Winter 2005

Multi-Modal Enhancement Techniques for Visibility Improvement of Digital Images

Li Tao
Old Dominion University

MULTI-MODAL ENHANCEMENT TECHNIQUES FOR VISIBILITY IMPROVEMENT OF DIGITAL IMAGES

By

Li Tao
B.S. Sichuan University, Chengdu, China
M.S. Sichuan University, Chengdu, China

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

ELECTRICAL AND COMPUTER ENGINEERING

OLD DOMINION UNIVERSITY

December 2005

Approved by:

Vijayan K. Asari (Director)

Stephen A. Zahorian (Member)

Shunichi Toida (Member)

Min Song (Member)

ABSTRACT

MULTI-MODAL ENHANCEMENT TECHNIQUES FOR VISIBILITY IMPROVEMENT OF DIGITAL IMAGES

Li Tao
Old Dominion University
Director: Dr Vijayan K. Asari
Defense Date: December 12, 2005

Image enhancement techniques for visibility improvement of 8-bit color digital images based on spatial domain, wavelet transform domain, and multiple image fusion approaches are investigated in this dissertation research.

In the category of spatial domain approach, two enhancement algorithms are developed to deal with problems associated with images captured from scenes with high dynamic ranges. The first technique is based on an illuminance-reflectance (I-R) model of the scene irradiance. The dynamic range compression of the input image is achieved by a nonlinear transformation of the estimated illuminance based on a windowed inverse sigmoid transfer function. A single-scale neighborhood dependent contrast enhancement process is proposed to enhance the high frequency components of the illuminance, which compensates for the contrast degradation of the mid-tone frequency components caused by dynamic range compression. The intensity image obtained by integrating the enhanced illuminance and the extracted reflectance is then converted to a RGB color image through linear color restoration utilizing the color components of the original image. The second technique, named AINDANE, is a two step approach comprised of adaptive luminance enhancement and adaptive contrast enhancement. An image dependent nonlinear transfer function is designed for dynamic range compression and a multiscale image dependent neighborhood approach is developed for contrast enhancement. Real time processing of video streams is realized with the I-R model based technique due to its high speed processing capability while AINDANE produces higher quality enhanced images due to its multi-scale contrast enhancement property. Both the algorithms exhibit balanced luminance, contrast enhancement, higher robustness, and better color consistency when compared with conventional techniques.

In the transform domain approach, wavelet transform based image denoising and contrast enhancement algorithms are developed. The denoising is treated as a maximum *a posteriori* (MAP) estimator problem; a Bivariate probability density function model is introduced to explore the interlevel dependency among the wavelet coefficients. In addition, an approximate solution to the MAP estimation problem is proposed to avoid the use of complex iterative computations to find a numerical solution. This relatively low complexity image denoising algorithm implemented with dual-tree complex wavelet transform (DT-CWT) produces high quality denoised images. A wavelet transform based contrast enhancement technique is developed based on the correlation between the modified wavelet coefficients and a corresponding change in image quality. Comparison of the proposed technique with a curvelet based method shows that the new technique is extremely fast while providing a similar quality in the resulting images.

Both pixel-based and region-based image fusion methods are investigated for the purpose of incorporating much more scene information into a single image based on the information contained in multiple registered (aligned) source images. The three features in the wavelet transform based multiresolution fusion scheme proposed in this dissertation are: 1) a match measure obtained in the spatial domain which is applied to guide the fusion process, 2) a double-thresholding scheme which is adopted when combining corresponding pixels or regions, and 3) a multi-resolution segmentation process conducted on the match measured images. The proposed image fusion schemes are also implemented with DT-CWT. A prototype DVI (driver visibility improvement) system is developed based on the pixel-level image fusion algorithm.

Research work is progressing for the development of an algorithm which can make robust and optimal balance between the dynamic range compression of low frequency components and the enhancement of high frequency components. A neighborhood dependent coefficient shrinkage function based on estimated noise content is being developed for this purpose. In addition, the interlevel and intralevel dependency will be considered into the wavelet transform based image enhancement algorithms. Furthermore, advanced MR segmentation techniques based on texture analysis, gradient analysis, watershed methods and statistical methods are also being investigated for performing optimal fusion of multiple images.

ACKNOWLEDGEMENTS

My foremost thanks go to my advisor, Dr. Vijayan K. Asari. His invaluable academic guidance and inspiration carried me through the course of my Ph.D study and dissertation research. His encouragement, patience, and concern, helped me through many difficult times. His insights, wisdom and professionalism helped to shape my research skills and ethics.

I would also like to thank my dissertation committee members, Dr. Stephen A. Zahorian, Dr. Shunichi Toida, and Dr. Min Song for their willingness to review my dissertation. Their precious review comments and suggestions are sincerely appreciated for helping me to improve the dissertation.

I thank all my colleagues in the Computational Intelligence and Machine Vision Laboratory. I appreciate all the vivid discussions we had on various topics and cherish all the happy hours we spent together all these years. I always feel very lucky for being a member of this fantastic group.

Last but not least, special thanks go to my husband, Dr. Bing Xiao and my parents and sister in China for always being there when I needed them. Their caring, understanding and support helped me throughout my life here as a foreigner.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I
INTRODUCTION

1.1 Motivation for the Dissertation Research

Image enhancement is an important topic in digital image processing. It can help human viewers and computer vision algorithms obtain more accurate information from enhanced images. The visual quality and certain image properties, such as brightness, contrast, signal-to-noise ratio, resolution, edge sharpness, and color accuracy can be improved by the enhancement process. Many image enhancement algorithms have been developed based on various digital image processing techniques and applications. They can be developed in either the spatial domain or frequency domain.

Image enhancement techniques can be classified into various groups using different functionalities and criteria. For instance, image enhancement can be conducted on a single image using only the information contained in the original image while it can also be performed to create an enhanced image based on the information obtained from multiple input images. For example, image fusion is one of the enhancement techniques used to process multiple input images. On the other hand, image enhancement algorithms can also process images in different signal representation domains, such as the 2-D spatial domain, multiresolution (MR) representation in spatial domain (e.g., Gaussian and Laplacian pyramids), frequency domains (e.g., FFT domain and DCT domain), and spatial-frequency domain (e.g., wavelet transform domain). Due to different properties of various image processing techniques employed in image enhancement algorithms, each algorithm may have certain specialties compared to other algorithms in terms of capabilities, performance, speed, robustness, computation load, algorithm complexity, and so on. Therefore, it is necessary to investigate different image processing techniques to develop new image enhancement algorithms or improve existing algorithms for the purpose of improving the visibility in scenes and strengthening our capability to deal with various image processing and computer vision applications.

This dissertation research is dedicated to develop innovative image enhancement techniques for improving the visibility of low quality digital images caused by high dynamic range scene irradiance, noise, poor contrast, and very low illumination (low light condition). New image enhancement algorithms have been proposed and implemented based on various image processing techniques, and they are briefly introduced in the next section.

1.2 Summary of the Dissertation Contributions

The major research contributions achieved in this dissertation research are the following:

- Two new spatial domain nonlinear image enhancement algorithms are developed to solve the problems created by high dynamic range scene irradiance. Both algorithms deal with the issues of dynamic range compression, contrast improvement and color consistency. The first algorithm is originated from the concept of illuminance-reflectance (I-R) model of the scene irradiance. An adaptive dynamic range compression and an image dependent enhancement of mid-tone frequency components are performed on the estimated illuminance. The original reflectance value is combined with the treated illuminance for recreating the enhanced intensity image. The second enhancement algorithm, named AINDANE (Adaptive and Integrated Neighborhood Dependent Approach for Nonlinear Enhancement of color images) is a two-step approach comprising of adaptive luminance enhancement and an adaptive contrast enhancement processes. This configuration provides flexible control to the image enhancement process. A linear color restoration process is performed on the enhanced intensity image to restore its color information and maintain color consistency with the original color image. AINDANE provides better performance in terms of the visual quality of the image compared to that by the I-R model based algorithm due to its multiscale contrast enhancement mechanism. However, real-time video enhancement can be successfully realized using the I-R model based algorithm due to its relatively low computational complexity. Both algorithms are implemented in a newly developed software package named 'IESuite' for still image and video enhancement. An embedded pocket PC based application is also developed.

- New wavelet transform based image enhancement techniques are developed for image denoising and image contrast enhancement. Wavelet coefficients extracted from the original image are appropriately modified to obtain reconstructed images with reduced noise and improved contrast. Image denoising is treated as a statistical problem with a bivariate pdf (probability density function), which is used to model the interlevel statistical dependency among wavelet coefficients. An approximate solution of the bivariate shrinkage function for estimating the 'clean' image with reduced noise is developed to avoid sophisticated computations. The results of the proposed denoising algorithm are observed to be equivalent to the best results in the literature. Contrast enhancement is implemented by performing a nonlinear or a piece-wise linear transformation to the detailed wavelet coefficients. Dynamic range compression of the approximation coefficients is also performed to provide luminance enhancement to the images. Both algorithms are implemented using the dual-tree complex wavelet transform (DT-CWT) technique for optimal results.

- New wavelet transform based image fusion schemes, with a modified match measure calculation, image segmentation and decision making techniques, are developed. Both pixel-based and region-based approaches are investigated for performing multi-image fusion. In the proposed algorithms, match measure is computed in the spatial domain for reducing the computations and yielding more accurate results. Moreover, in the region-based fusion scheme, the regions are created by applying multi-resolution (MR) image segmentation on the match measure image. In both fusion schemes, corresponding pixels or regions are combined in wavelet domain using weighted averaging with the weight factors determined by match measure and activity measure. New double thresholding scheme based on the match measure value is designed to determine the values of weight factors according to their activity measures. More image features from source images can be appropriately incorporated into the fused images. Both wavelet transform based fusion algorithms are also implemented using DT-CWT, and a prototype image fusion software package named 'Image Fusion (IF)' is developed in a PC platform for fusion of still images and video files. The fusion algorithms are successfully applied for fusing multi-sensor, multi-illumination, and multi-focus images.

1.3 Organization of the Dissertation

The remaining chapters in this dissertation are organized as follows:

Chapter II presents a review of the conventional nonlinear image enhancement techniques, such as Retinex-based methods. These include the algorithms proposed in the field of computer graphics that have to deal with rendering of high dynamic range radiance map, similar to the problem of displaying high dynamic range images in the image processing field. A review of image enhancement techniques using wavelet transform focusing on image contrast enhancement and image denoising are also presented. This chapter also reviews the multi-resolution image fusion techniques which are either pixel based or region based.

In Chapter III, we present the two new spatial domain image enhancement algorithms. I-R model based algorithm will be introduced first followed by AINDANE. For both algorithms, the details of nonlinear dynamic range compression, single-scale or multi-scale neighborhood dependent contrast enhancement, and color restoration will be described and explained. Two algorithms will be separately discussed using various experimental results concerning their capability, adaptiveness, robustness, processing speed, and comparison with other techniques. Image statistics are also used to evaluate the effect of image enhancement in spatial domain. Comparison with other advanced techniques are performed and discussed in terms of clarity, processing speed, and color consistency.

Chapter IV presents wavelet transform based algorithms for image denoising and contrast enhancement. In image denoising, the statistical model considering the interlevel dependency of wavelet coefficients is introduced, and the derivation of the corresponding estimation problem of the 'clean' image is discussed. An approximate solution to the estimation problem is also presented. Finally, quantitative evaluation method is used to compare the proposed algorithm with other existing methods. In contrast enhancement, the correlation between the modified wavelet coefficients and the corresponding change of image quality is investigated. Nonlinear and piece-wise linear transfer functions for modifying the wavelet coefficients are developed to improve the local contrast in images. In addition, we propose to apply the dynamic range compression technique to the approximation images for luminance enhancement. This wavelet-based enhancement technique is implemented using DT-CWT and compared with a similar technique implemented with curvelet transform.

Chapter V presents wavelet transform based multi-resolution (MR) image fusion schemes for visibility improvement. Both pixel and region based approaches are investigated. The fusion scheme for each type of image fusion will first be introduced followed by detailed description of each step. A new match measure method and a new combination scheme are proposed, and a new region classification method is presented based on MR image segmentation applied on the match measure images. Both types of image fusion algorithms are implemented with DT-CWT technique. Experimental results and discussions are presented in this chapter. A new performance evaluation criterion is developed to evaluate and compare the efficiency of the proposed algorithms with the conventional techniques in the literature.

Chapter VI introduces the application systems developed in this dissertation work. Spatial domain algorithms based software packages for enhancement of still images and videos on both desktop/notebook PC and pocket PC are developed. Another software package for fusion of still images and video files for a prototype DVI (Drivers Visibility Improvement) system is also developed based on the pixel-level based image fusion algorithm proposed in this dissertation.

Finally, Chapter VII presents a summary of the major contributions of this dissertation work and suggestions for future work.

CHAPTER II

LITERATURE REVIEW

2.1 Overview of Nonlinear Image Enhancement in the Spatial Domain

Retinex based algorithms are effective techniques dealing with dynamic range compression and color constancy, which are developed from E. Land's theory [1-3] of human visual perception of lightness and color. Since the introduction of Retinex, several variants [4-7] on the original method have been developed mainly to improve the computational efficiency while preserving the basic principles. However, those methods are not widely used because of comparatively low processing speed and some issues concerning the visual quality of enhanced images. Nevertheless, more algorithms and applications have been developed based on Retinex due to its deep understanding of the lightness and color perception [8, 9].

MSRCR (Mutiscale Retinex with Color Restoration)[10-14], proposed by Z. Rahman, *et al*, is a newly developed and widely cited image processing technique which is a Retinex based algorithm using logarithmic compression and spatial convolution to implement the idea of Retinex. It aims to synthesize local contrast enhancement, color constancy, and lightness/color rendition for digital color image enhancement. MSRCR generally works well with various types of images. However, it also has some drawbacks that need to be tackled for approaching optimal performance [15-18]. Since the standard MSRCR algorithm needs to process all spectral bands of color images, it takes a long time to enhance large images. Thus MSRCR is hard to use in real time applications without a hardware implementation of the algorithm. In addition, the nonlinear color restoration process may produce colors, which are not predictable and sometimes make images look unnatural. For images having a dark subject with a very bright background, MSRCR seems to have difficulty providing sufficient luminance enhancement for the subject without post-enhancement treatment, and MSRCR may also cause the decrease in the luminance of the background, which can deteriorate the quality of the enhanced images if no post-treatment is performed. Finally, images enhanced by MSRCR may have some artifacts appearing at the boundaries with a large luminance change between the bright and dark regions. This so-called "halo effect" is associated with the convolution.

Histogram equalization (HE) is a well-established technique for image enhancement. However, HE only works well for scenes that have uni-modal or weakly bi-modal histograms (i.e. very dark or very bright scenes). For scenes with strongly bi-modal histograms (i.e. scenes that contain both very dark and very bright regions), HE performs poorly. Therefore, adaptive histogram equalization (AHE) was introduced [19]. AHE, also called localized or windowed HE, produces a local contrast enhancement by performing HE within a window whose size is adjustable for an optimized result. AHE definitely performs better than normal HE. However, the contrast enhancement is so strong that two major problems arise: intensive noise enhancement in "flat" regions and "ring" artifacts at strong edges [20]. To deal with those problems, a generalized version of AHE, contrast limiting AHE (CLAHE) was designed [21]. CLAHE has more flexibility in controlling the local contrast enhancement by selecting the clipping level of the histogram. Undesired noise amplification is reduced. In addition, the boundary artifacts can also be reduced using background subtraction [22]. Multi-scale AHE (MAHE) is the most advanced variation of HE. Unlike traditional single scale techniques, wavelet-based MAHE is capable of modifying/enhancing image components adaptively based on their spatial-frequency properties [23]. Those advanced HE variations generally have very strong contrast enhancement, which is especially useful in feature extraction applications like medical imaging for diagnosis. They are not commonly used in processing color images probably because their strong contrast enhancement may lead to excessive noise or artifacts and cause the image to look unnatural.

In the field of computer graphics [24-29], various algorithms have been developed to deal with a similar problem: how to display a high dynamic range image on a display device with limited dynamic range. However, the techniques developed in both areas may not be shared due to the following two reasons. First, in image processing, the input is an image that has been degraded and recorded by an imaging device of limited dynamic range. In computer graphics, the input is an undistorted array of simulated real-world luminance with high dynamic range. Second, in image processing, the task is to enhance the visibility of imperfect images by compressing the dynamic range and improving the contrast. The subjective correspondence with the original view of the scene generally cannot be maintained. In computer graphics, however, the subjective correspondence needs to be maintained. Visibility and contrast are simulated to produce visually accurate, not enhanced (changed) images.

Larson et al. [24] developed a tone reproduction operator that preserves visibility of high dynamic range scenes using a new histogram adjustment technique, based on the population of local adaptation of luminance in a scene. To match a subjective viewing experience, the method incorporates models for human contrast sensitivity, glare, spatial acuity, and color sensitivity. This technique and other similar techniques, developed for computer graphic applications, is not suitable for image enhancement due to its global processing approach and lack of contrast enhancement, which may lead to feature loss or degradation at some areas in the image. Chiu et al. [25] were the first in computer graphics to consider that tone mapping should be spatially non-uniform; in other words, the tone mapping should be neighborhood dependent. This method and MSRCR are similar in spirit. Therefore, they share the same problem of halo artifacts. Schlick [26] proposed a simple and fast non-uniform tone mapping scheme which is dependent on the intensity of each pixel itself. It is not adaptive enough to account well for contrast enhancement in image processing.

In order to eliminate the notorious halo effect, Tumblin and Turk [27] introduced the low curvature image simplifier (LCIS) hierarchical decomposition of an image. Each stage in this hierarchy is computed by solving partial differential equations inspired by anisotropic diffusion. At each hierarchical level, the method segments the image into smooth (low-curvature) regions while stopping at sharp discontinuities. The hierarchy describes progressively smoother components of the image intensity (luminance). A set of image details are obtained by subtractions between images at adjacent levels. Tumblin and Turk attenuate the smooth components and reassemble the images to create a low-contrast version of the original while compensating for the wide changes in the illumination field. This method drastically reduces the dynamic range, but tends to overemphasize fine details. The algorithm is computationally intensive and requires the selection of at least 8 different parameters. Pattanaik et al. [28] presented a tone mapping algorithm, which is made more realistic by incorporating human visual perception behavior into the model. They performed the image decomposition by the Laplacian pyramid (difference-of-Gaussian pyramid) approach. This technique produces dynamic range compressed images with good tonality and accurate color rendering. However, the halo effect in the images produced by their algorithm is much more severe than that produced by MSRCR.

2.2 Overview of Wavelet Transform Based Image Processing

The fourier transform has long been the mainstream of transform-based signal processing since the late 1950s. However, a more recent transformation, called the wavelet transform, has shown great promise for various image processing applications which include image denoising, contrast enhancement, compression, segmentation and classification, etc. Although wavelet transform cannot replace Fourier transform, it really has some unique signal processing capabilities which are not possessed by Fourier transform. Unlike the Fourier transform whose basis functions are sinusoids, wavelet transforms are based on small waves of varying frequency and limited duration. This allows wavelet coefficients to provide both spatial (or temporal) and frequency information (i.e. space-frequency or time-frequency analysis), whereas the non-local Fourier transform gives only frequency information. Mallat [30] first proposed multiresolution analysis (MRA) by using the wavelet transform to provide a new powerful approach to signal processing and analysis. Multiresolution theory unifies techniques from several fields, including subband coding from signal processing [31], quadrature mirror filtering (QMF) from speech recognition [32], and pyramidal image processing [33]. Due to the close links to these techniques and its specials properties, the wavelet transform has been investigated for many applications which include prediction and filtering [34], estimation [35], image denoising [36], image coding or compression [37], image enhancement [38], detection and classification [39], and synthesis [40]. Most importantly, the wavelet transform has been adopted in the state-of-art image and video compression standards like JPEG-2000 and MPEG-4 [41, 42].

2.2.1 Wavelet transform based image denoising

The visual quality of images is frequently corrupted by noise introduced during the image acquisition and transmission phases. The goal of image denoising is to remove the noise, for both aesthetic and image processing reasons, while retaining important image features as much as possible. Traditionally, denoising is achieved in the spatial domain by approaches such as Wiener filtering, which is the optimal estimator in the sense of mean squared error (MSE) for stationary Gaussian processes. However, the requirement for a stationary and accurate statistical model of the underlying process leads to poor performance on natural images which violate these assumptions. As a result, applying the Wiener filter to natural images tends to produce blurred edges. In practice, adaptive methods [43, 44] are mostly used. Adaptive methods are fast and can effectively suppress noise for many natural images. More importantly, their adaptivity allows them to work for non-stationary processes. The main problem with such methods is their assumption that natural images are independent random processes, which usually is not true.

Since the last decade, wavelet transform has been studied extensively for suppressing noise in natural images [36, 45-48] because of its effectiveness and simplicity. It is now well-known that wavelet transforms with some regularity have strong decorrelation abilities, and can well represent an image with just a few large coefficients. Therefore, it is far more reasonable to assume that the wavelet coefficients are more independent than original spatial domain pixels. This explains why impressive denoising results have been achieved by simply thresholding or shrinking each wavelet coefficient independently [36, 45, 46]. Indeed, this kind of approach has much better results than the traditional methods [43, 44], both subjectively and objectively.

The pioneering work by Donoho *et al* [49] inspired the research on wavelet transform based denoising. They tried to constrain the smoothness of denoised signals and to find the asymptotically optimal minimax regression estimator, which turned out to be a simple thresholding process but outperformed many previous denoising methods (e.g. [43, 50]). This idea was later developed by Starck *et al* [51] for image denoising. Donoho's work was mainly based on deterministic signal analysis. Because noise is naturally a stochastic process, many researchers adopted probabilistic models and Bayesian inference approaches to achieve denoising in the wavelet domain. The representative models include the non-Gaussian independence model [36], hidden Markov tree model (HMT) [52], Gaussian scale mixture model (GSM) [45] and Markov field model [46]. These ideas for image denoising are summarized as follows:

1. Non-Gaussian independence models, typically generalized Gaussian distribution (GGD) and Gaussian mixture models, are used to characterize the distribution of wavelet coefficients [36, 47, 53, 54]. Although the resulting denoising equations are relatively simple, they are often effective in terms of MSE and impressive denoising results have been obtained. The main problems are the Gibbs-like artifacts around edge areas due to ignoring strong remaining correlations in these areas.

2. Gaussian models exploiting inter- and/or intra- scale correlations are used by [55-58]. For image denoising, a significant decrease of MSE (relative to the above independence model) was reported when noise variances are relatively high. However, these models ignore the non-Gaussian nature of wavelet coefficients and due to limited amount of data available for estimating the needed parameters, artifacts can often be seen due to parameter estimation errors.

3. Non-Gaussian models exploiting inter- and/or intra- scale correlations are proposed in, for example [45, 59, 60, 52]. These models include the well-known hidden Markov tree (HMT) models [52, 61, 62], Markov field models [59, 60, 63, 64] and scale mixture Gaussian (GSM) models [45]. These models are sophisticated and have very powerful representation strengths. They can be used not only for image denoising, but also for segmentation [65], detection [39], and enhancement [38]. However, for image denoising, the formulations using these models usually do not result in closed-form solutions and high-complexity numerical iterative methods have to be used. Another problem with these models is none of them can represent edge areas very well and thus have unavoidable artifacts in these areas. Recently, Kingsbury [66-68] proposed DT-CWT which has approximate shift invariance and better directional selectivity compared to other types of wavelet transforms. Therefore, DT-CWT is believed to exhibit better performance in image processing than other wavelet transforms, and some improved results have been obtained [69-71].

2.2.2 Wavelet transform for image enhancement

Wavelet-based image enhancement is mainly used to enhance the perceptual sharpness of an image but this method can generally improve the image local contrast too. Due to its spatial-frequency analysis capability, wavelet transform is well suitable for augmenting the edge features in images. In a standard scenario of an image enhancement experiment for improving image sharpness, for example, when a given input image is blurred with known degradation model, restoration techniques can be applied together with other frequency enhancement techniques [72, 73]. The enhancement scheme described here can be applied as an additional enhancement utility.

Given a blurred image, classic image enhancement methods will try to recover the lost high-frequency components so that lost image features can be completely or partially restored and the processed image will look sharper and more pleasing to human observers. Traditionally this was performed by so-called unsharp filters [73], which are linear (and generally shift-invariant) processors. It is well known that linear shift invariant (LSI) filters modify only the existing frequencies but cannot generate new frequency components, and thus cannot recover the lost high-frequency components in principle. Nonlinear filtering methods were also studied by several authors [74-77]. However, so far there are only *ad hoc* solutions and designing general-purpose nonlinear filters remains difficult.

