing (ICIP)*, 2008.

[103] S. D. Neve, B. Goossens, H. Luong and W. Philips, "An improved HDR image synthesis algorithm," in *IEEE International Conference on Image Processing (ICIP)*, 2009.

[104] B. Goossens, H. Luong, J. Aelterman, A. Pizurica and W. Philips, "Reconstruction of high dynamic range images with Poisson noise modeling and integrated denoising," in *IEEE International Conference on Image Processing (ICIP)*, 2011.

[105] W. Yao, Z. G. Li, S. Rahardja, S. S. Yao and J. H. Zheng, "Half-quadratic ragularization based de-noising for high dynamic range image synthesis," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010.

[106] R. C. Bilcu, S. Alenius and M. Vehvilainen, "Adaptive local tone mapping of color images," in *Proceedings of European Signal Processing Conference (EUSIPCO)*, 2012.

[107] A. O. Akyuz, R. Fleming, B. E. Riecke, E. Reinhard and H. H. Bulthoff, "Do HDR displays support LDR content?: A psychophysical evaluation," *ACM Transactions on Graphics (TOG),* vol. 26, no. 3, pp. 38:1-38:7, 2007.

[108] R. Mantiuk, S. Daly and L. Kerofsky, "Display adaptive tone mapping," *ACM Transcations on Graphics (TOG),* vol. 27, no. 3, pp. 68:1-68:10, 2008.

[109] E. H. Land and J. J. McCann, "Lightness and retinex theory,"

Journal of the Optical Society of America, vol. 61, no. 1, pp. 1-11, 1971.

[110] D. Jobson, Z. U. Rahman and G. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing (TIP),* vol. 6, no. 3, pp. 451-462, 1997.

[111] J. Kuang, G. M. Johnson and M. D. Fairchild, "icam06: A refined image appearance model for hdr image rendering," *Journal of Visual Communication and Image Representation,* vol. 18, no. 5, pp. 406-414, 2007.

[112] G. Guarnieri, S. Marsi and G. Ramponi, "High dynamic range image display with halo and clipping prevention," *IEEE Transactions on Image Processing,* vol. 20, no. 5, pp. 1351-1362, 2011.

[113] B. Gu, W. Li, M. Zhu and M. Wang, "Local edge-preserving multiscale decomposition for high dynamic range image tone mapping," *IEEE Transactions on Image Processing,* vol. 22, no. 1, pp. 70-79, 2013.

[114] H. R. Sheikh, A. C. Bovik and G. Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Transactions on Image Processing,* vol. 14, no. 12, pp. 2117-2128, 2005.

[115] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing,* vol. 20, no. 5, pp. 1185-1198, 2011.

[116] Z. Wang and X. Shang, "Spatial pooling strategies for perceptual image quality assessment," in *IEEE International Conference on Image Processing (ICIP)*, 2006.

[117] Y. Li, L. Sharan and E. H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *ACM Transactions on Graphics (TOG),* vol. 24, no. 3, p. 836–844, 2005.

[118] Q. Shan, J. Jia and M. Brown, "Globally optimized linear windowed tone mapping," *IEEE Transactions on Visualization and Computer Graphics,* vol. 16, no. 4, pp. 663-675, 2010.

[119] O. Pirinen, A. Foi and A. Gotchev, "Color high dynamic range

imaging: The luminance–chrominance approach," *International Journal of Imaging Systems and Technology - Special Issue on Applied Color Image Processing,* vol. 17, no. 3, pp. 152-162, 2007.

[120] Z. Wang, A. C. Bovik, H. Sheikh and E. Simonceli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing,* vol. 13, no. 4, pp. 600-612, 2004.

[121] R. C. Gonzalez and R. E. Woods, Digital image processing, Prentice Hall, 2002.

[122] A. Glantz, A. Krutz, M. Haller and T. Sikora, "Video coding using global motion temporal filtering," in *IEEE International Conference on Image Processing (ICIP)*, 2009.

[123] K. J. Oh, S. Yea, A. Vetro and Y. S. Ho, "Depth reconstruction filter for depth coding," *IET Electronics Letter,* vol. 45, no. 6, pp. 305-306, 2009.

[124] F. J. Richards, "A flexible growth function for empirical use," *Journal of Experimental Botany,* vol. 10, no. 2, pp. 290-301, 1959.

[125] Z. G. Li, W. Yao, S. L. Xie and S. Rahardja, "Robust image similarity indices via intensity mapping functions and image partition," in *APSIPA ASC*, Singapore, 2010.

[126] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman and E. Shechtman, "Robust patch-based hdr reconstruction of dynamic scenes," *ACM Transactions on Graphics (TOG),* vol. 31, no. 6, pp. 203:1-203:11, 2012.

[127] R. Mantiuk, K. J. Kim, A. G. Rempel and W. Heidrich, "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Transactions on Graphics (TOG),* vol. 40, no. 14, pp. 40:1-40:14, 2011.

[128] T. O. Aydin, R. Mantiuk, K. Myszkowski and H. P. Seidel, "Dynamic range independent image quality assessment," *ACM Transactions on Graphics (TOG),* vol. 27, no. 3, pp. 69:1-69:10, 2008.

Zijian Zhu

High Quality High Dynamic Range Imaging

UNIVERSITY OF
EASTERN FINLAND

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences

ISBN 978-952-61-1171-1 (printed)
ISSNL 1798-5668
ISSN 1798-5668
ISBN 978-952-61-1172-8 (pdf)
ISSN 1798-5676 (pdf)