Recently, several multiscale image enhancement approaches have been proposed with interesting results [78, 79, 38, 80]. Image enhancement in a multiscale context can be considered as the estimation of coefficients in high frequency subbands based on those in lower-frequency subbands. All of these approaches attempt to utilize the inter-scale dependency (mainly related to edges) to extrapolate lost high-frequency components. Greenspan *et al.* [78] and Burt *et al.* [79] used zero-crossings of the second derivative of smoothed images to locate edges, and based on the ideal step edge model they estimated the inter-scale relations of edges. They then used these relations to estimate edges in finer scales from those of the low-frequency subbands. Kinebuchi and Woo [38, 80] assumed a different approach: they first used a hidden Markov tree model (HMT) [52, 61] to infer the probability of each hidden state and corresponding variances. Then a Gaussian mixture model (GMM) (corresponding to the hidden states) is used for each wavelet coefficient and the wavelet coefficients in the highest subband are generated randomly (by sampling) using the estimated state probabilities and variances. In estimating variances, the property of exponential decay of variances was assumed [81] with roughly estimated exponents.

The image enhancement techniques previously reviewed are mainly used to recover lost high frequency components, but may also help improve local image contrast. However, those enhancement techniques are different from contrast enhancement which needs no restoration of lost high frequency components but which do need enhancement of certain existing high frequency components. Multi-scale image contrast enhancement has been proposed and implemented using either wavelet [82] or curvelet [83] transform. In [82], Velde proposed enhancing the faintest edges and keeping untouched the strongest by modifying the wavelet coefficients in a similar way as the coefficient thresholding in wavelet denoising. However, multi-scale contrast enhancement technique based on wavelet transform has not been extensively investigated due to the limitations of real DWT. Due to the improved capabilities of DT-CWT, better contrast enhancement results are expected by using DT-CWT instead of normal DWT.

2.3 Overview of Multiresolution Image Fusion

Image fusion can be broadly defined as the process of combining multiple input images into a smaller collection of images, usually a single one, which contains the relevant information from the inputs. In order to better understand a scene from an image, not only the position and geometry, but more importantly, the semantic interpretation matters. In this context, the word relevant should be considered in the sense of 'relevant with respect to the task the fused images will be subject to', in most cases high-level tasks such as interpretation or classification. Since the fused image generally possesses more scene information than any single input image, image fusion can also be considered as an image enhancement process. For example, fusing multi-focus images (e.g. images with various focal depths) can produce a fused image which exhibits clear details of all objects in the scene.

The image fusion process can take place at different levels of information representation. A common categorization method is to distinguish between pixel, feature (or region) and symbol level [84], although indeed these levels can be combined themselves [85]. Image fusion at pixel-level represents the lowest processing level referring to the merging of measured physical parameters (usually pixel intensity) [86, 87]. It generates a fused image in which each pixel is often determined from a set of pixels (a window) in the various sources. Fusion at feature-level requires first the extraction (e.g. by image segmentation procedures) of the features contained in the input images [88, 89]. Those features can be identified by characteristics such as size, shape, contrast, and texture. The fusion is thus based on those extracted features and enables the detection of useful features with higher confidence. Fusion at symbol level allows the information to be effectively combined at the highest level of abstraction [90, 91]. The input images are usually processed individually for information extraction and classification. This results in a number of symbolic representations which are then fused according to decision rules which reinforce common interpretation and resolve differences. The choice of the appropriate level depends on many different factors such as data sources, application and available tools. At the same time, the selection of the fusion level determines the necessary pre-processing involved. For instance, fusing data at pixel-level requires co-registered images at subpixel accuracy because the existing fusion methods are sensitive to misalignment.

Currently, many image fusion applications employ pixel-based methods. The advantage of pixel-level image fusion is that the source images contain the original physical information, and the fused image generally exhibits more scene information than those produced by fusion at feature level. Furthermore, the algorithms are rather easy to implement and time efficient. It should be noted that an important pre-processing step in pixel-fusion methods is image registration, which ensures that the information from each source is referring to the same physical structures in the real-world. However, for many image fusion applications, it is more meaningful to combine objects rather than pixels. As an intermediate step from pixel based toward object-based fusion schemes, one may consider region-based approaches. Such approaches have the additional advantage that the fusion process becomes more robust and avoids some of the well known problems in pixel-level fusion such as blurring effects and high sensitivity to noise and misregistration.

MR analysis techniques have been extensively used in image fusion due to its strong capability of representing multi-scale image features. The first MR image fusion approach was proposed in the literature by Burt [92]. His implementation used a Laplacian pyramid and a sample-based maximum selection rule with activity measure defined by the magnitude of the coefficient. Toet [93] presented a similar algorithm but using the ratio-of-low-pass pyramid. His approach is motivated by the fact that the HVS is based on contrast, and therefore, a fusion technique which selects the highest local luminance contrast is likely to provide better details to a human observer. Another variation of this scheme is obtained by replacing the linear filters by morphological ones [94, 95]. Burt and Kolczynski [86] proposed to use the gradient pyramid together with a combination algorithm that is based on an activity and a match measure. In particular, they define the activity of as a local energy measure and match measure as a normalized correlation computed in a window with a size of either 1×1, 3×3 or 5×5. The combination process is a weighted average of corresponding pixels with the weights being determined by the decision process for each pixel location at each level and band with a threshold. The decision process works in such a way that in case of a poor match (no similarity between the inputs), the source coefficient having the largest activity will yield the composite value, while otherwise, a weighted sum of the sources coefficients will be used. The authors claim that this approach provides a partial solution to the problem of combining components that have opposite contrast, since such components are combined by selection. In addition, the use of window-based (versus pixel-based) operations and the gradient pyramid provide greater stability in noise, compared to the Laplacian pyramid based fusion.

Ranchin and Wald [96] presented one of the first wavelet-based fusion systems. This approach is also used by Li et al. [97]. Their implementation considers the maximum absolute value within a window as the activity measure associated with the pixel centered in the window. For each position in the transform domain, the maximum selection rule is used to determine which of the inputs is likely to contain the most useful information. This results in a preliminary decision map which indicates, at each position, which source should be used in the combination. This decision map is then subject to a consistency verification. In particular, Li et al. apply a majority filter in order to remove possible wrong selection decisions caused by impulsive noise. The authors claim that their scheme performs better than the Laplacian pyramid-based fusion due to the compactness, directional selectivity and orthogonality of the wavelet transform. Wilson et al. [98] used a DWT fusion method and a perceptual-based weighting based on the frequency response of the HVS. Indeed, their activity measure is computed as a weighted sum of the Fourier transform coefficients of the wavelet decomposition, with the weights determined by the contrast sensitivity. They define a perceptual distance between the sources and use it together with the activity to determine the weights of the wavelet coefficients from each source. Actually this perceptual distance is directly related to the matching measure: the smaller the perceptual distance, the higher the matching measure. The final weighting process is guided by the perceptual distance and activity measure with a distance threshold. When the perceptual distance is larger the threshold, maximum selection rule is used to determine which coefficient will be chosen. Otherwise, a weighted averaging process will be applied with the weight being determined by the relative relation between the two coefficients. In the experimental results presented by the authors, the fused images obtained with their method are visually better than the ones obtained by fusion techniques based on the gradient pyramid or the ratio-of-low-pass pyramid.

Koren et al. [99] used a steerable wavelet transform for the MR decomposition. They advocate their choice because of the shift-invariance and no-aliasing properties this transform offers. For each frequency band, the activity is a local oriented energy. Only the components corresponding to the frequency band whose activity is the largest are included for reconstruction. Liu et al. [100] also used a steerable pyramid but rather than using it to fuse the source images, they fuse the various bands of this decomposition by means of a Laplacian pyramid. In [101], Rockinger considered an approach based on a shift-invariant extension of the DWT. The detail coefficients are combined by a maximum selection rule, while the coarse approximation coefficients are merged by averaging. Due to the shift-invariance representation, the proposed method is particularly useful for image sequence fusion, where a composite image sequence has to be built from various input image sequences. The author shows that the shift-invariant fusion method outperforms other MR fusion methods with respect to temporal stability and consistency.

Pu and Ni [102] proposed a contrast-based image fusion method using also the DWT. They measure the activity as the absolute value of what they call directive contrast which is the ratio between the coefficient in one subband and the coefficient at the same position in the reference subband, and they use a maximum selection rule as the combination method of the wavelet coefficients. They also proposed an alternative approach where the combination process is performed on the directive contrast itself. Li and Wang [103] examined the application of discrete multiwavelet transforms to multisensor image fusion. The composite coefficients are obtained through a pixel-based maximum selection rule. The authors showed experimental results where their fusion scheme performs better than those based on comparable scalar wavelet transforms. Another MR technique is proposed by Scheunders [104] where the fusion consists of retaining the modulus maxima [105] of the wavelet coefficients from the different bands and combining them. Noise reduction can be applied during the fusion process by removing noise related modulus maxima. In the experiments presented, the proposed method outperforms other wavelet-based fusion techniques.

CHAPTER III

NONLINEAR IMAGE ENHANCEMENT IN THE SPATIAL DOMAIN

In this chapter, we investigate the image visibility improvement in the spatial domain based on image processing of a single image. Two new nonlinear image enhancement algorithms are developed: illuminance-reflectance model (I-R) and adaptive integrated neighborhood dependant approach for nonlinear enhancement (AINDANE). Both of them have to deal with the three major technical problems which have been addressed in the previous chapter: dynamic range compression, local contrast enhancement, and color consistency. The derivation of both algorithms will be introduced in detail and discussed with experimental results and performance comparison with other techniques.

3.1 Illuminance-Reflectance Model Based Algorithm

Illuminace-reflectance model is a physical description of the creation of a radiance map of real world scenes. It divides the object radiance into two parts: the light intensity (illuminance) incident on object surface and the light refection properties (reflectance) of the object surface. The separation of illuminance and reflectance provides a method to process images for the purpose of obtaining an improved visual perception of those scenes for human viewers or computer vision algorithms.

An effective method for estimation of illumiance and reflectance is developed. The proposed algorithm uses a windowed-inverse-sigmoid function to adaptively compress the dynamic range of the illuminance while the reflectance is unchanged to maintain important image features. In addition, the mid-frequency components of the reflectance are enhanced using a new center-surround method to improve the image contrast to obtain a well-balanced result between global lightness rendition and local contrast quality. The low-complexity and adaptiveness enable the new algorithm to be suitable for real time and mobile applications while consistent and robust performance can be achieved. The applications of this algorithm are presented in Chapter VI.

3.1.1 I-R Algorithm

The structure of the I-R algorithm is illustrated in Fig. 3.1. It is composed of four major steps: (1) illuminance estimation and reflectance extraction; (2) adaptive dynamic range compression of illuminance; (3) adaptive mid-tone frequency components enhancement and (4) image restoration, which combines illuminance and reflectance to recover the enhanced intensity image and then performs color restoration to obtain color images.

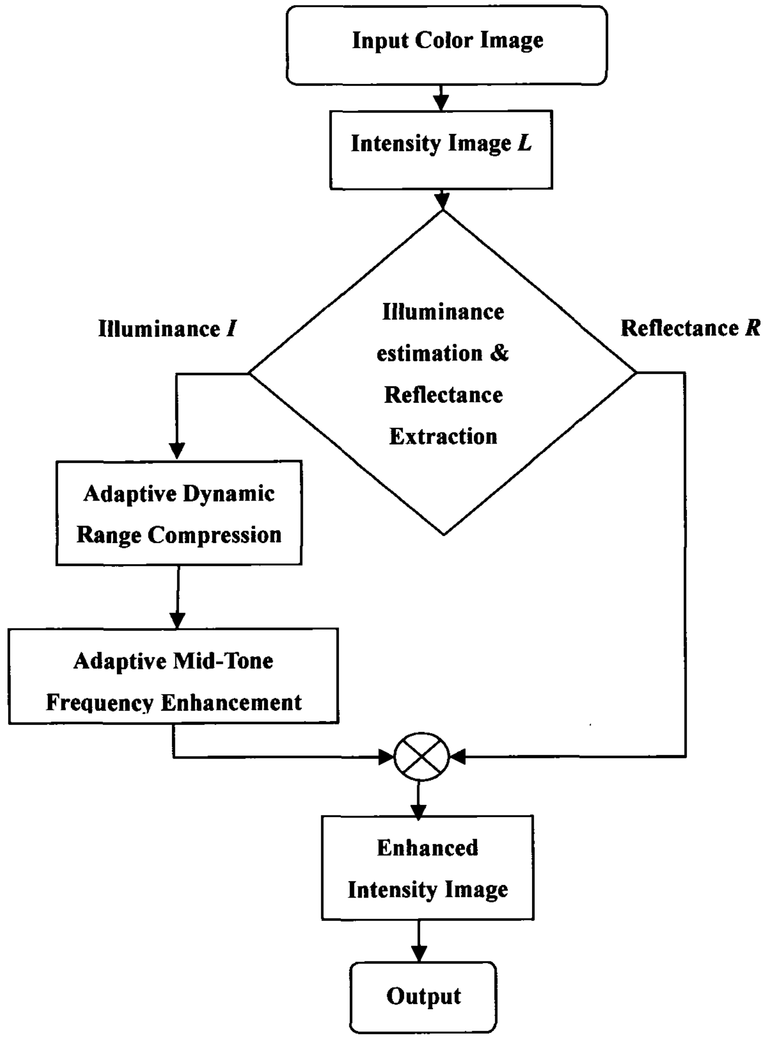


Figure 3.1 Structure of the proposed algorithm for color image enhancement.

For input color images, the intensity image $L(x,y)$ can be obtained using either one of the following two methods:

$$L(x,y) = 0.2989 \cdot r(x,y) + 0.587 \cdot g(x,y) + 0.114 \cdot b(x,y) \tag{3.1a}$$

or

$$L(x,y) = \max\left[r(x,y), g(x,y), b(x,y)\right] \tag{3.1b}$$

where r, g and b are the RGB components of color images in RGB color space. The first method, the NTSC standard, is commonly used for converting color images to grayscale images. The second method is in fact the definition of the value (V) component in HSV color space. The latter method is used in the I-R algorithm. The reason for this will be explained later in section 3.1.2.

The I-R algorithm is based on the commonly accepted assumptions about image formation and human vision behavior. First, the image intensity $L(x, y)$ can be simplified and formulated as:

$$L(x,y) = I(x,y)R(x,y) \tag{3.2}$$

where $R(x, y)$ is the reflectance and $I(x, y)$ is the illuminance at each position (x, y) in the scene. Second, the illuminance I is assumed to be contained in the low frequency components of the image while the reflectance R mainly represents the high frequency components of the image. This assumption can be easily understood considering that R generally varies much faster than I in most regions of an image with few exceptions, like the shadow boundaries where an abrupt change of I exists. In addition, in a real world scene, the dynamic range of the illumination variation can be several orders larger than the dynamic range of the reflectance. Therefore, compressing the dynamic range of the illuminance is an effective way for image enhancement. Finally, there is a widely accepted conclusion about human vision, namely that human eyes respond to local changes in lightness rather than to global brightness levels. Therefore, it is possible to keep the visually important features represented by reflectance while the image's dynamic range (induced mainly by illuminance) can be compressed.

Accurate estimation of illuminance of a scene from an image is a difficult task. Many techniques have been developed to deal with this problem [5,8,9]. In the I-R algorithm, the Gaussian low-pass filtered result of the intensity image is used as the estimation of the illuminance. In the spatial domain, this filtering process is actually a 2D discrete convolution with a Gaussian kernel, which can be mathematically expressed as

$$I(x,y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} L(m,n)G(m+x,n+y) \tag{3.3}$$

where I is the illuminance, L is the original intensity image and G is the 2D Gaussian function with size $M \times N$. Gaussian kernel (mask) G is defined as

$$G(x,y) = q \cdot \exp\left(\frac{-\left(x^2+y^2\right)}{c^2}\right) \tag{3.4}$$

where q is the factor for normalizing G by

$$\sum_x \sum_y q \cdot \exp\left(\frac{-\left(x^2+y^2\right)}{c^2}\right) = 1 \tag{3.5}$$

and c is the scale (Gaussian surround space constant), which determines the size of the neighborhood. In our algorithm, c = 2~5 is commonly used. Finally, illuminance is normalized as

$$I_n(x,y) = \frac{I(x,y)}{255} \tag{3.6}$$

for 8-bit depth images.

After the illuminance I is obtained using Eq. (3.3), the reflectance R is computed using Eq. (3.2). One example showing the results of I and R from an image is provided in Fig. 3.2(b) and 3.2(c) with the original intensity image shown in Fig. 3.2(a). It can be observed from those images that the illuminance comprises the low- and mid-tone frequency information of the image. Because this is not the real illuminance as defined in physics but an approximation, this estimated illuminance actually is a combination of both illuminance and the low /mid- tone frequency components from the reflectance. The visually salient features (high frequency reflectance) and a small part of illumination information are contained in the reflectance R in which the major illumination effect is removed. Therefore, important image features can be retained even after the dynamic range compression of illuminance. Based on those observations, reflectance and illuminance are also regarded as details and base.

The adaptive dynamic range compression of illuminance is realized in our algorithm using the windowed inverse sigmoid (WIS) function. The sigmoid function is defined as

$$f(v) = \frac{1}{1+e^{-av}} \tag{3.7}$$

This function is used as the intensity transfer function for dynamic range compression by performing the following steps described in Eqs. (3.8) - (3.10).

$$I_n' = I_n[f(v_{\max}) - f(v_{\min})] + f(v_{\min}) \tag{3.8}$$

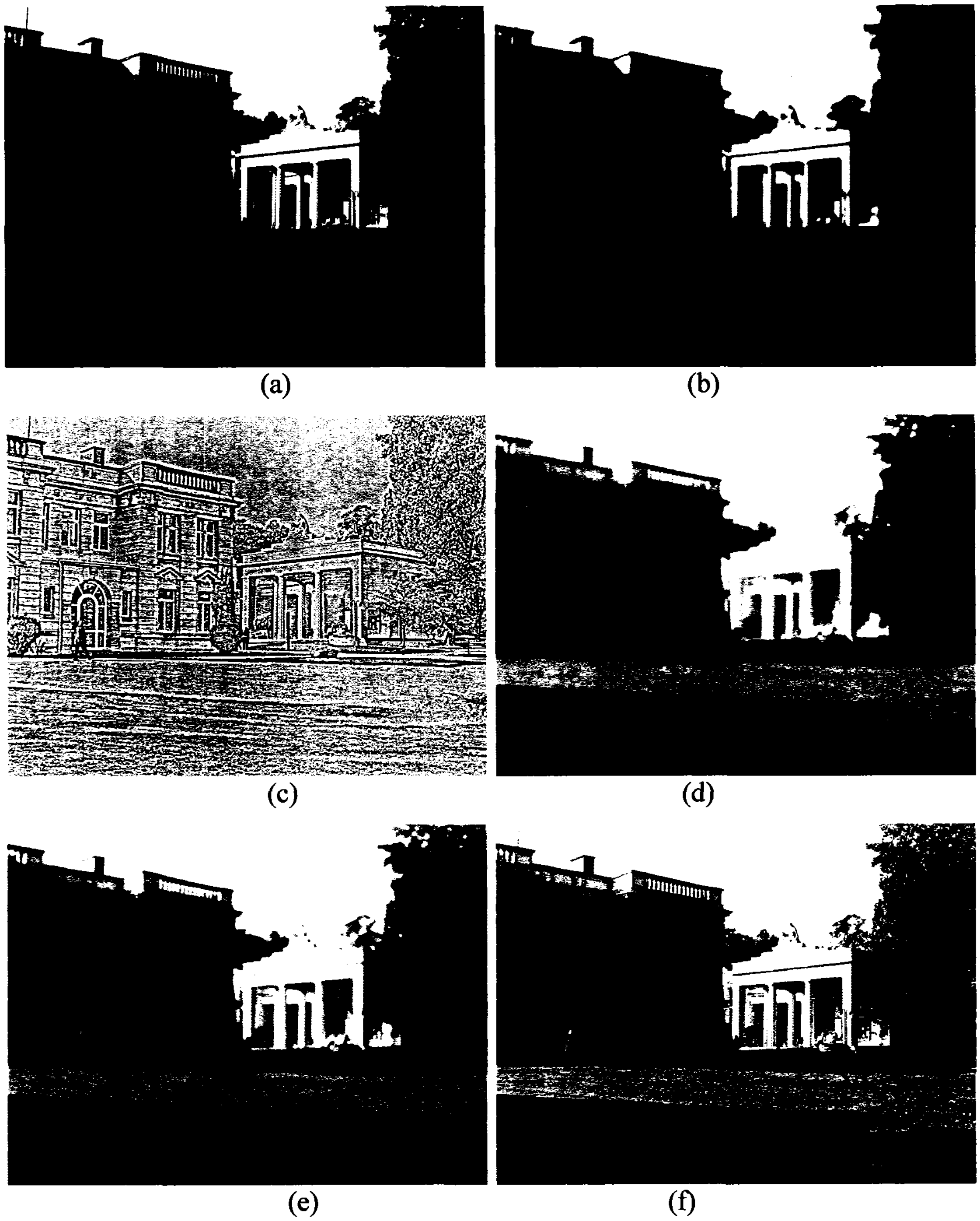


Figure 3.2 Flowchart and intermediate results of I-R algorithm:
(a) original image; (b) estimated illuminance; (c) extracted reflectance; (d) dynamic range compressed illuminance; (e) mid-tone frequency components enhancement from (d); (f) enhanced image.

$$I_n'' = \frac{1}{a}\ln\left(\frac{1}{I_n'} - 1\right) \tag{3.9}$$

$$I_{n,enh} = \frac{I_n'' - v_{\min}}{v_{\max} - v_{\min}} \tag{3.10}$$

where Eq. (3.8) is for linearly mapping the input range [0 1] of normalized illuminance I_n to the input range $[f(v_{min})\ f(v_{max})]$ for WIS. Eq. (3.9) is the inverse sigmoid function. Eq. (3.10) is applied to normalize the output illuminance I_n'' to range [0 1]. Parameters v_{max}, and v_{min} are used to tune the curve shape of the transfer function. One example of the dynamic range compressed illuminance $I_{n,enh}$ is shown in Fig. 3.2(d).

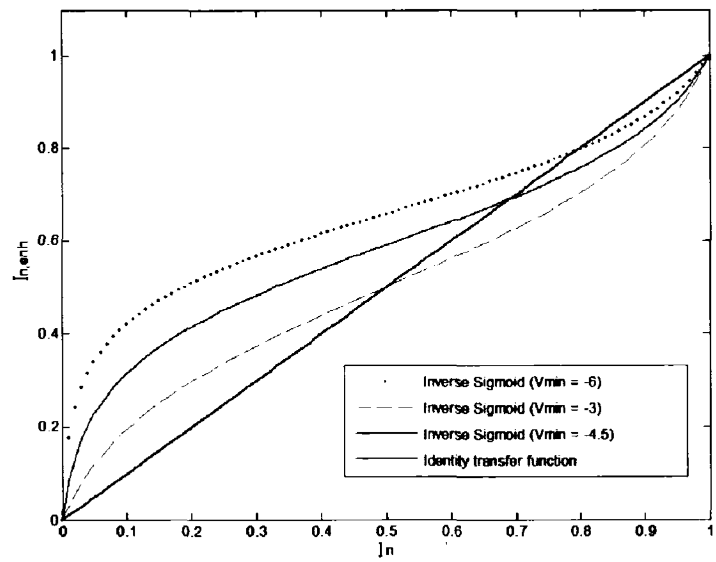


Figure 3.3 WIS with different parameters applied for intensity transformation.

A set of various curve shapes of WIS transfer function is provided in Fig. (3.3) with $a = 1$ in Eq. (3.7). Those curves are produced by Eqs. (3.8) - (3.10) with different parameters: $v_{max} = 3$, $v_{min} = -6$ for the dotted red curve; $v_{max} = 3$, $v_{min} = -3$ for the dashed red curve; and $v_{max} = 3$, $v_{min} = -4.5$ for the solid red curve. Identity transfer function (blue line) is also provided for comparison. The inverse sigmoid function can be used to decrease the intensity of the over-bright pixels while dramatically increasing the intensity of dark pixels. Therefore, dynamic range compression of the illuminance is realized.

The parameter v_{min}, v_{max} and a can be manually adjusted by users to tune the dynamic range compression. In the I-R experiments, v_{max} and a are generally set to 3 and 1, respectively, while the value of v_{min} is also made image dependent to obtain adaptive control over the dynamic range compression. In our algorithm, L_{max} is the maximum intensity of the original image, v_{min} is empirically determined by the global mean L_m of the intensity image L based on image enhancement experiments, using

$$v_{\min} = \begin{cases} -6 & for \;\; L_m \le 25\%L_{max} \\ \dfrac{L_m - 70}{80} \times 3 - 6 & for \;\; 25\%L_{max} < L_m < 60\%L_{max} \\ -3 & for \;\; L_m \ge 60\%L_{max} \end{cases} \tag{3.11}$$

It has been noted that the illuminance also contains the mid-tone and low frequency components from reflectance which has been degraded during the dynamic range compression. For an original image with low contrast or a slowly-varying reflectance map, the degradation of mid-frequency features may be rather severe and thus it can impair the quality of output images. Therefore, a new center-surround type of local contrast enhancement method is proposed to compensate for this degradation. This mid-tone frequency enhancement is carried out using the following two equations

$$I'_{n,enh}(x,y) = I_{n,enh}(x,y)^{E(x,y)} \tag{3.12}$$

where the exponent is given by

$$E(x,y) = R_e(x,y)^P = \left(\frac{L_{conv}(x,y)}{L(x,y)} \right)^P . \tag{3.13}$$

$I'_{n,enh}(x, y)$ is the illuminance after mid-tone frequency enhancement and $R_e(x, y)$ is the ratio between $L(x, y)$ and its low-pass version $L_{conv}(x, y)$ which is computed through the same operations as in Eqs. (3.3) through (3.5), but with a larger scale c (representing a lower cut-off frequency) that is 10 by default. One example of illuminance image after dynamic range compression and mid-tone frequency enhancement is shown in Fig. 3.2(e). Similar to v_{min}, P is also made image dependent to achieve adaptive control of the process, and its value is determined by the global standard deviation σ of the input intensity image $L(x, y)$ as:

$$P = \begin{cases} 2 & for \;\; \sigma \le 30 \\ -0.03\sigma + 2.9 & for \;\; 30 < \sigma \le 80 \\ 1/2 & for \;\; \sigma > 80 \end{cases} . \tag{3.14}$$

According to this definition, P and σ have a linear relationship with $\sigma \in (30, 80]$. This empirical relationship is determined based on a large number of image enhancement experiments and is found quite robust for natural images. Here, the global standard deviation σ of $L(x, y)$ is considered as an indication of the contrast level of the original intensity image. It should be noted that P also can be manually adjusted by users to tune the contrast enhancement process.

The mid-tone frequency components enhancement described in Eqs. (3.12) and (3.13) can also be regarded as an intensity transformation which is illustrated in Fig. 3.4. Since $I_{n,enh}$ is normalized to 1, $I_{n,enh}(x, y)^{E(x, y)}$ will be larger than $I_{n,enh}(x, y)$ if $E(x, y)$ is less than 1 (e.g. the center pixel is brighter than surrounding pixels leading to $R(x, y) < 1$). Otherwise, if $E(x, y)$ is larger than 1 (e.g. the center pixel is darker than the surrounding pixels with $R(x, y) > 1$), $I_{n,enh}(x, y)^{E(x, y)}$ will be smaller than $I_{n,enh}(x, y)$. In this way, the local contrast (or details) of the compressed illuminance image can be improved. Here, the ratio $R(x, y)$ is obtained from the original intensity image $L(x, y)$ and its low-pass filtered result $L_{conv}(x, y)$ but not from the estimated illuminance I, since the reflectance's mid-tone and low frequency information has been lost or degraded in I.

The parameters of P and v_{min} are introduced in this algorithm to increase the adaptiveness and flexibility of the algorithm. Meanwhile, they can also improve the robustness of the enhancement algorithm to produce consistent results for images captured under various types of lighting conditions.

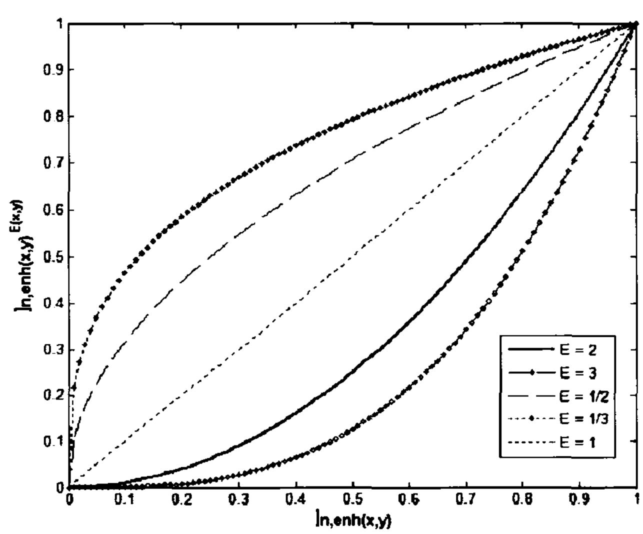


Figure 3.4 Transformation for mid-tone frequency enhancement of illuminance.

The enhanced intensity image L' is finally obtained by combining the illuminance $I'_{n,enh}$ and reflectance R using Eq. (3.2) as $L'(x,y) = I'_{n,enh}(x,y)R(x,y)$. One example of the enhanced intensity image is presented in Fig. 3.2(f). For color images, a linear color restoration process based on the chromatic information of the original image is applied to converting the enhanced intensity image to RGB color image. The RGB values (r', g', b') of the restored color image are obtained as

$$r' = \frac{L'}{L}r \qquad g' = \frac{L'}{L}g \qquad b' = \frac{L'}{L}b \tag{3.15}$$

where L and L' are the intensity of the original image and enhanced image, respectively. Thus, the color consistency between the original color image and the enhanced color can be achieved.

3.1.2 Results and Discussion

In this section, several important features of the algorithm will be discussed with experimental results. Then the performance of the proposed algorithm will be compared with other enhancement techniques both by visual analysis and by the verification in face detection. Finally, robustness of the I-R algorithm and statistical analysis of its performance will be introduced to illustrate the image enhancement effects.

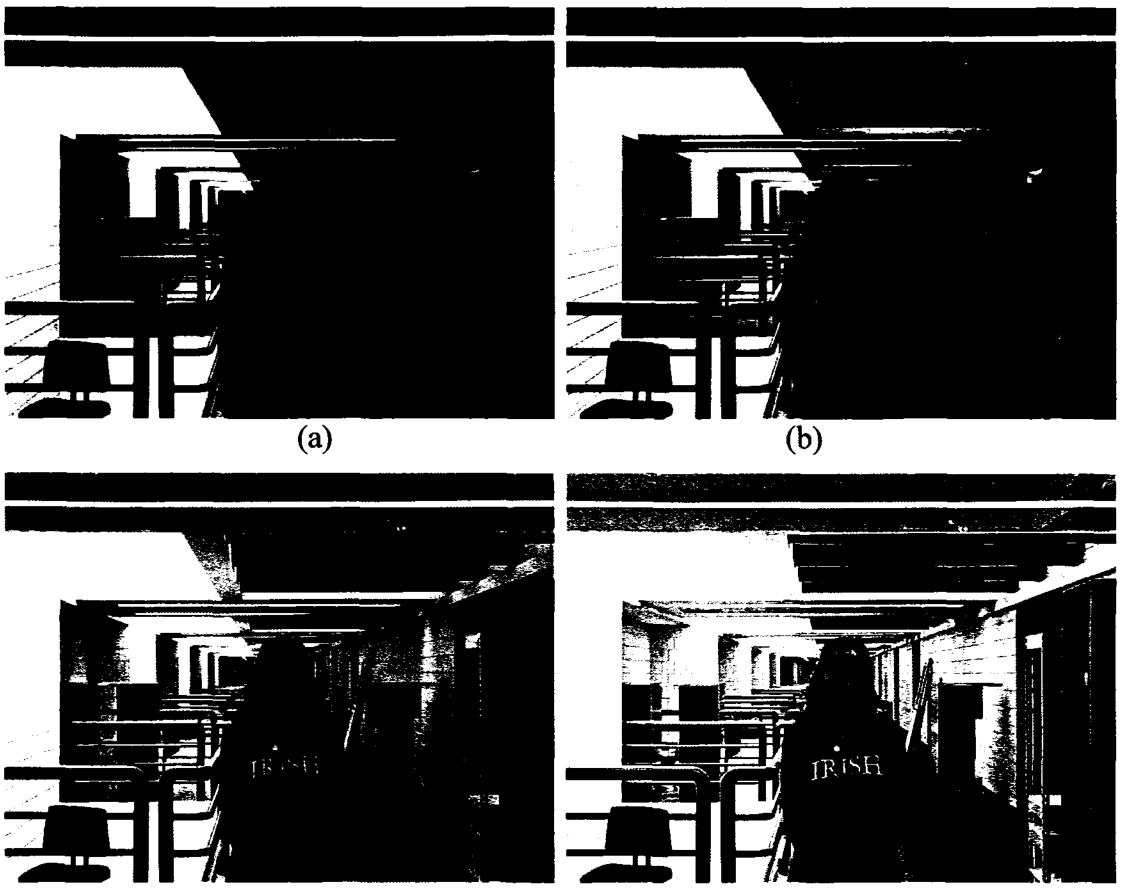
The proposed algorithm processes intensity/luminance images for image enhancement of color images. In fact, the algorithm can also be applied to treat each spectral band of color images to perform image enhancement. However, processing all color bands separately is definitely more complex and more time consuming than only processing the intensity image. Multi-band processing is not found significantly superior to intensity processing in terms of the quality of enhanced images. Therefore, intensity-only processing is still useful and preferred for certain image processing applications when rapid processing or low computational load is needed, like real time video processing.

In the proposed algorithm, two methods (NTSC and HSV) are applied to extract intensity images from RGB color images. These two methods produce slightly different results which can be seen in Fig. 3.5. The images produced by Eq. (3.1a) (NTSC) exhibit a lower brightness level in intensity images and produces inconsistent or shifted colors like the red range colors in the enhanced image from those in original image. Therefore, Eq. (3.1b) (HSV) is preferred for extraction of intensity images from color images.

(a) (b) (c) (d) (e)

Figure 3.5 (a) original image; (b) intensity image obtained by Eq. (3.1b); (c) intensity image obtained by Eq. (3.1a); (d) enhanced color image from (b); (e) enhanced color image from (c).

The curve shape of WIS can directly affect the dynamic range compression of the illuminance. Fig. 3.3 illustrates how the curve shape changes by adjusting the value of v_{min} manually or image dependent for adaptive control. The effect of v_{min} is illustrated in Fig. 3.6 with a sample image treated with various v_{min} values, one of which shown in Fig. 3.6(c) is set by Eq. (3.11). The image's illuminance enhancement can be readily adjusted by tuning v_{min}. The value of v_{min} set by Eq. (3.11) produces better results than manual adjustment results [Fig. 8(b) and 8(d)] without over or under enhancement of brightness.



(a) (b) (c) (d)

Figure. 3.6 Image enhancement with different ν_{min}: (a) original image; (b) enhanced image with $\nu_{min} = -3$; (c) enhanced image with $\nu_{min} = -4.9$, set by Eq. (3.11); (d) enhanced image with $\nu_{min} = -7$. Note: This sample image is provided by [16].

The parameter P in Eq. (3.14) is used to tune the enhancement for mid-tone frequency components to compensate the image contrast which is poor in the original image or has been significantly degraded due to dynamic range compression. The effect of P on image enhancement is illustrated in Fig. 3.7 with the sample image processed with various P values. P is determined by Eq. (3.14) for Fig. 3.7(c), which produces a more appropriate result compared to the other two enhanced images with manually set P values. For example, the wood grains are well enhanced without producing severe halo effect in Fig. 3.7(c), but an obvious halo effect is observed in Fig. 3.7(d).

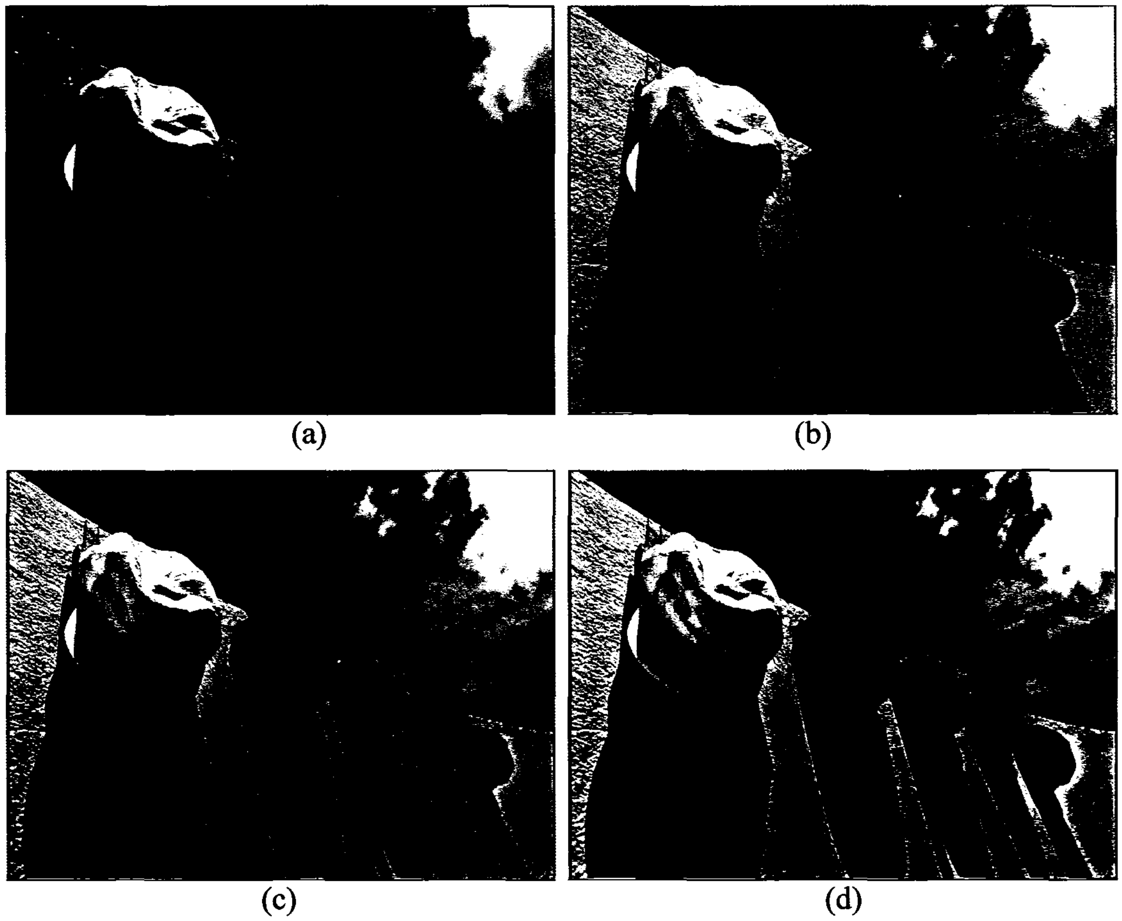
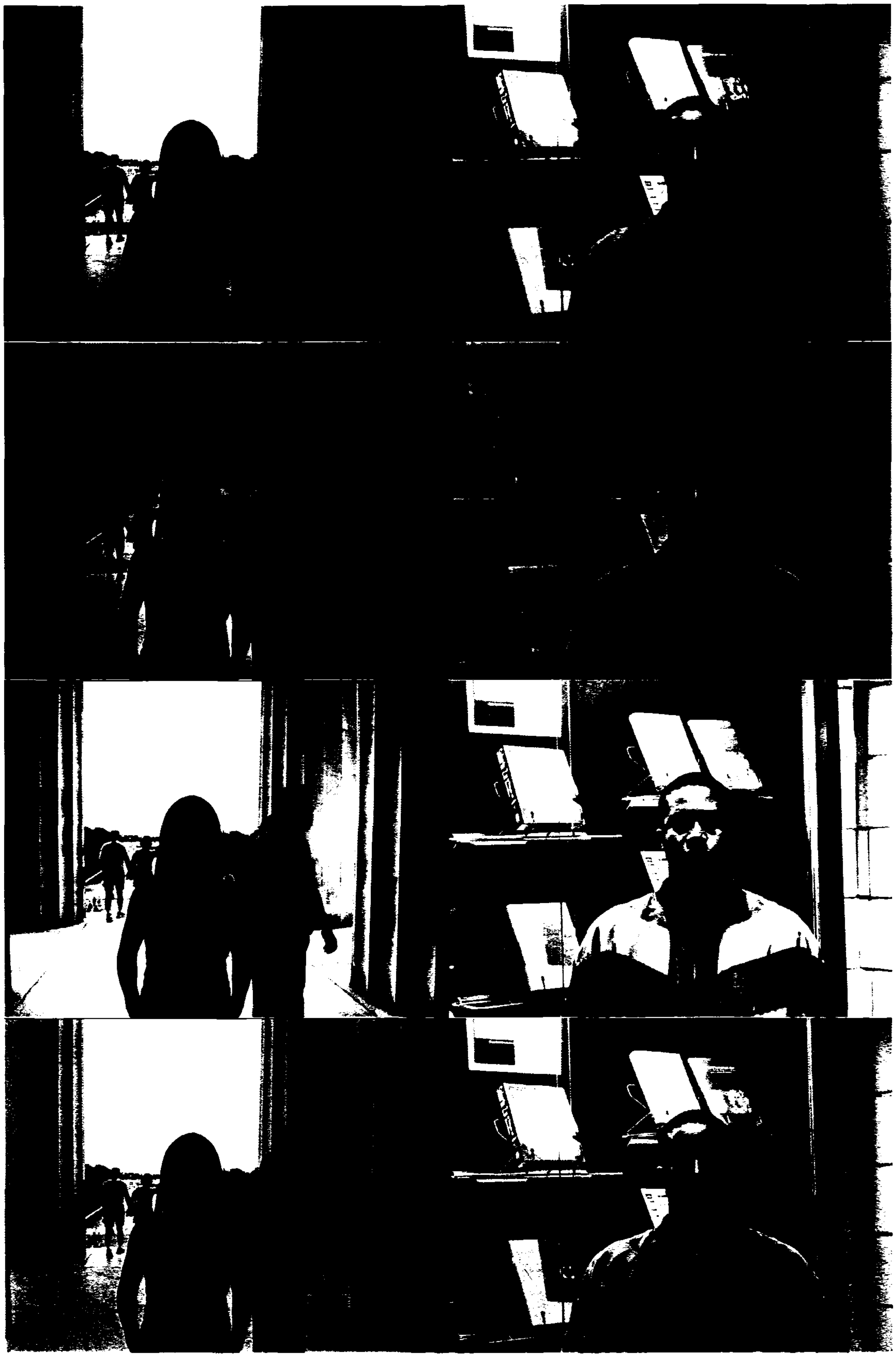


Figure 3.7. Image enhancement with different P: (a) original; (b) enhanced image with $P = 1/2$; (c) enhanced image with $P = 1$, set by Eq. (3.14); (d) enhanced image with $P = 2$.

In Fig. 3.8, two sets of sample images are provided for performance comparison among MSRCR, Retinex and the I-R methods. Retinex is realized using the Matlab® code provided in [2] by Frankle and McCann. A commercial software PhotoFlair® (www.truview.com) is used to implement MSRCR. It can be observed that the images processed by I-R method demonstrate a higher visual quality than those processed by MSRCR and Retinex. The I-R method yields better color accuracy and better balance between the luminance and contrast across the whole image due to its contrast.



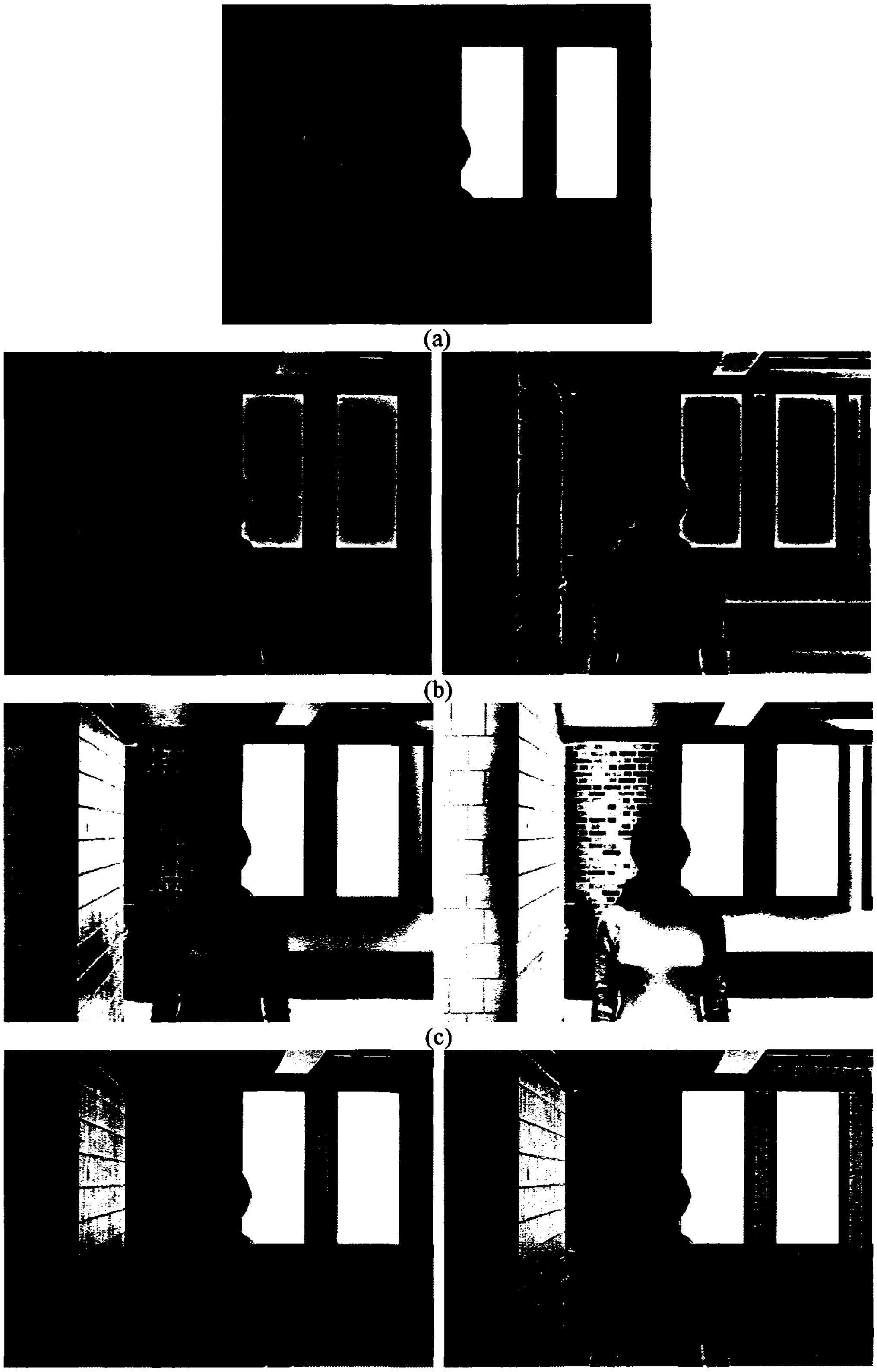
(a)

(b)

Figure 3.8 Comparison among different techniques: (a) original images (first row), by MSRCR (second row), by Retinex (third row), and by the I-R method (last row); (b) the images layout is the same as in (a).

However, its color correction capability creates incorrect colors (e.g. the right side enhanced results in both set (a) and set (b)) and the lightness rendition looks unnatural in some areas (e.g. the wall in the right side image in set (a)). In addition, both MSRCR and Retinex show more visible halo effect than the I-R algorithm, which also affects the color and lightness rendition. It is also found that both MSRCR and Retinex seem to have some difficulty enhancing images captured under non-uniform lighting conditions. They produce insufficient luminance enhancement in the dark regions with a high-lighted background, like the woman in the image in set (a) and the lawn and trees in set (b); or produce an over-enhancement effect to the bright regions from the original image, like the chair surface enhanced by Retinex in set (a). This problem consistently occurs in similar types of images. Moreover, MSRCR performs even worse in lightness rendition by incorrectly decreasing the brightness of some of the high-lighted areas (e.g. the book cover in set (a), the surface of the chair and the sky in set (b)), which yields a tonality much different from that of the original. On the other hand, our proposed algorithm generally performs well on the test images showing a more balanced result between luminance enhancement and local contrast enhancement without incorrect colors being created.

The robustness of this algorithm is evaluated by performing image enhancement on the input image twice: the first enhancement processing is carried out on the original image, then the second enhancement processing is performed on the output image obtained from the first enhancement. An example of this process is presented in Fig. 3.9 accompanied by the results produced by MSRCR and Retinex. Obviously, after double enhancement, the proposed algorithm produces the minimal change from the first enhancement result while image quality is degraded in images produced by MSRCR and Retinex, in which incorrect lightness, color rendition, halo effect, and image noise become much more visible after the second enhancement. It is believed that the higher robustness of our algorithm is due to its adaptiveness in processing.



(a)

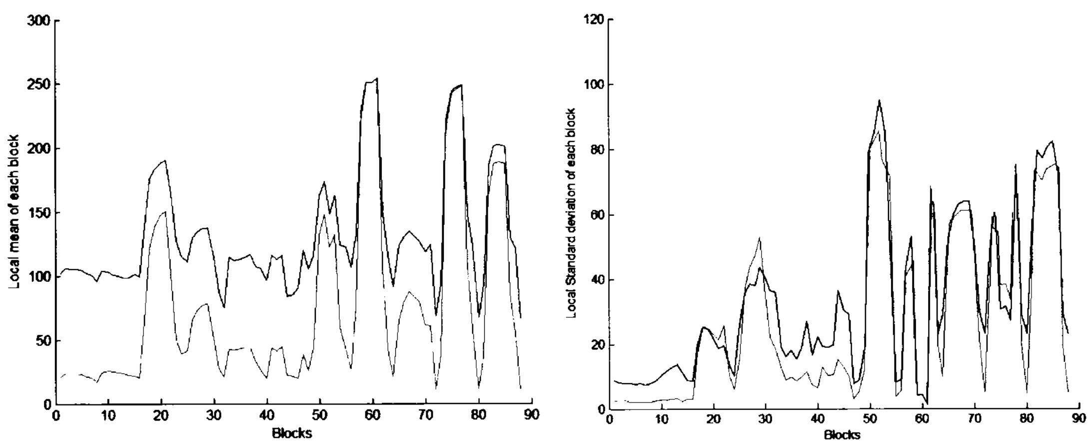
(b)

(c)

(d)

Figure 3.9 Robustness evaluation: (a) original image, (b) MSRCR, (c) Retinex and (d) the proposed algorithm. Left: enhanced results from the original images (enhanced once); right: enhanced results from the left column images (enhanced twice).

The statistical properties of images, image mean and the zonal standard deviation, are used to assess the visual quality of images in terms of brightness and contrast which are directly associated with those statistical parameters. The local brightness is measured by the image local mean while the local contrast is evaluated by taking the regional standard deviations. As shown in Fig. 3.10(a), we plotted the local mean and local standard deviations of all blocks (block size: 55 × 55 pixels) of a sample image (Fig. 3.9(a)) before and after image enhancement. It can be seen that in some regions of the sample image, the luminance enhancement is dramatic (eg. the shadows of the wall and the lady) but in some regions it is almost the same as the original luminance (eg. the regions of the window). This is exactly matched with the scheme of our nonlinear luminance enhancement stage, which can adaptively compress the dynamic range. For the local contrast enhancement, the effect is obvious in the regions where the luminance distribution is relatively uniform in the original image.



(a) (b)

Figure 3.10 Statistical analysis of the image [Fig. 3.9(a)] before and after image enhancement. (a): red curve and blue curve are the local mean of each block of the original and the enhanced image, respectively; (b): green curve and black curve are the local standard deviation of each block of the original and enhanced image, respectively.

3.2 Adaptive and Integrated Neighborhood Dependent Approach for Nonlinear Enhancement (AINDANE)

In this dissertation, another new non-linear image enhancement algorithm named AINDANE is proposed based on the essential features about the visibility improvement for 8-bit digital color images. Low visibility is generally presented in images as dark shadows, over-bright regions and blurred details. All these are all related to the luminance and contrast properties of images; therefore, it is a logical way to develop an image enhancement algorithm based on the processing of the luminance and contrast of images.

The proposed algorithm consists of two separated processes: adaptive luminance enhancement and adaptive local contrast enhancement, which are applied in a successive way to enhance images. Luminance enhancement is equivalent to dynamic range compression while the local contrast enhancement is a multi-scale neighborhood dependent process that is intended to preserve visual details and approximate the tonality of the original image. In other advanced algorithms, like the original Retinex [2] and multi-scale Retinex with color restoration (MSRCR) [3], both processes are implemented together. The separation of the two processes provides AINDANE more flexibility and capability to tune and control the whole image enhancement process.

For enhancement of color images by AINDANE, color images are first converted to intensity images prior to luminance and contrast enhancements for faster processing and consistent color rendition. The averaged luminance information of neighboring pixels, which is needed for local contrast enhancement, can be obtained using 2D discrete convolution. In order to obtain better results, image enhancement techniques should be made image dependent (adaptive). It requires that the image enhancement process can be tuned (controlled) by certain properties of images, such as the statistical information. After luminance and contrast enhancements have been performed, a linear color restoration process is applied to convert the intensity images back to color images using the chromatic information of the original image. Compared to other image enhancement techniques, AINDANE produces better image quality with well-balanced luminance and contrast as well as accurate and natural color rendition.

3.2.1 AINDANE Algorithms

First, color images in the RGB color space are converted to intensity (grayscale) images using the NTSC standard method defined in Eq. 3.1(a) that is for obtaining the luminance (intensity) information of color images on additive color device. Image intensity $I(x, y)$ is then normalized as

$$L_n(x, y) = \frac{L(x, y)}{255}. \tag{3.16}$$

Intensity images are treated by an enhancement process to elevate the intensity values of low-intensity pixels using a specifically designed non-linear transfer function defined by

$$L_n' = \frac{\left(L_n^{(0.75z+0.25)} + (1-L_n)\cdot 0.4\cdot(1-z) + L_n^{(2-z)}\right)}{2} \tag{3.17}$$

It can be observed from Eq. (3.17) that the non-linear transfer function is image dependent with a parameter z, which is related to the image histogram and is defined as

$$z = \begin{cases} 0 & \text{for} \quad L_{CDF} \le 20\% L_{\max} \\ \dfrac{L_{CDF} - 50}{100} & \text{for} \quad 20\% L_{\max} < L_{CDF} \le 60\% L_{\max} \\ 1 & \text{for} \quad L_{CDF} > 60\% L_{\max} \end{cases}$$

where L_{CDF} is the intensity level corresponding to a cumulative distribution function (CDF) of 0.1. L_{max} is the maximum intensity of the original image. That is, when more than 90% of all pixels have intensity higher than 60% L_{max}, z is 1. If 10% or more of all pixels have intensity lower than $20\%L_{max}$, z is 0. For all other cases, when the grayscale of 10% or more of all pixels are higher than $20\%L_{max}$ and lower than 60% L_{max}, $z = (L_{CDF} - 50)/100$. Obviously, L_{CDF} is used as an indication to determine how dark the 10% of pixels in an image are. If they are really dark (e.g. $L_{CDF} < 50$), luminance needs to be enhanced more. If they are not that dark (e.g. $L_{CDF} \approx 100$), less luminance enhancement will be needed. If most of the pixels have sufficient brightness (e.g. $L_{CDF} > 150$), no luminance enhancement will be needed. Fig. 3.11 illustrates the CDF of an intensity image with respect to the gray level L_{CDF}. The range of z and the relationship between z and L_{CDF} are determined empirically based on image enhancement experiments and authors' judgment. In addition, z can be a user adjustable parameter for manually tuning the luminance enhancement process.

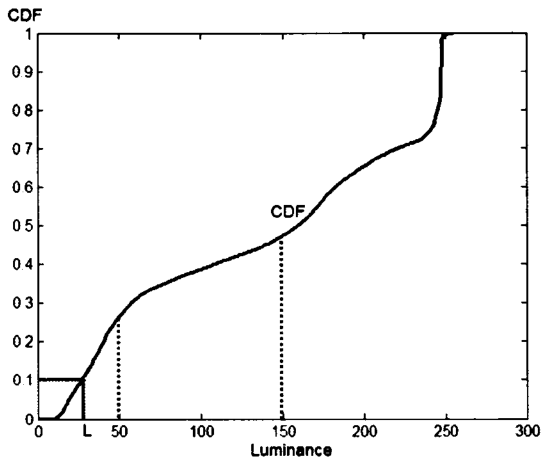


Figure 3.11 CDF of an intensity image.

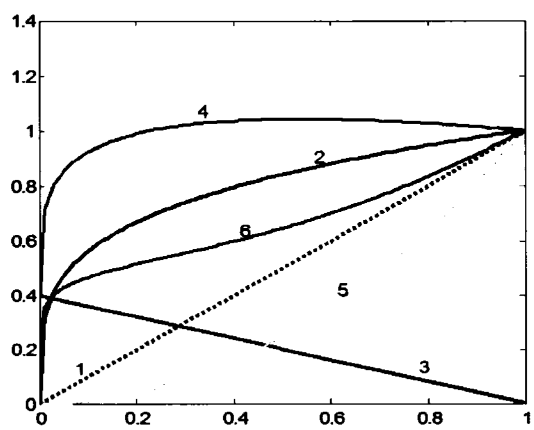


Figure 3.12 Nonlinear transfer function for luminance enhancement.

The transfer function is actually a combination of three simple mathematical functions and it is graphically shown as the curve 6 ($z = 0$) in Fig. 3.12 with a dotted straight line (line 1) for comparison, which in fact represents the identity transformation. The first two terms in Equation (3.17) are plotted as curve 2 and line 3, respectively, and the summation of them yields curve 4. The last term in Equation (3) is shown in the graph as curve 5. The addition of curve 4 and curve 5 after normalization (division by 2 in Equation (3)) produces the transfer function shown as curve 6. It can be seen that this transformation largely increases the luminance of darker pixels (regions) while brighter pixels (regions) are less enhanced. Thus this process also serves as dynamic range compression. Therefore, the line shape of the transfer function is important no matter what mathematical functions are used. Simple functions are applied for faster computation. Similar curve shapes produced by other types of functions will have similar effects on luminance enhancement if they are used as the transfer function. It should be noted that this transfer function is designed for the purpose of luminance enhancement and it can also provide appropriate dynamic range compression from which good enhanced results will be obtained through the contrast enhancement process. The effect of z on the transfer function is illustrated in Fig. 3.13. As z approaches 1, the transfer function curve gets closer to the identity transformation. The graph indicates that brighter images (with larger z) have less luminance enhancement in order to prevent over-enhancement.

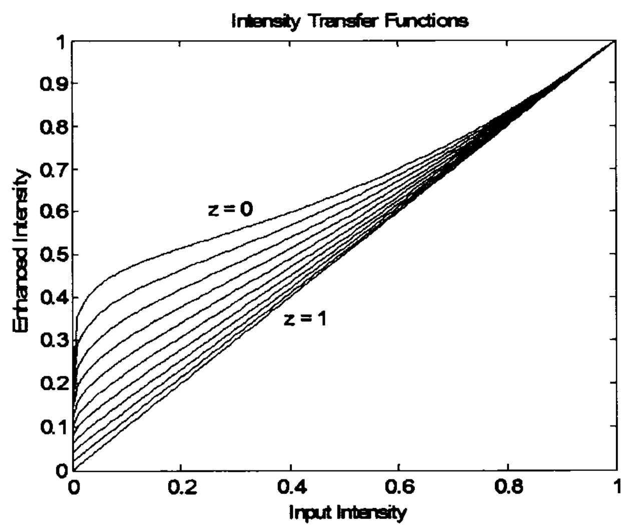


Figure 3.13 Nonlinear transfer functions with different z values.

After luminance enhancement, the contrast enhancement process is applied to restore the contrast of the luminance-enhanced images, which has been degraded during the previous process (the images look gray-out). The restored contrast may be even higher than that of original images for high visual quality. However, the normal global contrast enhancement technique is unable to fulfill that request. It simply increases the luminance for bright pixels and decreases the luminance for the dark pixels. As a result, the dynamic range can be significantly expanded. On the other hand, this method has limited performance for bringing out fine details where adjacent pixels have small luminance differences because the surrounding pixels are not considered when one pixel is being processed. Therefore, a surrounding pixel (neighborhood) dependent contrast enhancement technique must be implemented to achieve sufficient contrast for image enhancement without losing the dynamic range compression. While an image is processed with such a method, pixels with the same luminance can have different outputs depending on their neighboring pixels. When surrounded by darker or brighter pixels, the luminance of the pixel being processed (the center pixel) will be boosted or lowered respectively. In this way, picture contrast and fine details can be sufficiently enhanced while dynamic range expansion can be controlled without degrading image quality.

The luminance information of surrounding pixels is obtained using 2D discrete spatial convolution with a Gaussian kernel, which in essence is one type of neighborhood averaging. A Gaussian kernel is used due to its closeness to the way in which the human visual system works. This computation is described by Eqs. (3.3) to (3.5). In practice, convolution in the spatial domain can be computed by multiplication in the frequency domain. After the surrounding intensity information is obtained by a 2D convolution, the center pixel's intensity is compared with the convolution result. These two operations are both carried out on the original image. If the center pixel's intensity is higher than the average intensity of surrounding pixels, the corresponding pixel on the luminance-enhanced image (L_n') will be increased, otherwise it will be decreased. As a result, the contrast of the luminance-enhanced image can be adaptively improved without counteracting the effect of luminance enhancement.

The center-surround contrast enhancement is performed as per the following two equations:

$$S(x,y) = 255 \cdot L_n'(x,y)^{E(x,y)} \tag{3.18}$$

where the exponent is defined by

$$E(x,y) = r(x,y)^P = \left(\frac{L_{conv}(x,y)}{L(x,y)} \right)^P \tag{3.19}$$

$S(x, y)$ is the pixel intensity after contrast enhancement and $r(x, y)$ is the intensity ratio between $L_{conv}(x, y)$ and $L(x, y)$. P is an image dependent parameter, which is used to tune the contrast enhancement process. If the contrast of original image is poor, P will be larger and further increase the contrast enhancement. P is determined by the global standard deviation σ of the input intensity image $I(x, y)$ as

$$P = \begin{cases} 3 & for \quad \sigma \leq 3 \\ \dfrac{27-2\sigma}{7} & for \quad 3 < \sigma < 10 \\ 1 & for \quad \sigma \geq 10 \end{cases}$$

According to this definition, P is 1 if σ is greater than or equal to 10, and P equals to 3 when σ is less than or equal to 3. For all other cases, there is a linear relationship between the power P and σ. This relationship is determined based on image enhancement experiments. Here, the global standard deviation σ of $L(x, y)$ is considered as an indication of the contrast level of the original intensity image. It should be noted that P can be changed by users to manually adjust the contrast enhancement process.

The contrast enhancement process defined in Eqs. (3.18) and (3.19) is actually an intensity transformation process and can be understood using Fig. 3.14. Since L_n' is normalized to 1, $L_n'(x, y)^{E(x, y)}$ will be larger than $L_n'(x, y)$ if $E(x, y)$ is less than 1 (i.e. the center pixel is brighter than surrounding pixels leading to $r(x, y) < 1$). Otherwise, if $E(x, y)$ is larger than 1 (i.e. the center pixel is darker than the surrounding pixels with $r(x, y) > 1$), $L_n'(x, y)^{E(x, y)}$ will be smaller than $L_n'(x, y)$. In this way, the contrast of the luminance-enhanced image can be improved. Here, the ratio $r(x, y)$ is obtained from the original intensity image $L(x, y)$ and its low pass filtered result $L'(x, y)$, since the contrast information in the luminance enhanced image has been changed and degraded during the nonlinear luminance enhancement process.

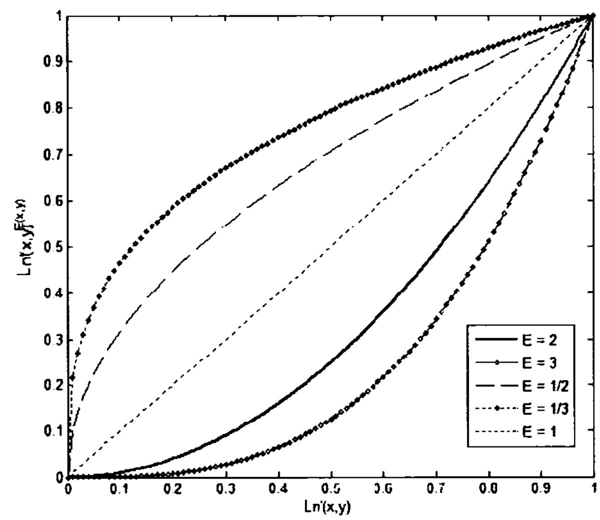


Figure 3.14 Intensity transformation for contrast enhancement.

For better results of image enhancement, contrast enhancement is performed with multiple convolution results using different scales. The final output is a linear combination of the contrast enhancement results based on multiple scales. A convolution with a small scale, such as a few neighboring pixels, can provide luminance information about the nearest neighborhood pixels, while the convolution with a large scale comparable to the image dimensions can provide the information about the large-scale luminance variation over the whole image. Generally, contrast enhancement with smaller scale convolutions tend to enhance local contrast or fine details, and the contrast enhancement with larger scale convolutions can produce a global tonality closing to the original image for smooth and natural looking results. A medium scale can provide a mixture of both details and image rendition. Convolutions with multiple scales can yield much more complete information on the image's luminance distribution, and hence lead to a much more balanced image enhancement. However, if faster processing or certain special effect is wanted, a single-scale convolution may be used for contrast enhancement. For example, only a medium scale convolution can be applied to achieve fast processing while the result is still acceptable. The contrast enhancement with multi-scale convolutions can be described by the following equations:

$$G_i(x,y) = K \cdot e^{\left(\frac{-(x^2+y^2)}{c_i^2}\right)} \tag{3.20}$$

$$L_{conv,i}(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} L(m,n)\, G_i(m+x, n+y) \tag{3.21}$$

$$E_i(x,y) = r_i(x,y)^P = \left(\frac{L_{conv,i}(x,y)}{L(x,y)} \right)^P \tag{3.22}$$

$$S_i(x,y) = 255 \cdot L_n{}'(x,y)^{E_i(x,y)} \tag{3.23}$$

$$S(x,y) = \sum_i w_i S_i(x,y) \tag{3.24}$$

where c_i (i = 1, 2, 3, ...) represents different scales and w_i is the weight factor for each contrast enhancement output $S_i(x, y)$. By default, $w_i = 1/n$, i = 1, 2, 3, ... n (n is the number of scales), based on our image enhancement experiments. n = 3 is typical and yields enhanced images having a well-balanced and natural visual effect. Both fine details and overall tonality can be accounted for in the output images. In this work, the three scales, 5, 20 and 120, are commonly used, or a simple guide can be followed: the smallest scale is 1-5% of the image size, the medium and the largest scales are 10-15% and 25-45% of the image size, respectively. c_i, w_i and n are usually fixed in our image enhancement experiments while they can still be changed to obtain optimal result or certain special effect.

So far, both luminance and contrast enhancements have been performed in the luminance space. The enhanced color image can be obtained through a linear color restoration process based on the chromatic information contained in the input image. Mathematically, the color restoration process for images in RGB color space can be expressed as

$$S_j(x,y) = S(x,y)\frac{L_j(x,y)}{L(x,y)} \cdot \lambda_j \tag{3.25}$$

where $j = r, g, b$ represents the R, G, B spectral band respectively, and S_r, S_g and S_b are the RGB values of the enhanced color image. A parameter λ is introduced here in order to manually adjust the color hue of the enhanced color images. λ is a constant smaller than but very close to 1, which takes different values in different spectral bands. When all λs are equal to 1, Eq. (3.25) can preserve the chromatic information of the input color image for minimal color shifts.

3.2.2 Results and Discussion

The proposed algorithm has been applied to enhance a large number of digital images for performance evaluation and comparison with other algorithms. Some typical results as well as detailed discussion about various characteristics of the algorithm are presented in this section.

Images enhanced with various parameter values are illustrated in Fig. 3.15. The effects of those parameters are clearly shown and are in agreement with the description provided in the previous section. If the parameters are manually adjusted, image quality can be changed to obtain optimized result. With the adaptiveness implemented in our algorithm, the parameter adjustment can be conducted automatically according to the quality of the original image. Automatic tuning can produce results better than or at least equivalent to those obtained with default parameter values.

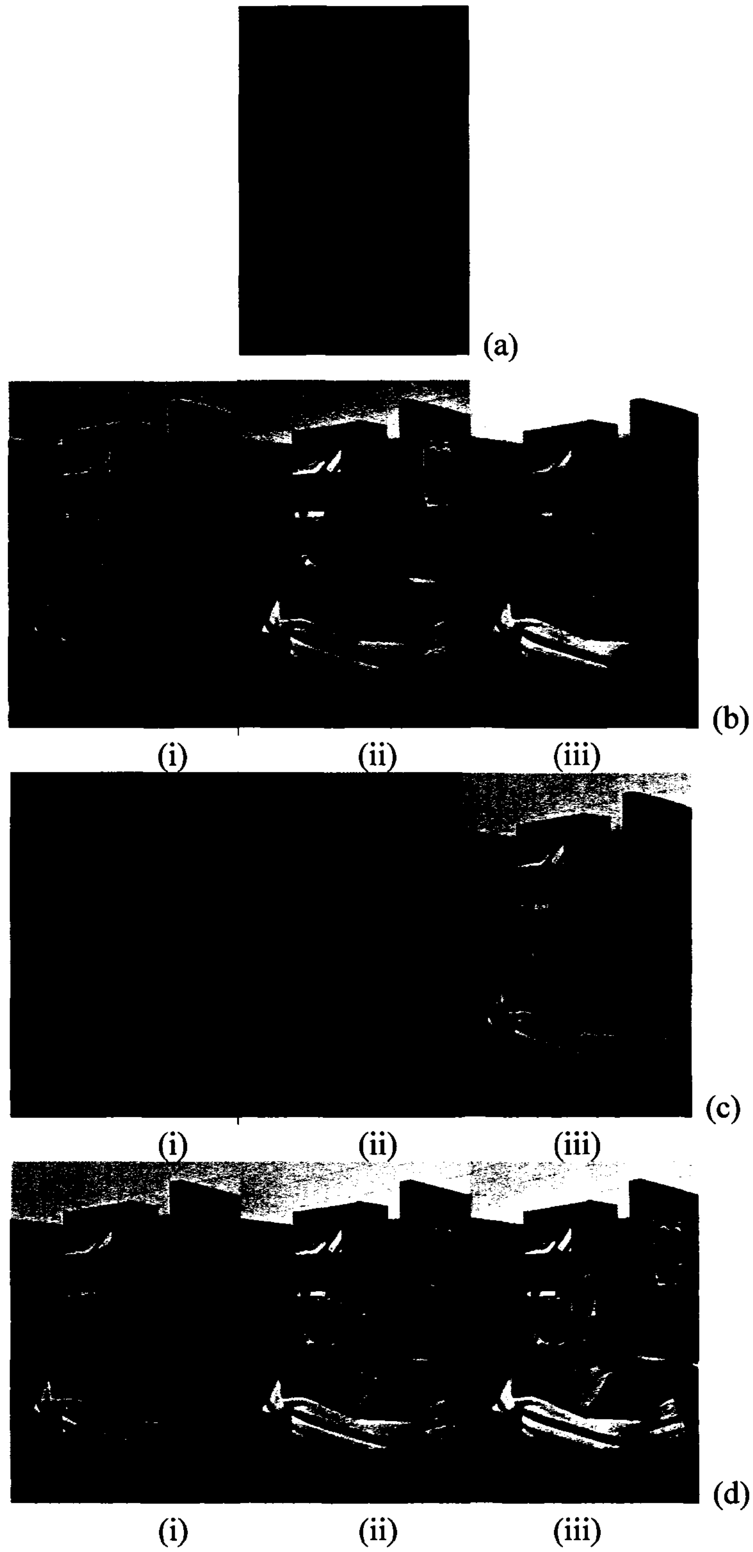


Figure 3.15 Image enhancement results with various parameter values: (a) original image; (b) images enhanced using single scale convolution and with z=0 and P=1: (i) c=5, (ii) c=20, and (iii) c=240; (c) images enhanced by multiscale c and with different z values and P=1: (i) z=1, (ii) z=0.5, and (iii) z=0; and (d) images enhanced by multiscale c and with different P values and z=0: (i) P=1, (ii) P=2, and (iii) P=3.

(a)

(b)

(c)

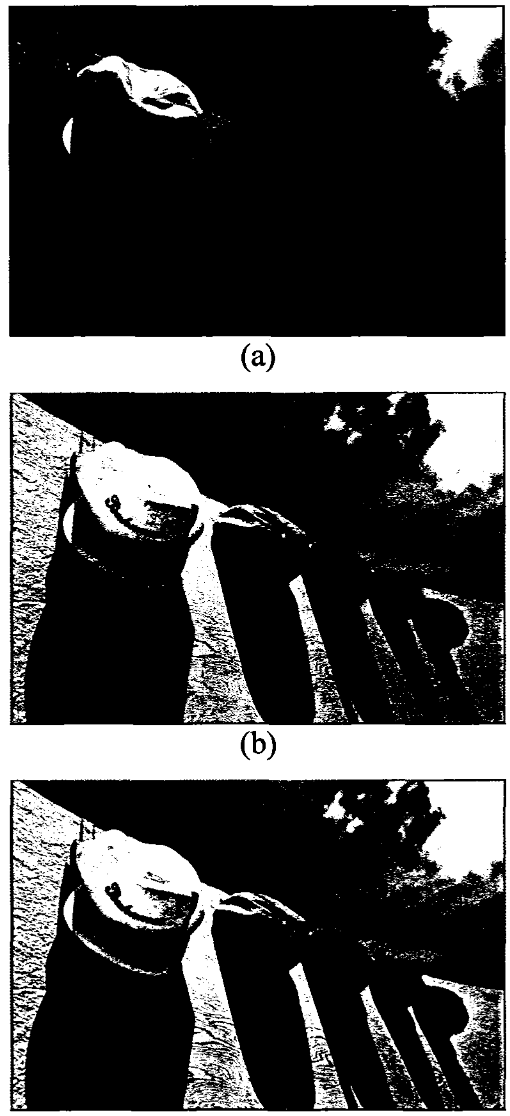
Figure 3.16 Luminance enhancement with adaptive factor z of different values: (a) original image, (b) result obtained from INDANE (z=0), and (c) result obtained from AINDANE ($z \neq 0$).

The AINDANE algorithm has been applied to process a large number of digital images taken by digital cameras under varying lighting conditions. The enhanced images have good quality, with fine details, well-balanced contrast and luminance across the whole image, and natural color rendition of appropriate color saturation. AINDANE has various adjustable parameters, which have been finely tuned by conducting several experiments to achieve consistently high quality results for different types of images. Some of the parameters are image dependent and are introduced to make the algorithm more adaptive. For example, the parameter z in Eq. (3.17) (nonlinear transfer function) is determined by the image's histogram and is used to adjust the luminance enhancement to avoid over or insufficient enhancement of luminance. The example in Fig. 3.16 shows the effect of z on enhanced images. The original image was taken with sufficient illumination. The enhanced image processed by INDANE [32] (without adaptiveness, $z = 0$ and $P = 1$) looks worse than the original in terms of contrast and color saturation. However, AINDANE ($z \neq 0$) produced a better result in which the person's face looks more real with outlines and edges being clearly brought out. The general contrast is improved by AINDANE.

P is another image dependent parameter used for contrast enhancement in AINDANE. It is used to ensure that the contrast of low-contrast images will be increased to obtain good visual effect. The sample image in Fig. 3.17 is the same one as we use in the I-R algorithm for illustrating the effect of P on contrast enhancement. We can see although P has different form in the two algorithms due to different scheme, it always can provide similar good and robust contrast enhancement result for the same sample image. The original image was captured with insufficient illumination under the caps and has poor contrast on the wood board. AINDANE with adaptive control parameter ($P \neq 1$) produces an enhanced image with much higher contrast than that obtained by INDANE that has no adaptive-ness ($P = 1$). Grains in the wood region became obvious and maintain the overall luminance level.

It can be found that AINDANE shares certain similarity with MSRCR. However, AINDANE has several advantages over MSRCR in terms of color rendition and flexibility in algorithm tuning. In MSRCR, each spectral band is individually processed, and color restoration for improving color constancy is realized using a non-linear process. These processes might produce some color artifacts, which cannot be predicted from the original image, making enhanced images look unnatural or incorrect in colors. As in MSRCR, dynamic range compression and contrast enhancement are implemented jointly, which makes the algorithm more difficult for tuning than AINDANE where the two processes are separated and can be tuned independently.

In both enhancement techniques, the convolution result is compared with the center pixel in the form of the intensity ratio. However, AINDANE is completely different from MSRCR on the way it uses this ratio. AINDANE performs a power-law transformation to the luminance-enhanced image using the ratio as the



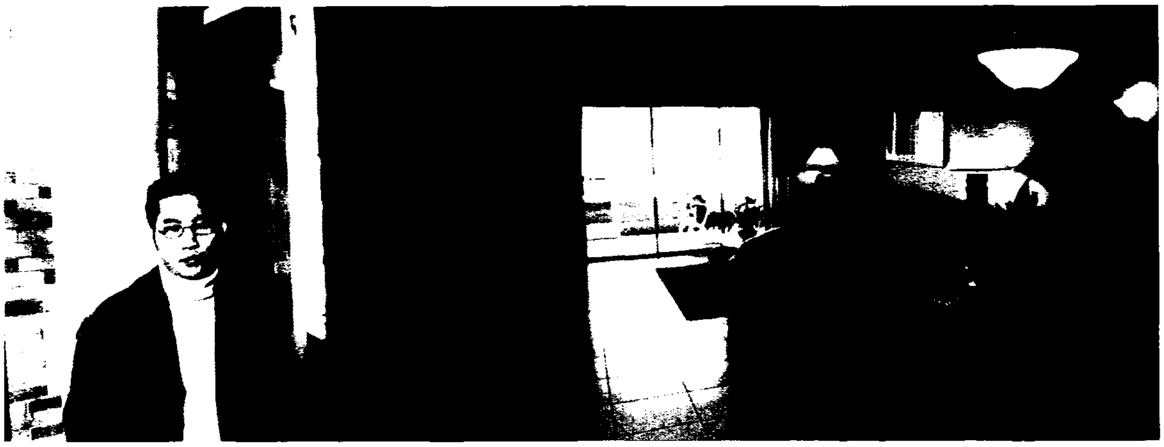
(a)

(b)

(c)

Figure 3.17 Contrast enhancement with adaptive factor P of different values: (a) original, (b) result obtained by INDANE (P=1), and (c) result obtained by AINDANE ($P\neq 1$).

exponent. As a result, the image's contrast can be enhanced. MSRCR however, perform a logarithmic transformation to the ratio followed by a gain-offset process, and both dynamic range compression and contrast enhancement are accomplished at the same time.

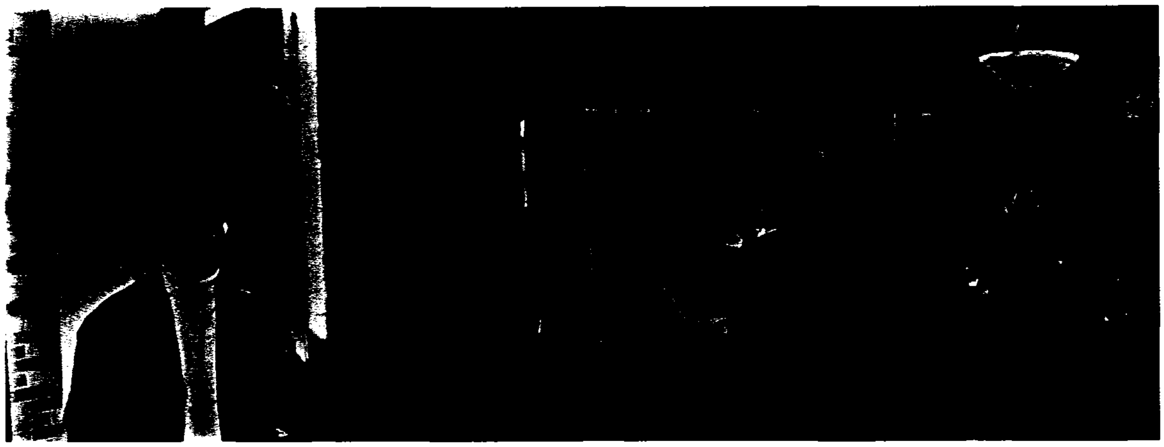


(a)

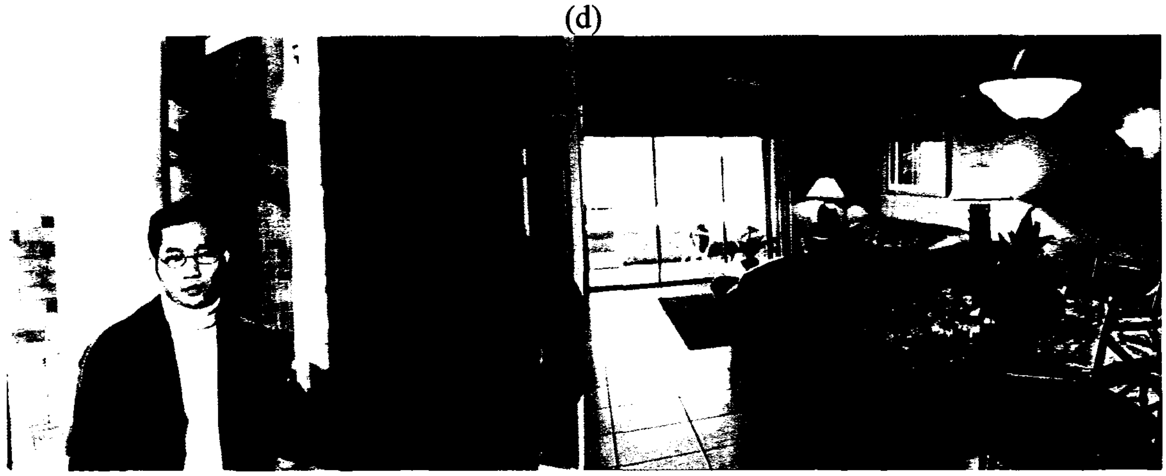
(b)

(c)

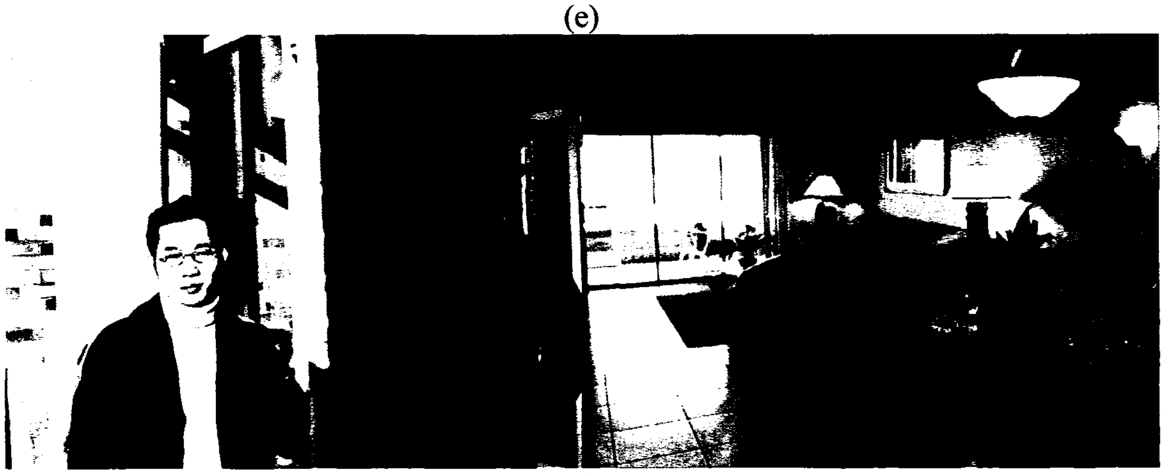
Figure 3.18 Image enhancement comparison: (a) original images, (b) AHE, (c) retinex, (d) MSRCR, (e) INDANE, and (f) AINDANE using two sample images captured with a non-uniform lighting environment.



(d)



(e)



(f)

Figure 3.18 (Continued).

In Fig. 3.18, two sample images are provided for comparison among the performance of Retinex, MSRCR, INDANE, AINDANE and AHE. Retinex was realized using the Matlab® code provided in Reference [44] by Frankle and McCann. A commercial software PhotoFlair® (www.truview.com) was used to implement MSRCR. It can be observed that the images processed by AINDANE have a higher visual quality than those processed by MSRCR, AHE and INDANE. AINDANE yields higher color accuracy and a better balance between the luminance and contrast across the whole image due to its adaptive-ness and flexibility involved in the processing. In the first sample image, AHE brings out the person standing in the shadow. However, artifacts appear on the white pillar and the face of the person in sunshine gets even dimmer. The whole image looks quite unnatural with such a brightness distribution. In the second sample image, AHE produces a lot of artifacts and noise, and luminance enhancement is not sufficient. The images enhanced by Retinex demonstrate high contrast and good illuminance. However, its color constancy works poorly in some places (The face of the man in the shadow is much whiter than it should be, and the color of different parts of the table top is inconsistent.) and the lightness rendition looks a little unnatural (The table top is too bright.). In addition, both MSRCR and Retinex show a much more visible halo effect than our proposed algorithm. It is also found that MSRCR seems to have some difficulty enhancing images in non-uniform lighting conditions. The luminance enhancement in the darker regions is insufficient. It is even worse in the sense that the brightness of some of the bright areas is incorrectly reduced to yield a tonality difference from the original. More comparisons between MSRCR, Retinex, and AINDANE are provided in Fig. 3.19. The previously discussed issues, incorrect lightness and color rendition as well as insufficient luminance enhancement, are still clearly visible. Obviously, MSRCR provides strong contrast enhancement but the luminance enhancement is poor. The luminance of high-brightness regions are even largely degraded after enhancement. In addition, the color rendition looks unnatural with high color saturation. Retinex performs much better than MSRCR. However, its color correction capability may also create incorrect colors. For example, the cloud in the left image is bleached although the cloud color is correct in the original image. Moreover, both MSRCR and Retinex provide insufficient luminance enhancement to the woman's face because of the highlighted background and the unbalancing between the luminance enhancement and contrast enhancement. On the other hand, the proposed algorithm generally performs well on those test images showing a more balanced result between luminance enhancement and contrast enhancement and no incorrect colors are created.

Figure 3.19 Image enhancement comparison: (a) originals, (b) MSRCR, (c) retinex, and (d) AINDANE on two sample images under non-uniform lighting environment.

3.2.3 Analysis of the I-R based Algorithm, AINDANE and MSRCR

From one point of view, the I-R model based algorithm, AINDANE and MSRCR may be considered similar techniques because they all make use of neighborhood dependent processing. However, the proposed algorithms are essentially different from MSRCR in terms of the manner in which images are processed and the performances.

First, let us discuss MSRCR. It can be observed that the single-scale Retinex (SSR) operation implemented in MSRCR (there are three SSR for each spectral band) is similar to the extraction of the scene reflectance previously discussed in the I-R model based algorithm except that logarithmic operation is used in SSR. The SSR process has two purposes: removal of the effect of illumination (i.e., compress the dynamic range) by eliminating the low frequency components and enhancement of the image contrast by the logarithmic operation. Due to these properties of SSR, MSR shows strong capability of contrast enhancement which may make images look unnatural as a result of over enhancement of image contrast. In addition, the loss of low frequency components may also make the lightness rendition different from that in the original image. Since the multiband process tends to decrease the color saturation, color restoration is then applied in MSRCR. However, the color restoration is found to produce incorrect colors in some output images although it is able to provide color correction (or color constancy) to the output images.

Compared to MSRCR, the proposed algorithms process images in a different way. For I-R model based algorithm, the high frequency components are obtained as the reflectance which is not degraded by the dynamic range compression. Thus the enhancement for high frequency components is not required. Only mid-tone frequency components are enhanced, and low frequency components do not need to be treated. Combined with single-band processing, the processing speed is greatly improved to achieve real time video processing maintaining the quality of the enhanced images. As for AINDANE, the entire image is processed with dynamic range compression which makes all frequency components degraded in terms of contrast. Therefore, multi-scale contrast enhancement is used to provide an all-round improvement of the image contrast. Multi-scale processing is also the reason why AINDANE performs better than I-R model based algorithm. In both algorithms, low frequency components are better kept compared to MSRCR, which makes the lightness rendition of the output images closer to that of the original images. This feature can let the output images look more natural. Due to the single band processing, AINDANE is also much faster than MSRCR. Finally, the linear color restoration used in our algorithms is able to prevent the incorrect color rendition appeared in MSRCR technique.

CHAPTER IV

IMAGE DENOISING AND CONTRAST ENHANCEMENT BASED ON WAVELET TRANSFORM

In this chapter, we present the investigation of two wavelet transform based image processing techniques: image denoising and contrast enhancement. The two techniques are based on the modification of wavelet coefficients to achieve their goals. For image denoising, it eliminates added noise information while, for image contrast enhancement, it enhances faint image features. The algorithm development of both techniques will be introduced and discussed separately in two sections. Both techniques are implemented using DT-CWT (dual tree complex wavelet transform) to obtain optimized results. The performance of both techniques is evaluated using objective assessment methods as well as the comparison with other techniques.

4.1 Wavelet Transform Based Image Denoising

The image denoising technique developed in this dissertation is based on the bivariate wavelet coefficient statistic model presented in [111] and the wavelet coefficient shrinkage method presented in [112].

4.1.1 Algorithm

In this section, we consider denoising an image corrupted by additive white Gaussian noise. The noisy image can be expressed as:

$$y = x + n \tag{4.1}$$

where n is independent Gaussian noise; x and y are clean and noisy images, respectively. The goal of image denoising is to estimate x as accurately as possible with a given y. To deal with this estimation problem in the original spatial domain is difficult. However, it can be simplified in a transformed domain. Considering an invertible linear transform denoted by T and applying it to Eq. (4.1), we have

$$Ty = Tx + Tn \tag{4.2}$$

If T is the wavelet transform, in the wavelet transform domain the problem can be formulated in the same manner as in Eq. (4.1):

$$u = v + n \quad (u = Ty \quad v = Tx \quad n = Tn) \tag{4.3}$$

where u, v and n are wavelet transformed versions (wavelet coefficients) of y, x and n in Eq. (4.1). Image denoising is a classical problem in estimation theory. In this dissertation, the maximum *a posteriori* (MAP) estimator will be used to estimate v from the noisy observation u.

The statistics of wavelet transform coefficients has been found to exhibit certain interlevel and intralevel dependency between coefficients. In our image denoising algorithm, the interlevel dependency is considered. To take into account the interlevel dependency between coefficients, we need to consider both wavelet coefficients and their parents in this problem. Thus Eq. (4.3) becomes:

$$u_1 = v_1 + n_1, \quad u_2 = v_2 + n_2 \tag{4.4}$$

where u_2 and u_1 are noisy observations of v_2 and v_1, and n_2 and n_1 are noise samples. The coefficients with subscript 2 are the corresponding parents of coefficients with subscript 1. The expressions in Eq. (4.4) can be combined using vector formulation as:

$$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{n} \tag{4.5}$$

where $\boldsymbol{u} = (u_1, u_2)$, $\boldsymbol{v} = (v_1, v_2)$, and $\boldsymbol{n} = (n_1, n_2)$.

The standard MAP estimator for Eq. (4.5) is

$$\hat{\boldsymbol{v}}(\boldsymbol{u}) = \underset{\boldsymbol{v}}{\mathbf{argmax}}[P_{\boldsymbol{v}|\boldsymbol{u}}(\boldsymbol{v} \mid \boldsymbol{u})] \tag{4.6}$$

which means that the estimation of $\boldsymbol{v}$ should maximize the conditional probability $P_{\boldsymbol{v}|\boldsymbol{u}}(\boldsymbol{v}|\boldsymbol{u})$. By using Bayes rule, we obtain:

$$\begin{aligned}\hat{\boldsymbol{v}}(\boldsymbol{u}) &= \underset{\boldsymbol{v}}{\mathbf{argmax}}[P_{\boldsymbol{u}|\boldsymbol{v}}(\boldsymbol{u} \mid \boldsymbol{v}) \cdot P_{\boldsymbol{v}}(\boldsymbol{v})] \\ &= \underset{\boldsymbol{v}}{\mathbf{argmax}}[P_{\boldsymbol{n}}(\boldsymbol{u} - \boldsymbol{v}) \cdot P_{\boldsymbol{v}}(\boldsymbol{v})]\end{aligned} \tag{4.7}$$

Eq. (4.7) can be further transformed into

$$\hat{\boldsymbol{v}}(\boldsymbol{u}) = \underset{\boldsymbol{v}}{\mathbf{argmax}}[\log(P_{\boldsymbol{n}}(\boldsymbol{u} - \boldsymbol{v})) + \log(P_{\boldsymbol{v}}(\boldsymbol{v}))]. \tag{4.8}$$

It can be seen that the estimation of $\boldsymbol{v}$ is formulated in terms of the probability density function (pdf) of the noise $P_{\boldsymbol{n}}$ and the pdf of the coefficients of true signal $P_{\boldsymbol{v}}$. Therefore, to use this equation to estimate the true signal, we must know both $P_{\boldsymbol{y}}$ and $P_{\boldsymbol{n}}$. The models of these density functions are essential for this estimation problem.

Here, we assume the noise is i.i.d (independent identical distribution) Gaussian, and the noise pdf $P_n(\boldsymbol{n})$ can be written as:

$$P_n(\boldsymbol{n}) = \frac{1}{2\pi\sigma_n^2} \cdot \exp\left(-\frac{n_1^2 + n_2^2}{2\sigma_n^2}\right) \tag{4.9}$$

where σ_n^2 is the noise variance (square of the noise standard deviation). It should be noted that Eq. (4.9) is only a simplified (or approximated) pdf for obtaining meaningful result. The equation is used to model the statistics of the wavelet coefficients with interlevel dependency, which in fact is not circularly symmetric. That means a more accurate noise pdf should have two different variances.

By combining Eq. (4.8) and (4.9), we have

$$\hat{\boldsymbol{v}}(\boldsymbol{u}) = \underset{\boldsymbol{v}}{\mathbf{argmax}}\left[-\frac{(u_1 - v_1)^2}{2\sigma_n^2} - \frac{(u_2 - v_2)^2}{2\sigma_n^2} + \log(P_v(\boldsymbol{v}))\right] \tag{4.10}$$

If $P_v(\boldsymbol{v})$ is assumed to be strictly convex and differentiable, Eq. (4.10) is equivalent to the two equations given below:

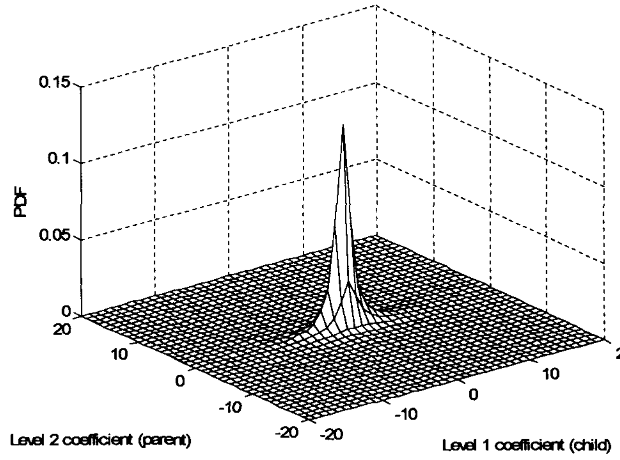
$$\frac{u_1 - \hat{v}_1}{\sigma_n^2} + \left[\frac{d}{dv_1}\log(P_v(\boldsymbol{v}))\right]_{\boldsymbol{v}=\hat{\boldsymbol{v}}} = 0 \tag{4.11}$$

$$\frac{u_2 - \hat{v}_2}{\sigma_n^2} + \left[\frac{d}{dv_2}\log(P_v(\boldsymbol{v}))\right]_{\boldsymbol{v}=\hat{\boldsymbol{v}}} = 0 \tag{4.12}$$

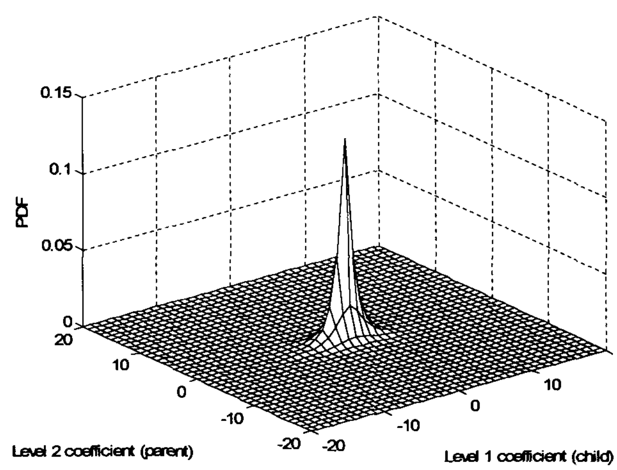
In order to accurately model the joint pdf of signal coefficients ($\boldsymbol{P_v}$), a computation described below is conducted. The joint histograms of level 1 and level 2 wavelet coefficients of 100 randomly selected natural images are calculated and the averaged joint histogram is presented in Fig. 4.1(a). From the experiment, it is found that the variances of the wavelet coefficients of natural images are quite different from level to level. Therefore, a joint pdf which has two adjustable marginal variances should be considered as [111]:

$$P_v(\boldsymbol{v}) = \frac{3}{2\pi\sigma_1\sigma_2} \cdot \exp\left(-\sqrt{3} \cdot \sqrt{\left(\frac{v_1}{\sigma_1}\right)^2 + \left(\frac{v_2}{\sigma_2}\right)^2}\right) \tag{4.13}$$

where σ_1 and σ_2 are variances of the child coefficient (v_1) and parent coefficient (v_2), respectively. Eq. (4.13) can be considered as a variant of bivariate Gaussian distribution or a variant of bivariate Laplace distribution. The two random variables v_1 and v_2 are uncorrelated but not independent. A plot of Eq. (4.13) (eg. $\sigma_1 = 2$, $\sigma_2 = 1$) is presented in



(a)



(b)

Figure 4.1 (a) Joint pdf of level 1 and 2 wavelet coefficients of natural images; (b) joint pdf defined in Eq. (4.13).

Fig. 4.1(b) for comparison with experimental data illustrated in Fig. 4.1(a). It can be observed that both distributions have a similar appearance. Data analysis also reveals that they are in good agreement. From this pdf, we have

$$\log(P_v(v)) = -\sqrt{3}\cdot\sqrt{\left(\frac{v_1}{\sigma_1}\right)^2 + \left(\frac{v_2}{\sigma_2}\right)^2} + \log\left(\frac{3}{2\pi\sigma_1\sigma_2}\right), \tag{4.14}$$

which can be further changed to

$$\frac{d}{dv_1}\log(P_v(v)) = -\frac{\sqrt{3}}{\sigma_1^2}\cdot\frac{v_1}{\sqrt{\left(\frac{v_1}{\sigma_1}\right)^2+\left(\frac{v_2}{\sigma_2}\right)^2}}, \tag{4.15}$$

and

$$\frac{d}{dv_2}\log(P_v(v)) = -\frac{\sqrt{3}}{\sigma_2^2}\cdot\frac{v_2}{\sqrt{\left(\frac{v_1}{\sigma_1}\right)^2+\left(\frac{v_2}{\sigma_2}\right)^2}}. \tag{4.16}$$

Substituting Eq. (4.15) and (4.16) into Eq. (4.11) and (4.12) produces

$$\hat{v}_1\cdot\left(1+\frac{\sqrt{3}\sigma_n^2}{\sigma_1^2 s}\right) = u_1 \tag{4.17}$$

and

$$\hat{v}_2\cdot\left(1+\frac{\sqrt{3}\sigma_n^2}{\sigma_2^2 s}\right) = u_2 \tag{4.18}$$

where

$$s = \sqrt{\left(\frac{\hat{v}_1}{\sigma_1}\right)^2+\left(\frac{\hat{v}_2}{\sigma_2}\right)^2}. \tag{4.19}$$

Eqs. (4.17) and (4.18) do not have an analytical solution, which means there is no simple expression for the bivariate shrinkage function (or MAP estimator). The solution can be found using iterative numerical methods [111] which, however, significantly increases the complexity of the denoising algorithm.

In our wavelet based image denoising algorithm, we propose a bivariate shrinkage function which can be considered as an approximate solution to Eqs. (4.17) and (4.18). which is written as

$$\hat{v}_1 = \frac{\left(\sqrt{u_1^2\sigma_1^2+u_2^2\sigma_2^2}-\sqrt{3}\sigma_n^2\right)_+}{\sqrt{u_1^2\sigma_1^2+u_2^2\sigma_2^2}}u_1 \tag{4.20}$$

where $\hat{v}_1$ is the estimated child coefficient. Accordingly, the deadzone (DZ) for coefficient shrinkage can be expressed as:

$$\mathrm{DZ} = \left\{(u_1,u_2) : \sqrt{u_1^2\sigma_1^2+u_2^2\sigma_2^2} \le \sqrt{3}\sigma_n^2\right\}. \tag{4.21}$$

A plot of Eq. (4.20) with $\sigma_1 = 7$, $\sigma_2 = 4$, $\sigma_n = 4$ is shown in Fig. 4.2. The flat ellipse region that has the coefficient estimate equal to zero is the deadzone defined by Eq. (4.21). This result is in good agreement with that obtained using the iterative method presented in [111], but our algorithm is much simpler in computing the shrinkage function.

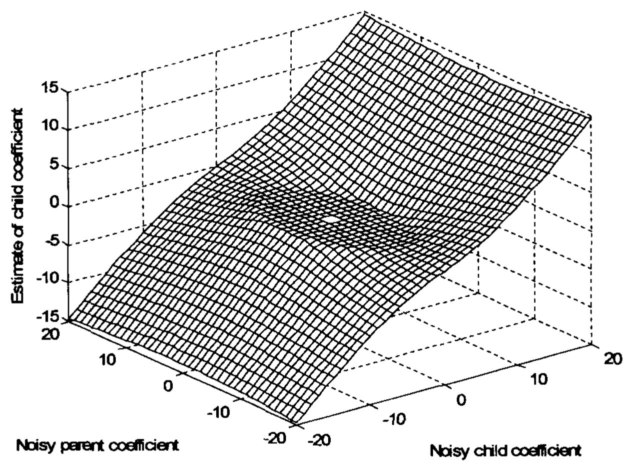


Figure 4.2 Bivariate shrinkage function proposed in Eq. (4.20).

The approximate solution in Eq. (4.20) is proposed using three facts. First, if we let $\sigma_1 = \sigma_2 = \sigma$, Eq. (4.20) becomes

$$\hat{v}_1 = \frac{\left(\sqrt{u_1^2 + u_2^2} - \dfrac{\sqrt{3}\sigma_n^2}{\sigma}\right)_+}{\sqrt{u_1^2 + u_2^2}} u_1. \tag{4.22}$$

Eq. (4.22) is actually the solution to Eqs. (4.17) and (4.18) under the condition $\sigma_1 = \sigma_2 = \sigma$. The dead zone defined in Eq. (4.22) is a circular region represented as $\sqrt{u_1^2 + u_2^2} - \dfrac{\sqrt{3}\sigma_n^2}{\sigma} \geq 0$. This result indicates that there exists a meaningful connection between Eq. (4.20) and Eq. (4.22). This connection makes us believe there is a certain validity of Eq. (4.20). The second fact is that the 3D surface produced by Eq. (4.20) (Fig. 4.2) is very close to that obtained using an iterative method [111]. This result suggests that the proposed approximate solution and the numerical solution describe similar mathematical relations among the variables. The last fact that supports the validity of Eq.

(4.20) is the image denoising experiments in which Eq. (4.20) yields very good results for various noisy images which contain different levels of noise. Therefore, it is reasonable to believe that those results are not just a matter of coincidence. Therefore, we believe that the validity of Eq. (4.20) can be verified by the three facts discussed above.

In order to compute $\hat{v}_1$ (see Eq. (4.20)), σ_1, σ_2 and σ_n must be known. σ_n is estimated from the wavelet coefficients in the subband of the lowest level with the highest resolution using a robust median estimator proposed in [113]

$$\hat{\sigma}_n^2 = \frac{\text{median}(|u|)}{0.6745}, \ u \in \text{subband } HH \tag{4.23}$$

where u is the noisy coefficient and HH is the subband of the highest resolution. σ_1, σ_2 can be estimated in an empirical way based on window operation. Recalling the observation model presented in Eq. (4.4), because v_1 and n_1 as well as v_2 and n_2 are independent of each other, we have

$$\sigma_{u_1}^2 = \sigma_1^2 + \sigma_n^2 \ \text{ and } \ \sigma_{u_2}^2 = \sigma_2^2 + \sigma_n^2 \tag{4.24}$$

where $\sigma_{u_1}^2$ and $\sigma_{u_2}^2$ are marginal variances of u_1 and u_2, respectively. From experiments, we observe that u_1 and u_2 can be modeled as Gaussian distributions with zero mean. Thus $\sigma_{u_1}^2$ and $\sigma_{u_2}^2$ can be computed empirically as in

$$\hat{\sigma}_{u_1}^2 = \frac{1}{N_W^2} \sum_{(x',y') \in W} u_1^2(x',y') \ \text{ and } \ \hat{\sigma}_{u_2}^2 = \frac{1}{N_W^2} \sum_{(x',y') \in W} u_2^2(x',y') \tag{4.25}$$

where W is a window in which the number of pixels is N_W^2. The window sizes used in the algorithm are either 7 × 7 or 5 × 5. The window based method produces location dependent variances. Theoretically, the window size can expand to the full subband size while N_w becomes the number of all the pixels in a subband. However, the experiments show better results obtained using window based operation. Using Eqs. (4.24) and (4.25), the standard deviation of v_1 and v_2 can be estimated as

$$\hat{\sigma}_1 = \sqrt{\left(\hat{\sigma}_{u_1}^2 - \hat{\sigma}_n^2\right)} \ \text{ and } \ \hat{\sigma}_2 = \sqrt{\left(\hat{\sigma}_{u_2}^2 - \hat{\sigma}_n^2\right)} \tag{4.26}$$

The denoising algorithm can be summarized as follows:

(1) Perform the wavelet transform on the noisy image up to level J (J = 5 is commonly used in our experiments) using DT-CWT.

(2) Calculate the noise variance using Eq. (4.23).

(3) For each subband in the decomposition levels up to level $J-1$

a) Calculate the marginal variances from the current subband and its parent subband (the subband one level is higher with the same orientation as the current subband) using Eq. (4.25).

b) Calculate σ_1 and σ_2 using Eq. (4.26).

c) Estimate the wavelet coefficients in the child subband using Eq. (4.20).

(4) Perform inverse DT-CWT transform on the estimated coefficients to obtain the denoised image. It should be noted that the wavelet coefficients in the highest level J have not been changed during the previous steps.

4.1.2 Results and discussion

Image denoising experiments have been conducted using the proposed algorithm. Three commonly used grayscale test images: Barbara, Boat and Lena shown in Fig. 4.3, are adopted in the experiments. First, the noisy test images with different levels of noise are obtained by adding white Gaussian noises with different standard deviations in the original test image. Then the denoising algorithm is applied to the noisy images to obtain denoised images. The performance of our algorithm is evaluated by visual inspection and comparison with published results produced by other developed algorithms.

We compared our proposed algorithm to various effective denoising algorithms published in literatures in terms of peak signal-to-noise ratio (PSNR) which is defined in decibels for 8-bit depth grayscale images.

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{256}{\varepsilon}\right) \tag{4.27}$$

where ε is the root-mean-square (rms) error given by

$$\varepsilon = \sqrt{\frac{1}{M}\sum\left(Y(x,y) - X(x,y)\right)^2} \tag{4.28}$$

where $X(x, y)$ is the image without the additive noise while $Y(x, y)$ is either the noisy image or denoised image. M is the number of pixels in the image.

Table 4.1 PSNR values of denoised images for different test images and noise levels (σ_n) (The values in the column of Original Noisy are values of noisy images before denoising).

	Original Noisy	BayesShrink in [114]	AdaptShr in [115]	HMT in [116]	LAWMAP in [117]	The system in [118]	SI-AdaptShr in [115]	CHMT in [119]	Proposed
Barbara									
$\sigma_n = 10$	28.15	30.86	-	31.36	31.99	33.35	-	-	**33.36**
$\sigma_n = 15$	24.55	24.65	29.96	29.23	29.60	31.10	31.14	-	**31.10**
$\sigma_n = 20$	22.18	22.14	28.36	27.80	27.94	29.44	29.52	-	**29.56**
$\sigma_n = 25$	20.15	20.17	27.23	25.99	26.75	28.23	28.33	-	**28.45**
$\sigma_n = 30$	18.65	-		25.11	25.80	-	-	-	**27.54**
Boat									
$\sigma_n = 10$	28.19	31.80	-	32.28	32.25	-	-	-	**33.06**
$\sigma_n = 15$	24.60	29.87	-	30.31	30.40	-	-	-	**31.20**
$\sigma_n = 20$	22.20	28.48	-	28.84	29.00	-	-	-	**30.01**
$\sigma_n = 25$	20.18	27.40	-	27.68	27.91	-	-	-	**29.00**
$\sigma_n = 30$	18.61	26.60	-	26.83	27.06	-	-	-	**27.96**
Lena									
$\sigma_n = 10$	28.16	33.32	-	33.84	34.10	34.96	-	34.9	**34.88**
$\sigma_n = 15$	24.57	31.41	32.39	31.76	32.23	33.05	33.41	-	**33.50**
$\sigma_n = 20$	22.19	30.17	31.07	30.39	30.89	31.72	32.12	-	**32.29**
$\sigma_n = 25$	20.16	29.22	30.70	29.89	29.89	30.64	31.11	29.9	**31.28**
$\sigma_n = 30$	18.64	28.48	-	28.35	29.05	-	-	-	**29.53**

The PSNR results of our algorithm and some other published algorithms are presented in Table 4.1. Performances of seven algorithms are compared with our algorithm in the table, which include BayesShrink [114], AdaptShrink [115], locally adaptive window-based denoising using MAP (LAWMAP) estimator [117], hidden Markov tree (HMT) model [116], undecimated wavelet transform presented in [115] and [118], and the DT-CWT [119]. It can be observed that our PSNR values are among the highest in the table, which verifies the effectiveness and performance of our algorithm although it is a relatively low-complexity image denoising algorithm.

(a)

(b)

(c)

Figure 4.3 Test images: (a) Barbara; (b) Boat; and (c) Lena.

Figure 4.4 Noisy images with σ_n =10 (left column) and denoised results (right column).

Figure 4.5 Noisy images with σ_n =30 (left column) and denoised results (right column).

Several examples of denoised images produced from noisy images with different noise levels are presented in Figs. 4.4 and 4.5. The noisy images are shown in the left column while the corresponding denoised results are provided in the right column. It can be seen that high quality image denoising is achieved with our proposed algorithm.

4.2 Wavelet Transform Based Image Contrast Enhancement

In this section, the image quality change caused by the modification of wavelet coefficients will be first investigated. Then an optimized transfer function for the modification of wavelet coefficients will be proposed to obtain high quality contrast enhancement for images.

The modification of the wavelet coefficients is one of the core issues in the wavelet transform based contrast enhancement. However, the characteristics of reconstructed images affected by the coefficient change has not been systematically studied. In this dissertation, image processing experiments have been conducted to investigate this critical problem.

We first introduce a specifically designed non-linear transfer function which is used to modify the wavelet transform coefficients. The function can be expressed as

$$w_o = \begin{cases} w_i & \textit{if } \ |w_i| \le S_1 \\ \left(\dfrac{|w'|}{S_2 - S_1}\right)^G \cdot |S_2 - S_1| + S_1 & \textit{if } \ S_1 < |w_i| < S_2 \ \textit{ and } \ w' > 0 \\ -\left(\dfrac{|w'|}{S_2 - S_1}\right)^G \cdot |S_2 - S_1| - S_1 & \textit{if } \ S_1 < |w_i| < S_2 \ \textit{ and } \ w' < 0 \\ w_i & \textit{if } \ |w_i| \ge S_2 \end{cases} \tag{4.29}$$

$$\text{with } \ w' = \begin{cases} w_i - S_1 & \textit{if } \ S_1 < |w_i| < S_2 \ \textit{ and } \ w_i > 0 \\ w_i + S_1 & \textit{if } \ S_1 < |w_i| < S_2 \ \textit{ and } \ w_i < 0 \end{cases},$$

where w_i and w_o are the input (original) and output (modified) wavelet coefficients, respectively. S_1 and S_2 are two nonnegative constants with $S_1 < S_2$. A plot of Eq. (4.29) is provided in Fig. 4.6 with $S_1 = 0$, $S_2 = 25$, and $G = 0.5$. In the same plot, the identity transformation is also illustrated for comparison. The plot shows that the value change is only made to the wavelet coefficients whose absolute values lie in the range $[S_1\ S_2]$ which corresponds to relatively small wavelet coefficients. The purpose of increasing the magnitude of those wavelet coefficients is to enhance the trivial image features where the intensity of bright pixels are increased while the intensity of dark pixels are decreased.

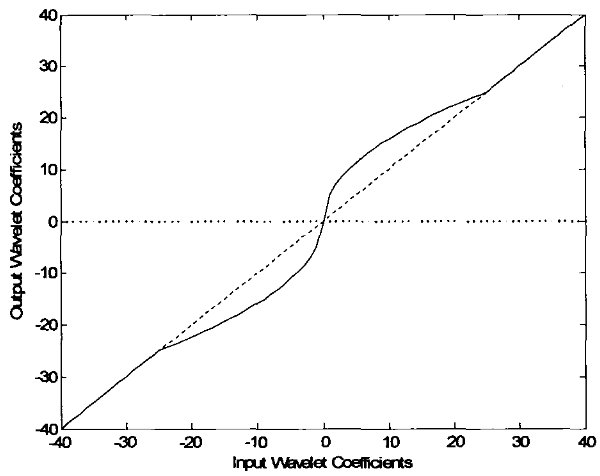
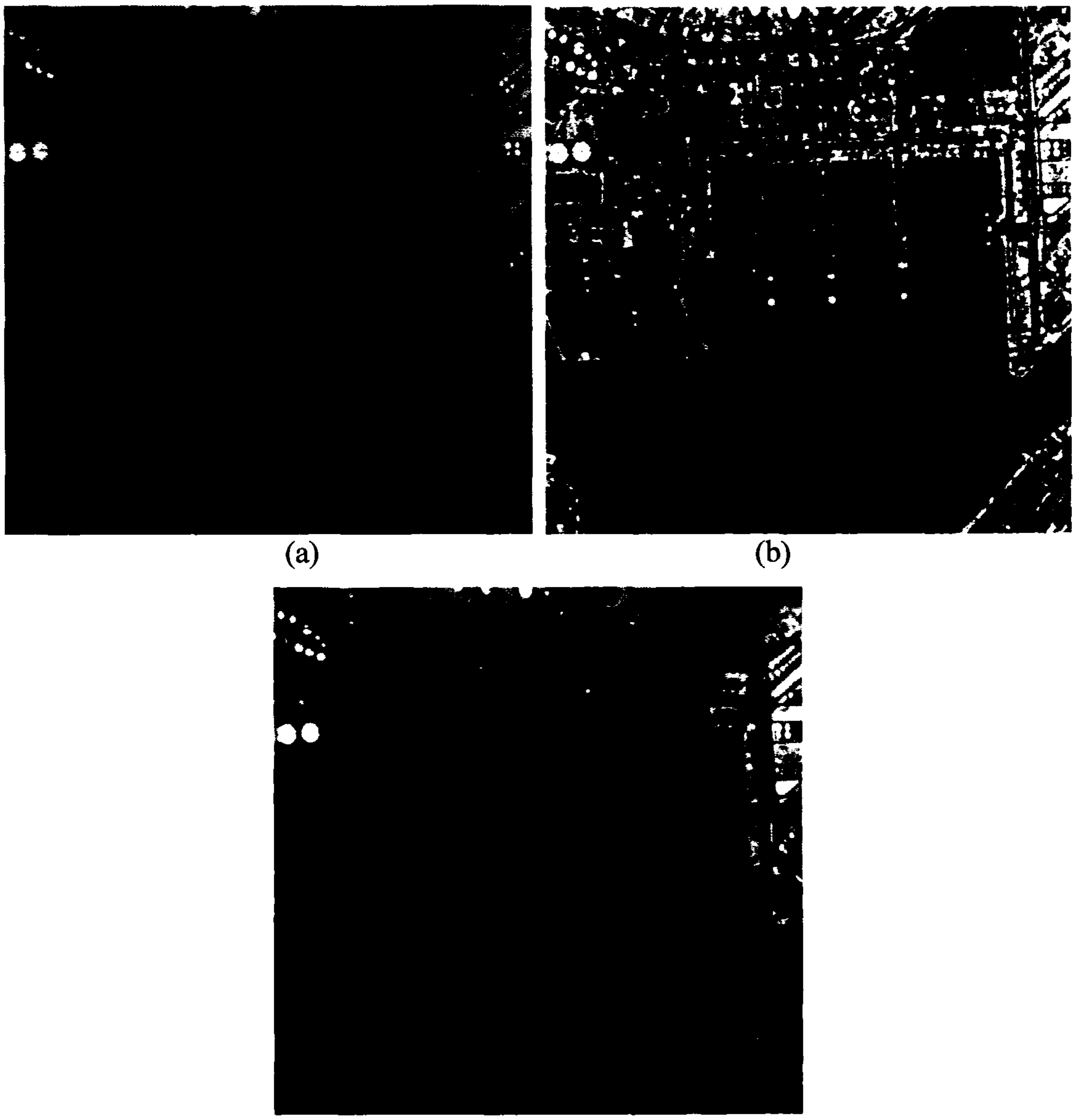


Figure 4.6 Transformation of wavelet coefficients.

The image quality change due to wavelet coefficient modification can be understood by using the sample images shown in Fig. 4.7, where the contrast enhanced image in (b) is obtained by reconstructing the modified wavelet coefficients from the original image. The coefficient modification is conducted in each subband using Eq. (4.29) with $G = 0.5$, $S_1 = 0$, and $S_2 = \max(|w_i|)$, in which w_i represents all the original coefficients in the subband. The image contrast has been largely improved by modifying (enhancing) the image features of all scales. But in this case, the noise is also enhanced. For comparison, an enhanced image produced by global contrast enhancement is also presented in (c) where many faint image features (e.g., the ripples and the features in the dark regions) have not been sufficiently enhanced. This is because wavelet transform is an multi-resolution (MR) decomposition and all image features with different spatial frequencies can be extracted and represented in the wavelet domain. On the other hand, the global contrast enhancement is unable to deal with all image details since its processing is not locally dependent.

Strong noise amplification can be observed in Fig. 4.7(b), which makes the image look slightly blurred and not clean. This is due to the large increase of those small wavelet coefficients close to the origin (see Fig. 4.6) because a significant part of those small coefficients is noise related. In order to suppress the unwanted noise, the transfer function needs to be tuned to avoid amplification of those small coefficients. In Fig. 4.8(a), a transfer function that is still produced by Eq. (4.29) but different from the one shown in Fig. 4.6 is presented with



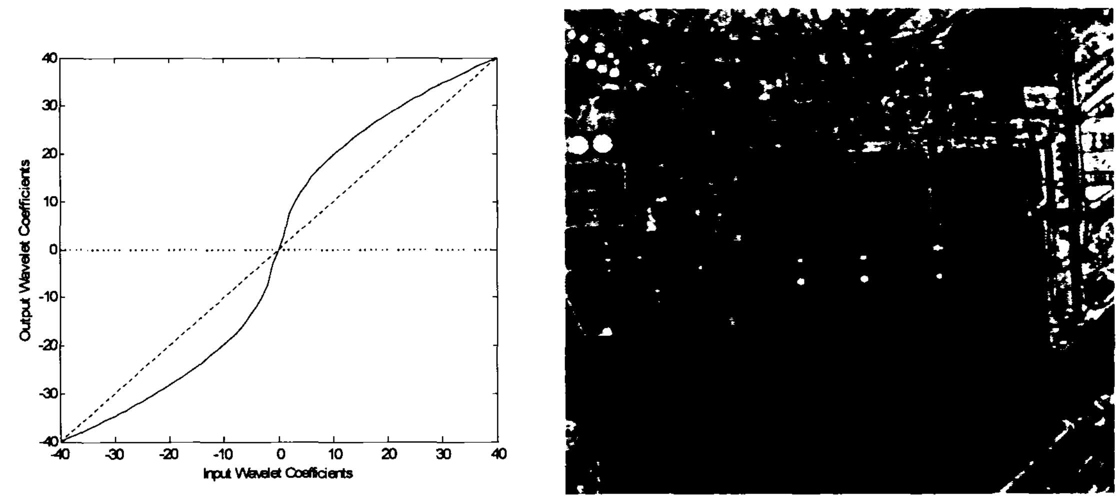
(a) (b)

(c)

Figure 4.7 Contrast enhancement result: (a) original low contrast image (by courtesy of [83]); (b) enhanced by wavelet coefficient modification; (c) enhanced by global contrast enhancement.

$$S_1 = c_1\sigma_n \tag{4.30}$$

where c_1 is a constant ($c_1 = 1$ for the plot in Fig. 4.8(a)) and σ_n is the noise standard deviation which is obtained using the robust median estimator from the lowest level (finest scale) coefficients (see Eq. (4.23)). The contrast enhanced image with $S_1 = \sigma_n$ ($\sigma_n = 0.9$ for the original image shown in Fig 4.7(a)) and $S_2 = \max(|w_i|)$ is shown in Fig. 4.8(b) which exhibits a much lower noise level compared to the image in Fig. 4.7(b). The default value of parameter c_1 is 1 while it is adjustable based on the image enhancement result.



(a) (b)

Figure 4.8 Contrast enhancement with noise suppression: (a) coefficient transfer function; (b) enhanced image.

Similar to S_1, the value of S_2 can be tunable based the quality of the enhanced image. The value of S_2 can be assigned in two ways. In the first way, S_2 is set based on the maximum absolute value (M_{wc}) of wavelet coefficients of each subband. For example, S_2 can be expressed as:

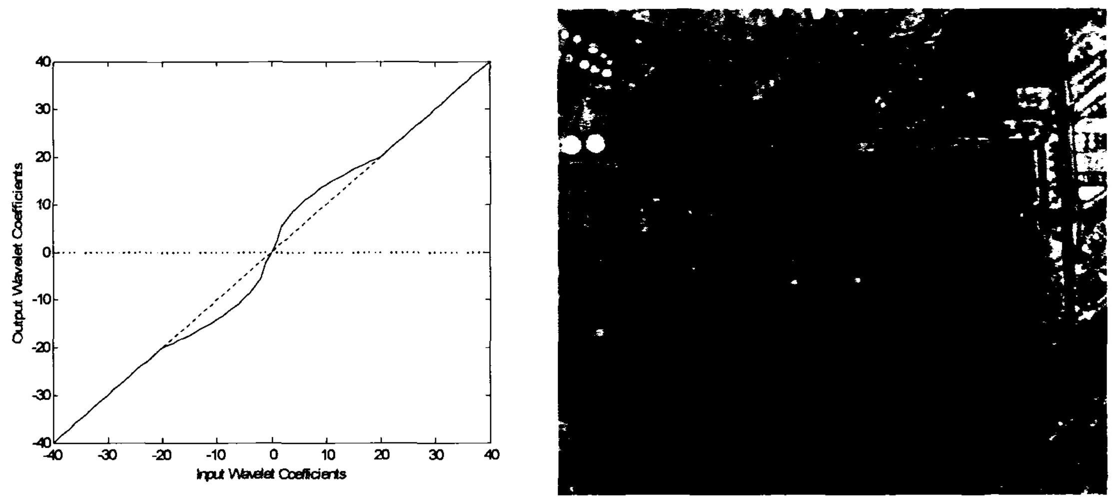
$$S_2 = c_{2m} \cdot M_{wc} \tag{31}$$

where c_{2m} is a constant less than or equal to 1. So if we choose $c_1 = 1$ and $c_{2m} = 0.5$, we amplify all coefficients with the absolute value between σ and half of the maximum absolute value of the subband. A plot of the transfer function with $c_1 = 1$ and $c_{2m} = 0.5$ is shown in Fig. 4.9(a) with the enhanced image shown in Fig. 4.9(b). Significant contrast loss can be observed as compared to Fig. 4.8(b) due to the lower amplification of the wavelet coefficients. However, the high-lighted features (corresponding to coefficients of large magnitudes) in Fig. 4.8(b) are also decreased to prevent over-enhancement.

In the second way for assigning the value of S_2, S_2 can be derived from the noise standard deviation. A simple implementation of that method can be written as

$$S_2 = c_{2s} \cdot \sigma_n \tag{4.32}$$

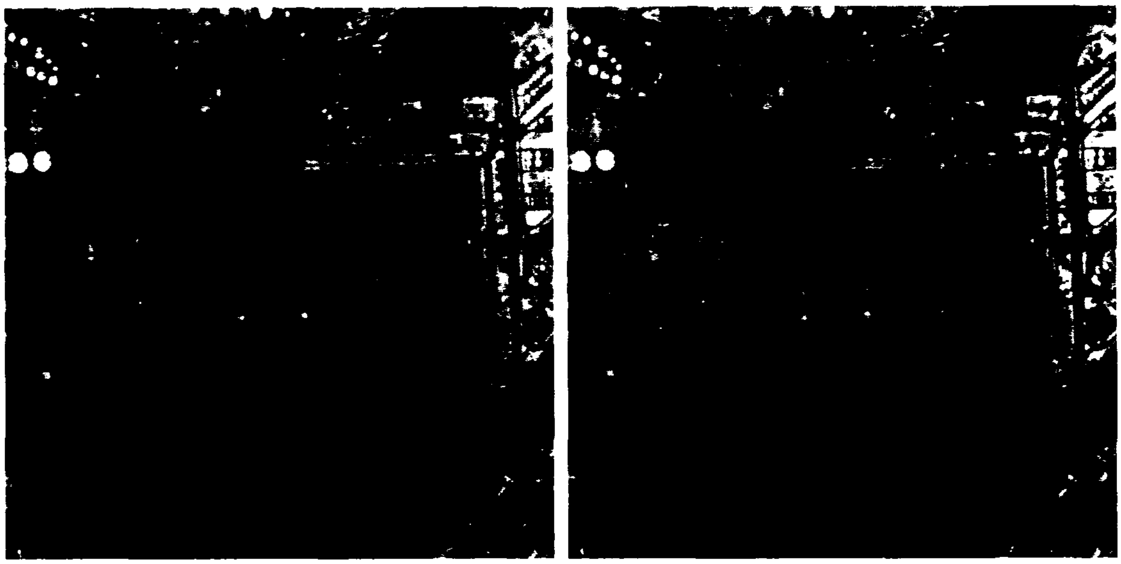
where c_{2s} is an adjustable constant. The advantage of this method is that S_2 is independent of M_{wc}, which means that to set the value of S_2 does not need the knowledge of M_{wc}. In addition, this method allows the user to have an estimation as to which part of the image information will be enhanced. For instance, by using $S_1 = \sigma_n$ and $S_2 = 10\sigma_n$, we can amplify all coefficients with a SNR range between 1 and 10.



(a) (b)

Figure 4.9 Contrast enhancement with $S_2 = 0.5M_{wc}$: (a) coefficient transformation function; (b) enhanced image.

Until now, the modification of wavelet coefficients is implemented only using a center symmetric transfer function. However, what effects can asymmetric transfer function impose on enhanced images? To answer this question, we apply only the positive or the negative part of the transfer function shown in Fig. 4.8(a) to the positive or negative coefficients, respectively. The results are presented in Fig. 4.10 where the contrast of the two images is obviously inferior to that of the image in Fig. 4.8(b). Moreover, the image artifacts (e.g., square shaped pattern and straight streaks) become more visible in the images in Fig. 4.10, especially in the image in Fig. 4.10(b).



(a) (b)

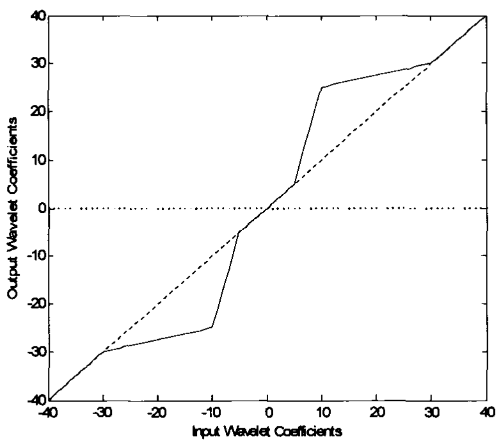
Figure 4.10 Images enhanced with asymmetric transfer function: (a) only positive coefficients processed; (b) only negative coefficients processed.

The line shape of the transfer function is also closely related to the characteristics of the enhanced images. For example, the parameter G can be adjusted to change the curvature (or nonlinearity) of the transfer function. Except that, the middle part (between S_1 and S_2) of the transfer function can be either a single nonlinear function (e.g., the transfer function in Eq. (4.29)) or a combination of linear and nonlinear functions. In addition, the middle part of the transfer function can be split into two segments, and both of them can be implemented using different mathematical functions. In our wavelet transform based contrast enhancement, a combination of two linear functions (two straight lines) is also designed to be used as the transfer function for coefficient modification, which is formulated as in

$$w_o = \begin{cases} w_i & \text{if } |w_i| < S_1 \\ (w_i - S_1)\cdot g_1 + S_1 & \text{if } S_1 \le |w_i| < S_m \text{ and } w_i > 0 \\ (w_i + S_1)\cdot g_1 - S_1 & \text{if } S_1 \le |w_i| < S_m \text{ and } w_i < 0 \\ (w_i - S_m)\cdot g_2 - S_m^o & \text{if } S_m \le |w_i| \le S_2 \text{ and } w_i > 0 \\ (w_i + S_m)\cdot g_2 - S_m^o & \text{if } S_m \le |w_i| \le S_2 \text{ and } w_i < 0 \\ w_i & \text{if } |w_i| > S_2 \end{cases} \tag{4.33}$$

with $g_1 = (S_m^o - S_1)/(S_m - S_1)$ and $g_2 = (S_2 - S_m^o)/(S_2 - S_m)$

where S_m and S_m^o are the coordinates of the joint of the two straight lines while g_1 and g_2 are the slopes of the two straight lines. One example of the transfer function in Eq. (4.33) is shown in Fig. 4.11(a). The two-segment configuration of the middle part of the transfer function provides a more flexible control over the modification of wavelet coefficients, such as which coefficients need to be modified and to which extent they need to be changed. One example of contrast enhancement using Eq. (4.33) is provided in Fig. 4.11(b). Compared to the results obtained by the nonlinear transfer function, the result in Fig. 4.12 is less noisy. However, some faint image feathers (e.g., sea waves) in the image in Fig. 4.11(b) are not enhanced as much as the images in Fig. 4.7 and 4.8.

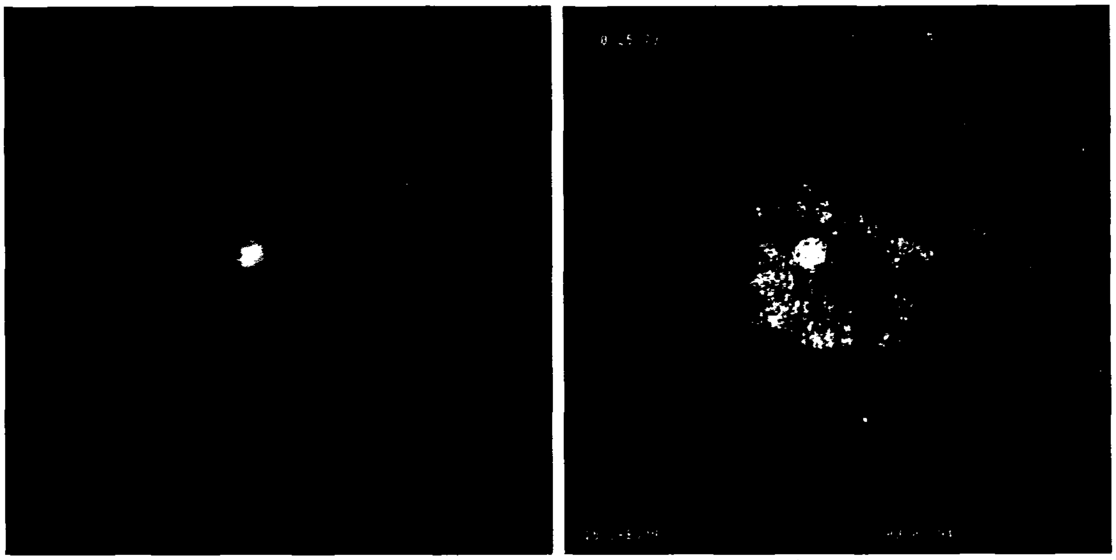


(a)

(b)

Figure 4.11 (a) Piece-wise linear transfer function; and (b) contrast enhanced image produced by linear transfer function.

All the previous discussions about the coefficient modification, including those in the literature, are only related to the detail coefficients while the treatment of the approximation coefficients seems to not be considered. This is because the approximation images only provide the information about the large-scale intensity variation across the whole image while the image contrast is mainly related to the local intensity variation. Therefore, processing of the approximation images will not help in improving the image



(a) (b)

Figure 4.12 Contrast enhancement with the processing of both detail and approximation coefficients: (a) original image; (b) enhanced image.

(a) (b)

Figure 4.13 Color image enhancement: (a) original image; (b) enhanced image.

contrast. However, it can improve the image quality in another way, e.g., brightness enhancement which can be realized by performing dynamic range compression on the approximation images, such as with the dynamic range compression techniques discussed in Chapter III. One example showing the processing of both detail coefficients and approximation coefficients is presented in Fig. 4.12. The detail coefficients are modified

Figure 4.14 Image enhancement produced by curvelet transform based contrast enhancement algorithm.

using Eq. (4.29) with $S_1 = \sigma_n$, $S_2 = M_{wc}$, and $G = 1/2$. The approximation coefficients are modified by using the Gamma correction with $\gamma = 0.5$. It can be observed that the image in Fig. 4.12 exhibits good luminance and contrast enhancement. Except enhancing grayscale images, the proposed algorithm can also process color images by only treating the intensity information of the image, and the enhanced color image is recovered using the same method discussed in Chapter III. One example of color image enhancement is shown in Fig. 4.13.

Finally, the proposed wavelet transform based contrast enhancement algorithm is compared with the curvelet transform based contrast enhancement algorithm described in [83]. The enhancement results produced by [83] are presented in Figure 4.14. The overall quality of those images is similar to that of our results. However, compared to curvelet transform, DT-CWT is much faster in processing.

CHAPTER V

ENHANCEMENT BY FUSION OF MULTIPLE IMAGES

The purpose of image fusion system is to integrate complementary and redundant information from multiple registered (aligned) source images to create a composite one that provides a more complete description of the scene than any individual source image. To develop a fusion algorithm, some of the challenges have to be properly dealt with. Considering the two registered images shown in Fig. 5.1, in the visual image it is very difficult to distinguish the person in camouflage from the background while the person is clearly visible in the infrared (IR) image. On the other hand, some objects in the scene (e.g. the fence), which are nearly imperceptible in the IR image, are readily discernible in the visual image. In order to properly merge the scene information from the source images, different types of scene information and various properties of the images need to be carefully classified and treated:

- *Complementary information*: some scene information appears in one source image but not in the others, e.g., the person in Fig. 5.1(b) or the fence in Fig. 5.1(a).
- *Common information*: scene information appears in all source images. However, same scene information may look different in different source images because of the dissimilar (or even opposite) contrast. For instance, the bushes along the bottom of the images and the house roof are represented differently in the two images in Fig. 5.1.
- *Properties of images*: source images captured from different types of sensors which generally have different sensing capability, different dynamic range and different resolution, noise, etc.

It has been commented that image fusion can be implemented at either the pixel level or the feature level based on MR decomposition. This chapter presents the improved MR image fusion algorithms which employ DT-CWT to obtain the MR decomposition of input images. Both pixel-based and region-based image fusion algorithms have been investigated and both of them will be discussed in this chapter. It should be noted that we confine our discussion to fusion of two registered source images with the output of a single fused image.

(a) (b)

Figure 5.1 Example of multisensor source images to be fused: (a) visual image; (b) IR image. Images courtesy of TNO Human Factors Institute, The Netherlands.

5.1 Pixel-Based Image Fusion

5.1.1 Algorithm

An improved pixel level image fusion technique based on (dual tree complex) DT-CWT MR decomposition has been developed. The fusion scheme is shown in Fig. 5.2 which processes two source images (A and B) and consists of 6 modules including both forward and inverse DT-CWT transforms. The *match measure* extracts image similarity information from the source images in the spatial domain while *activity measure* computes the pixel importance from the MR decompositions of the sources. The obtained information is then used by the *decision map* and *combination* to create the MR decomposition of the fused image. Finally, the inverse DT-CWT reconstructs the fused image. Each module is described and discussed in details in this section.

- *DT-CWT*: This module performs MR decomposition of the source images using DT-CWT. The detail coefficients will be used in the next step to compute an activity measure. If any source image is a color image, it has to be converted to a grayscale image prior to the wavelet transform. This is because one of the source images may be either a monochromatic or a grayscale image for many applications. In addition, even for fusion of color images, the separate processing of each spectral band is not practical because multi-band process needs much more computations and may readily produce image artifacts. To obtain optimized image fusion result, the source images are generally decomposed to level 5 in our experiments. If the source image size does not allow the MR decomposition up to 5 levels (i.e. the width or height of the source image is not equal to $a \cdot 2^5$ where a is a positive integer.), the source image will be resampled to the closest size which allows such MR decomposition. Moreover, if image enhancement of the source

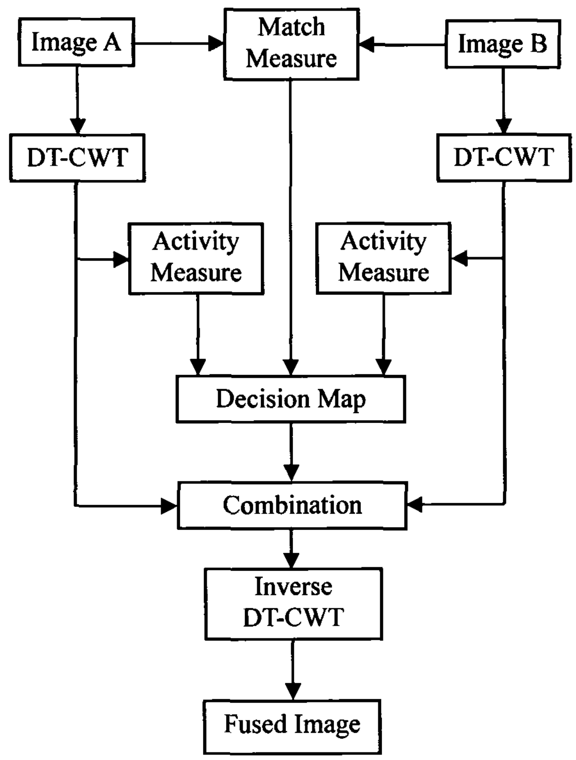


Figure 5.2 Pixel-based MR fusion scheme.

images is needed, it should be done before DT-CWT transform. Examples of level 1 subband images of the image shown in Fig. 5.1(a) are presented in Fig. 5.3. Higher grayscale value represents a larger absolute value of the complex wavelet transform coefficient, which then indicates a sharp intensity change in the original image. Moreover, the grayscale value at each pixel location is also independent on the orientation of the salient image feature. That is why the same image feature may appear different in different subband images which are produced by wavelets with various orientations.

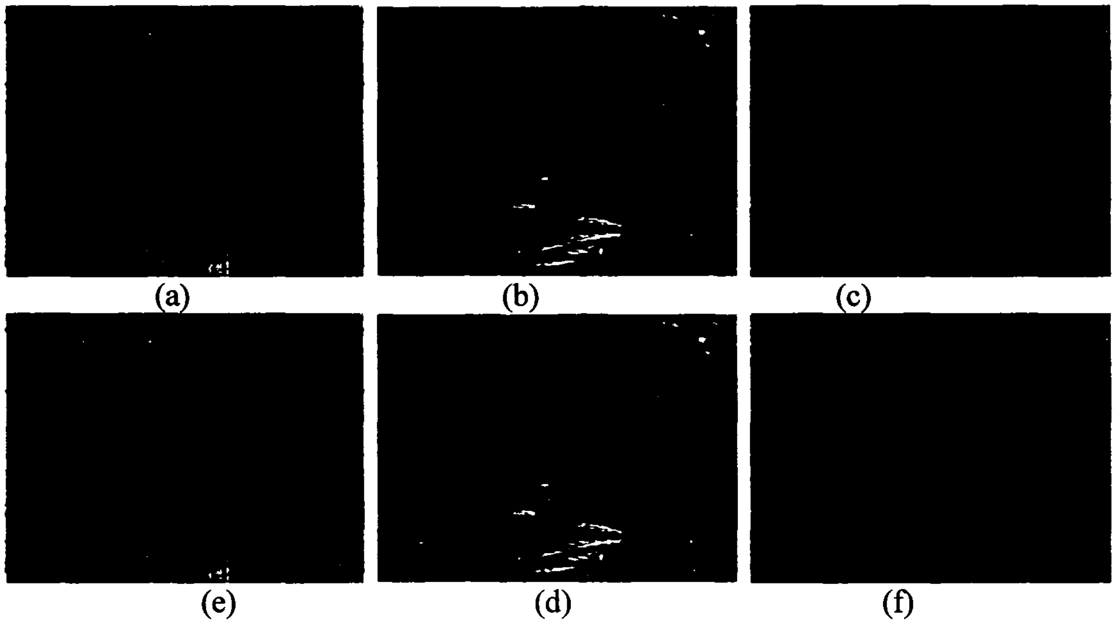


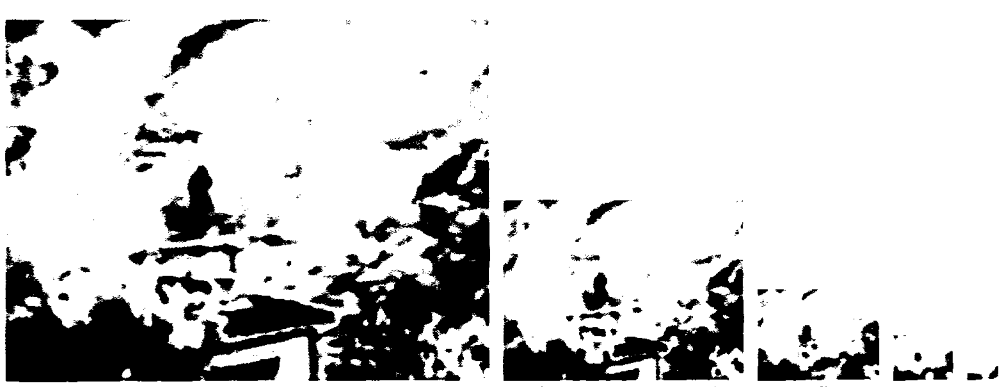
Figure 5.3 DT-CWT subband images of six orientations at level 1: (a) 15°; (b) 45°; (c) 75°; (d) -15°; (e) -45°; (f) -75°.

- *Match measure*: This module computes the similarity between corresponding pixels in both source images. Match measure is important for multi-sensor image fusion because the image from one imaging sensor may provide different type of information at each pixel location when compared with the image recorded by the other sensor. This issue has been discussed previously in this chapter. Match measure is critical in our fusion algorithm because it determines where the source images differ (or similar) and to which extent so that source images can be combined in an appropriate way. In order to properly compare the two corresponding pixels, neighborhoods surrounding the pixels should also be considered. In our image fusion scheme, match measure is defined as a normalized correlation averaged over a neighborhood of the samples as in:

$$m_j^{AB}(x,y)=\frac{2\sum\limits_{(x+m,y+n)\in w} I_j^A(x+m,y+n)I_j^B(x+m,y+n)}{\sum\limits_{(x+m,y+n)\in w}(I_j^A(x+m,y+n))^2+(I_j^B(x+m,y+n))^2} \tag{5.1}$$

where I_j^A and I_j^B are successively subsampled source images at level j, and w is the 5×5 neighborhood. The subsampled images should have the same matrix sizes as those of wavelet detail coefficients. The value of m^{AB} is an estimation of the similarity of image features at pixel level, e.g., $m^{AB}=1$ indicates identical patterns, $m^{AB}<1$ shows less similarity between features, $m^{AB} = 0$ indicates that the grayscale values of all the pixels in the two neighborhoods are zero.

We propose to compute match measure in the spatial domain instead of the transform domain for more accurate estimation of pattern similarity and less computational complexity. The conventional approach to calculating match measure is to apply Eq. (5.1) to all subbands of transformed source images in the wavelet domain. For instance, there are six subbands at each decomposition level produced by DT-CWT, thus Eq. (5.1) will be used six times at each level in the conventional method while the proposed approach only uses Eq. (5.1) once at each level. The match measure result of subsampled source images (Fig. 5.1) at level 1 is presented in Fig. 5.4 as a grayscale image. The values of match measure have been linearly scaled to the range [0 255] for display. The grayscale value at each pixel location represents the similarity between the neighborhoods which are surrounding the corresponding pixels on both subsampled source images at level 1. Higher grayscale value indicates higher degree of resemblance. It can be seen that the person in camouflage is one of the most dissimilar features between the two images in Fig. 5.1. The proposed match measure method will be further discussed in Section *5.1.2*.



(a) (b) (c) (d) (e)

Figure 5.4 Multi-level match measure results are shown as a grayscale images: (a)-(e) level 1 through level 5.

- *Activity measure*: This module computes the 'saliency' of each pixel in the transform domain. The meaning of saliency depends on the properties of source images and the objective of particular fusion application. Based on the fact that the human vision system (HVS) is primarily sensitive to local contrast changes (e.g. edges), most fusion algorithms compute the activity measure as some sort of energy calculation. In our fusion scheme, the magnitude of the detail coefficients is used to calculate the activity measure as in:

$$E_j^A = \sum_{k=1}^{6} \left|A_j^k\right|, \quad E_j^B = \sum_{k=1}^{6} \left|B_j^k\right| \tag{5.2}$$

where A_j^k and B_j^k are the complex detail coefficients of the two source images at level j and orientation k; E_j^A and E_j^B are activity measures at level j for both source images, which are the summation of the magnitudes of the detail coefficients of all 6 subbands at each decomposition level. Eq. (5.2) also has its physical meaning which can be understood by considering that each coefficient of a MR decomposition has a set of 'family-related' components in other orientation bands and other levels. They represent the image feature(s) at the same (or nearby) spatial location in the original image. Therefore, it is reasonable to take into account the coefficients in all subbands when the image property at one spatial location is being determined.

The activity measures obtained in Eq. (5.2) may need to be low-pass filtered to suppress the salient features caused by the impulsive noises in source images. In our fusion algorithm, a spatial convolution with a 3×3 Gaussian mask with a standard deviation of 0.5 is used for this purpose. Similar to the visualization of match measure, the activity measures of the subsampled source images (Fig. 5.1) at level 1 are presented in Fig. 5.5 as grayscale images. Higher grayscale indicates higher saliency which means a significant intensity change in source images. More detailed discussion of activity measure is provided in Section *5.1.2*

(a) (b)

Figure 5.5 Activity measure result of images shown in Fig. 5.1: (a) result for visible image, (b) result for IR image.

- *Combination*: This module performs the combination of the MR decompositions of the two source images to create a composite MR decomposition that will be transformed to spatial domain. In the implementation of our fusion scheme, a linear combination of the detail coefficients of the source images is used to obtain the detail coefficients of the fused image, which can be expressed as in:

$$F_j^k = C_j^A \cdot A_j^k + C_j^B \cdot B_j^k \tag{5.3}$$

where F_j^k is the detail coefficient of fused image at level j and orientation k; C_j^A and C_j^B are the weight factors at level j, which are dependent on match measure and activity measure but independent on orientation. The dependency of weigh factors on match and activity measures is determined by the decision map.

The combination of approximation coefficients of source images is conducted in a different manner compared to detail coefficients due to their different physical meanings. For example, detail coefficients with large magnitudes represent sharp intensity changes in images, such as edges, spots, lines and region boundaries. The approximation image, however, is a coarse representation of the original image and has reserved some of its properties like the average intensity and texture information. Thus, approximation coefficients with large magnitudes do not necessarily correspond to salient features. Accordingly, activity measure is not suitable for approximation images.

The approximation coefficients of the fused image are usually calculated using a simplified version of Eq. (5.3) where both weight factors are real constants (e.g. 0, 1/2, and 1) which are independent on match and activity measures but dependent on the quality of fused image and is application-orientated. In practice, a simple arithmetic average (i.e. both weight factors are 1/2) is often used to yield the composite approximation coefficients. However, there really are technical reasons behind this simple averaging, which are based on the assumption that the source images contain additive Gaussian noise and that, the given decomposition level is high enough. Important image features have already been captured by detail coefficient. Therefore, the approximation images contain mostly noise and averaging them reduces the variance of the noise while ensuring that an appropriate mean intensity is maintained. This issue is discussed in Sections *5.1.2*.

- *Decision map*: This module is the key point in image fusion, which determines the values of the weight factors in Eq. (5.3) based on the match and activity measures. Various decision making schemes have been developed [85-94]. However, most of them use single thresholding and consider only two cases. Thus no smooth transition is made between those two cases. In our image fusion scheme, we propose a double thresholding method to determine the weight factors in Eq. (5.3) in the following three cases:

$$\text{Case 1:} \quad \begin{cases} C_j^A = \dfrac{E_j^A}{E_j^A + E_j^B}, \\ C_j^B = \dfrac{E_j^B}{E_j^A + E_j^B}, \end{cases} \quad if \quad m_j^{AB} \geq 0.9$$

$$\text{Case 2:} \quad \begin{cases} C_j^A = \dfrac{E_j^A}{E_j^A + E_j^B} \cdot T + (1-T) \cdot W_A, \\ C_j^B = \dfrac{E_j^B}{E_j^A + E_j^B} \cdot T + (1-T) \cdot W_B, \end{cases} \quad if \quad 0.7 < m_j^{AB} < 0.9$$

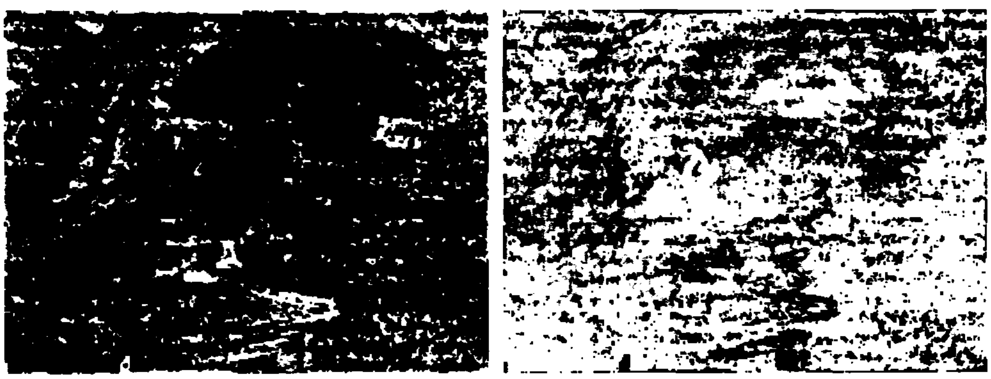
$$where \quad T = \frac{m_j^{AB} - 0.7}{0.9 - 0.7}, \quad and \quad \begin{cases} W_A = 1 \quad if \quad E_j^A \geq E_j^B, \quad otherwise \quad 0 \\ W_B = 1 \quad if \quad E_j^A < E_j^B, \quad otherwise \quad 0 \end{cases}$$

$$\text{Case 3:} \quad \begin{cases} C_j^A = 1 \quad C_j^B = 0, \quad if \quad m_j^{AB} \leq 0.7 \quad and \quad E_j^A \geq E_j^B \\ C_j^A = 0 \quad C_j^B = 1, \quad if \quad m_j^{AB} \leq 0.7 \quad and \quad E_j^A < E_j^B \end{cases}$$

In case 1, the match measure is high which represents very highly similar patterns, and the fused coefficient is the linear combination of the coefficients of both input images with the weight factors determined by the relative relation between the two activity measures. On the contrary, in case 3, the match measure is low, which represents very dissimilar patterns, and thus only the more salient feature (larger activity measure) is included in the fused image. Between these two extreme cases, case 2, which represents medium similarity, provides a smooth transition with the weight factors set by a linear combination of the two extreme cases. The coefficient T in case 2 modulates the relative importance between those two terms based on the relative position of match measure with respect to the two thresholds. If the match measure is close to 0.9, the weight factors are determined in a way more similar to case 1. Otherwise, if the match measure is close to 0.7, the weight factors are determined in a way more similar to case 3. Finally, the expressions discussed above should ensure that the sum of the two weight factors C_j^A and C_j^B be 1.

Examples of the decision maps of the subsampled source images (Fig. 5.1) are shown in Fig. 5.6. It can seen that more image information is merged in the fused image from the

IR image than the visual image since the decision map image since the IR image contains much more high-lighted pixels than those in the decision map for the visual image. This is because the visual image has a larger area composed of dark pixels compared to the IR image, which contains less scene information than those bright areas in the IR image. The gray areas appearing in both decision map images represent those similar areas between the two source images. In addition, a large number of discrete bright spots are uniformly scattered across almost the whole decision map image for the visual image while they are paired by dark spots in the decision map image for the IR image. These spots are created by either small-scale salient image features or by image noise. The decision map module is discussed in Section *5.1.2*.



(a) (b)

Figure 5.6 Decision maps for the level 1 wavelet coefficients of the source images: (a) visual image; (b) IR image.

- *Inverse DT-CWT*: This module concludes the image fusion process by performing inverse DT-CWT transform on the composite wavelet coefficients yielded by the *Combination* module to create the fused image. After this step, additional image processing may be needed as discussed below.

A color restoration process is needed to transfer the chromatic information of the source images to the fused image. If only one of the source images is a color image, the color information is certainly obtained from that image. However, if both source images are color images, the chromatic information should be obtained from one or both of the source images: If one, which image should be used? In our image fusion experiments, only the source image which contains more color information than the other will be used for color restoration so that the color artifacts can be avoided. Here the meaning of the word 'more' can be understood by considering a fusion of two visual images which are captured on exactly the same scene. One of the images is recorded in daytime while the other is captured at night with insufficient illumination. Obviously, the first image contains more color information of the scene for the second image has many dark regions where the color information is lost or degraded by noise.

A linear color restoration method is used in our algorithm, which can be expressed as in:

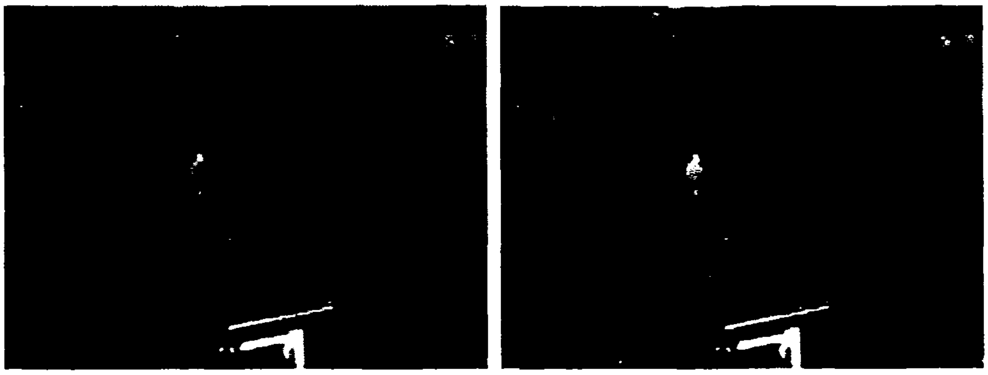
$$I_i^F = I^F \cdot \frac{I_i^A}{I^A}, \quad i = r, \ or \ \ g, \ or \ \ b \tag{5.4}$$

where image A is assumed to be the color image which contains more color information. Its RGB bands are I_i^A, and its intensity image is I^A which is used in the fusion process. Similarly, I^F is the intensity image (produced by inverse DT-CWT) of the fused image whose RGB components are I_i^F .

Since image fusion tends to degrade the image contrast, the fused image can be post-processed by some image enhancement algorithm to improve its quality, like the contrast enhancement and nonlinear image enhancement techniques discussed in the previous chapters.

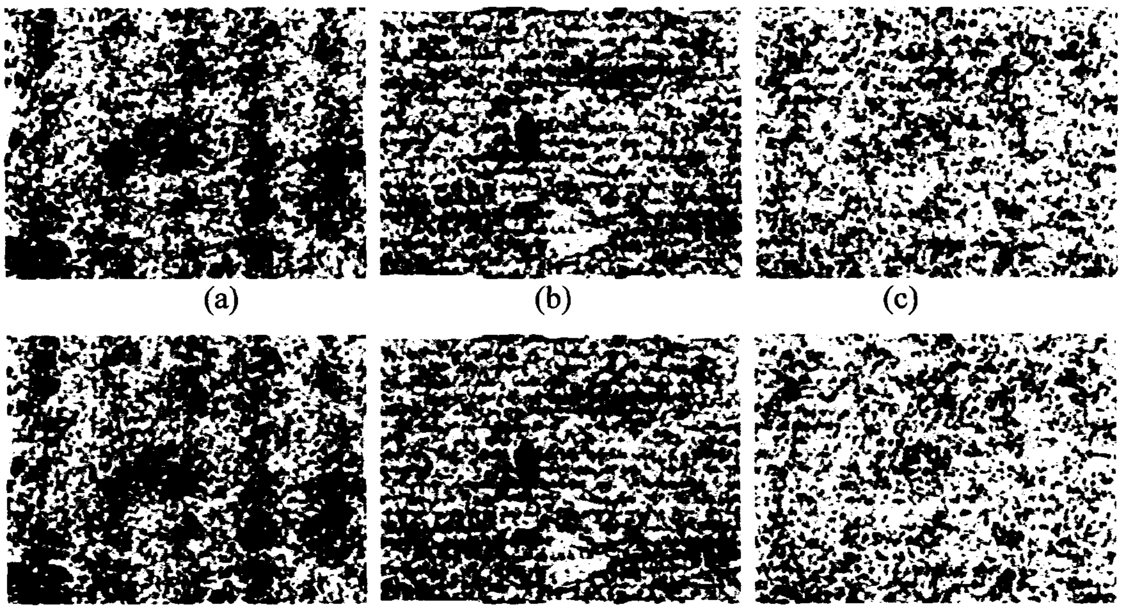
5.1.2 Results and Discussion

The image fusion results of the two source images shown in Fig. 5.1 are provided in Fig. 5.6. in which figure (a) is produced by our proposed fusion scheme and figure (b) is yielded by the well-known Maximum Selection scheme [92], which is implemented in our image fusion experiments using DT-CWT for the MR decomposition of source images. The two fused images look almost the same except some slight differences. One of them appears on the road that is located close to the upper left corner of the image. A more careful observation can reveal that Fig. 5.7(b) has slight higher contrast than the image in Fig. 5.7(a) which, however, is less noisy than the other. This is because the Maximum Selection rule is more suitable for selecting all salient features which also include noise while our fusion scheme tends to fuse the features from both source images, which can result in less noise and lower contrast.



(a) (b)

Figure 5.7. Image fusion results produced by: (a) the proposed scheme and (b) the Maximum Selection scheme.



(a) (b) (c)

(d) (e) (f)

Figure. 5.8 Match measure results obtained in wavelet transform domain between corresponding subbands: (a) 15°; (b) 45°; (c) 75°; (d) -15°; (e) -45°; (f) -75°.

In our proposed fusion scheme, the match measure is computed on the grayscale source images in spatial domain while, in the existing fusion algorithms, it is commonly computed on MR decomposed images in transform domain. Some examples of the match measure results obtained in wavelet transform domain are presented in Fig. 5.8. Those results are very different from what is shown in Fig. 5.4. However, the image fusion results shown in Fig. 5.7 demonstrate how our match measure computation method works appropriately without any unusual image artifacts or distortions found related to the new match measure approach. This result and higher computational efficiency provide our method some advantages over other match measure approaches. Moreover, this result also shows strong spatial correlation between the images in spatial domain and their transformed version in wavelet domain.

In wavelet transform based fusion schemes, the absolute values of the wavelet coefficients are commonly used to calculate activity measures. However, there are different methods to compute activity measures using wavelet coefficients. The mostly used method is to use the absolute value (magnitude) of the wavelet coefficient in each subband as the activity measure. Thus activity measure is orientation dependent while the activity measure defined in our fusion scheme (Eq. (5.2)) is independent of the orientation of wavelets. Based on the consideration that an image pixel is determined by all subbands at the pixel location. This type of activity measure calculation is intended to make all subbands at each pixel location behave in the same way while they are processed in image fusion. Therefore, there is only one decision map for each source image at each MR decomposition level in our fusion scheme, but there are six decision maps for each source image at each level in other fusion schemes which use DT-CWT and orientation dependent activity measures. In our image fusion experiments, the orientation dependent activity measure is also implemented in our fusion algorithm to investigate the difference between these two types of activity measure definitions. The experimental results are provided in both Fig. 5.9 and Fig. 5.10. In Fig. 5.9, the six decision maps for the visual image shown in Fig. 5.1(a) are presented while the final fused image is provided in Fig. 5.10. It is obvious that the decision maps shown in Fig. 5.9 are rather different from the decision map shown in Fig. 5.6(a) and the decision maps are not the same for all subbands. However, the fused image shown in Fig. 5.10 is almost the same as the one in Fig. 5.7(a) which is produced by the orientation independent decision map shown in Fig. 5.6(a). It is unexpected that the difference between these two types of decision maps does not result in any significant difference between the fused images. More experimental results presented in the remainder of this section will show similar phenomena.

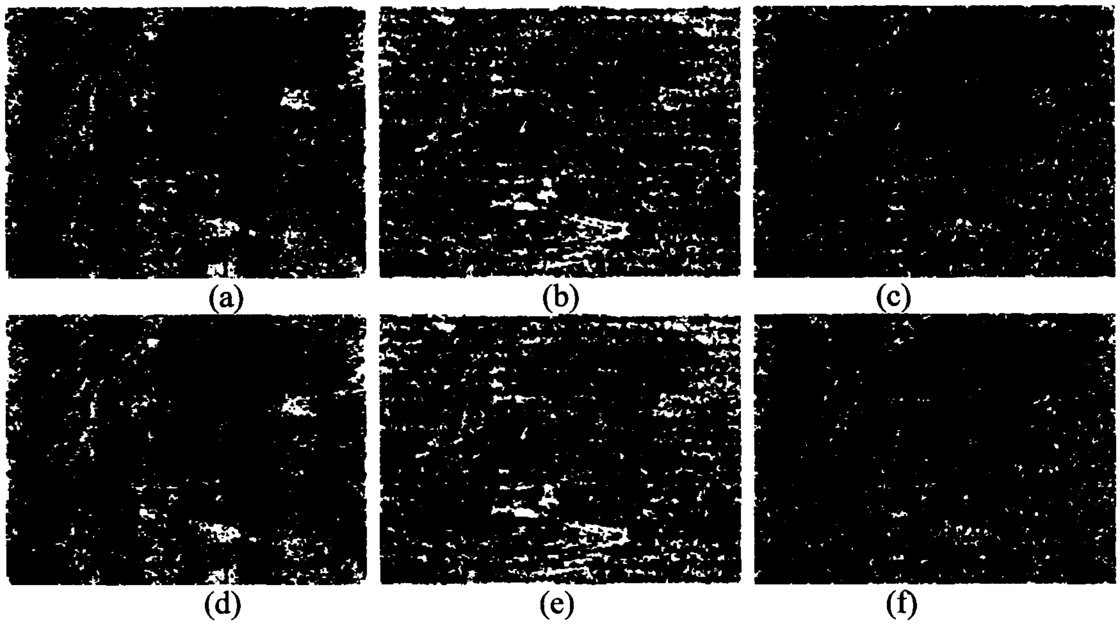


Figure 5.9 Decision maps for the six subbands of the visual image at level 1: (a) 15°; (b) 45°; (c) 75°; (d) -15°; (e) -45°; (f) -75°.

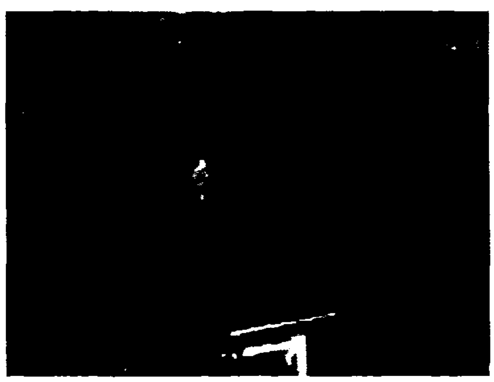
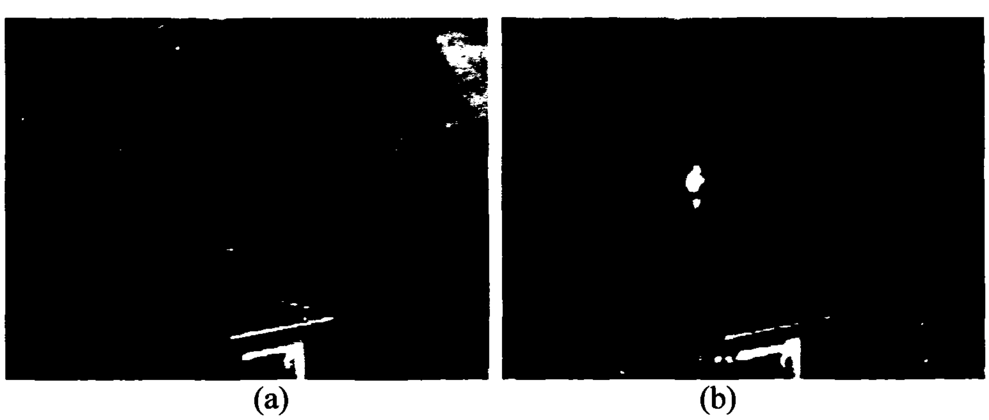


Figure 5.10 The fused image produced by subband dependent activity measure.

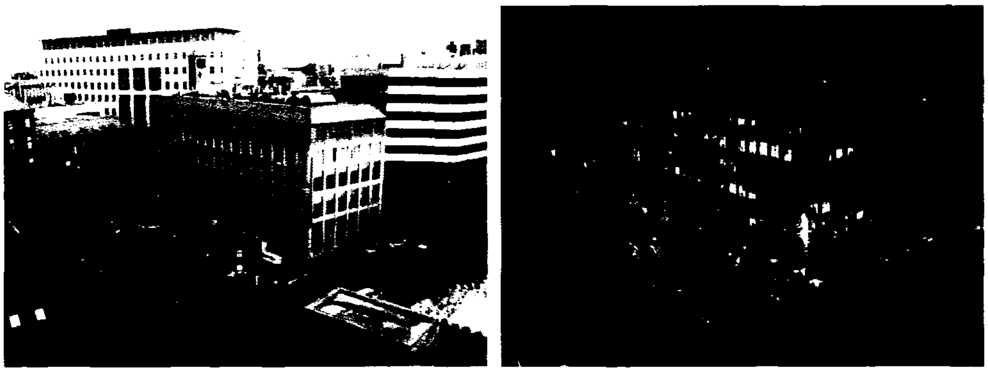
It has been mentioned in Section *5.1.1* that the combination of the approximation coefficients of source images are application and image quality orientated. This can be



(c) (d)

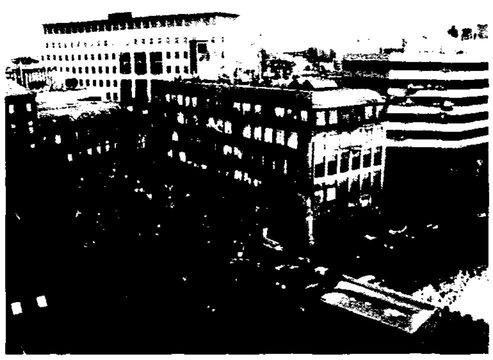
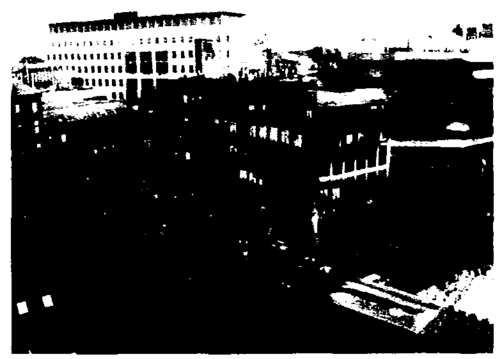
Figure 5.11 Image fusion results with composite approximation coefficients obtained in different ways: (a) from visible image; (b) from IR image; (c) determined by decision map and (d) maximum selection.

easily understood using the fused images shown in Fig. 5.11, in which the fused images are produced with approximation coefficients selected in different ways. The fused images in Fig 5.11(a) and (b) are created only using the approximation coefficients from image (a) and (b) in Fig. 5.1, respectively. In 5.11(a), the person becomes visible but some features in the IR image are missing in the image. In addition, the halo effect around the person is the most significant among all the fused image examples. On the contrary, in 5.11(b), although the features in the IR image are strongly represented with the removal of the halo effect, contrast reversal is very severe and makes the image look unnatural. 5.11(c) is created using the highest level decision map to select the approximation coefficients from both source images. The image quality would be very good if those unexpected dark regions could be eliminated. This image also shows that the decision maps created by the fusion scheme are not applicable to approximation coefficients. The approximation coefficients of 5.11(d) are obtained from both source images using the maximum selection rule. It has the highest visual quality among these



(a) (b)


(c) (d)

(e) (f)

(g)

Figure 5.12 Fusion of multi-illuminance color images. (See the description in page 90).

four images in Fig. 5.11. Compared to the image in Fig. 5.11(a) which is obtained using the averaged approximation coefficients from both source images. Image in Fig. 5.11(d) shows equivalent visual quality with no halo effect. However, some features are less represented and the contrast in some areas is degraded in Fig 5.11(d). Therefore, the averaging of source image approximation coefficients is still more preferred in most existing fusion schemes.

(a) (b)

Figure 5.13 Color image fusion obtained by: (a) maximum selection scheme; and (b) orientation dependent activity measure.

Fig. 5.12 shows an example of the fusion of multi-illuminance color images. The two source images are presented in Fig. 5.12(a) and 5.12(b) which are captured in the daytime and at night (i.e. under different illumination conditions), respectively. The image fusion process is conducted on grayscale images and color restoration is applied to convert the grayscale fused image to a color image. Figures 5.12(c) through (f) are obtained using the same image fusion process as the Fig.5.11(a) through (d) except the color restoration. Different selection methods are applied to choose the approximation coefficients for the fused images. Figure 5.12(g) is produced by averaging the approximation coefficients from both sources, which is the same as Figs.5.7(a) and 5.7(b). It can be easily determined that Fig.5.12(c) and 5.12(e) have the highest visual quality among all the samples and they are virtually the same image with only trace of difference. This result is different from what is observed in Fig. 5.11. This discrepancy lies in the fact that in the previous example, both source images contribute quite equally and both images have similar average intensity while in the second example, the day time source image, which is much higher in average intensity than the night image, contributes much more than the other image. In addition, the IR image in the previous example shows large scale image features which are contrast reversed compared to the other source image, but this problem does exist in the second example. Finally, another two fused images are presented in Fig. 5.13 for comparison, which are obtained using maximum selection scheme and orientation dependent activity measure, respectively. The result agrees with what is discussed previously in this section, which confirms the performance of our proposed fusion scheme and validity of the orientation independency of the activity measure and decision map.

5.2 Region-Based Image Fusion

5.1.1 Algorithm

A new region based MR image fusion algorithm has been proposed in this dissertation. The framework of this image fusion algorithm is presented in Fig. 5.14. The major contribution of the proposed algorithm is that pyramid image segmentation is first employed to achieve the MR segmentation of the multilevel match measure images which are computed at pixel level on source images in spatial domain. Then the MR segmentation of source images is conducted using the same segmentation maps obtained from the segmentation of match measure images. In the second step, MR decomposition of source images is performed using DT-CWT. In the third step, the wavelets coefficients of source images are combined in a region-based manner through a double-thresholding scheme governed by the match and activity measures of the corresponding regions. Finally, the fused image is reconstructed via the inverse wavelet transformation of the composite wavelet coefficients obtained in the last step. Experimental results demonstrate that the segmentation of match measure images better serves the purpose of image fusion.

The fusion scheme shown in Fig. 5.14 is similar to the pixel-based method shown in Fig. 5.2 because region-based method can be considered as an extension or generalized form of pixel-based method. Region-based image fusion deals with the merge of regions of the source images while pixel-based method works on fusion of individual pixels of the sources. If the region size decreases to a pixel, then the region-based method becomes a pixel-based method. The major difference between these two methods lies in the introduction of image segmentation in the region-based techniques. The essence of region-based fusion is the effort to integrate the objects (or the constituent parts of the objects), instead of individual pixels, to form the sources into the composite image. Therefore, image segmentation is the key component in the region-based fusion scheme, which is applied to find various objects in source images by classifying an image into different parts based on the intensity distribution and texture. Below, each functional modules included in the fusion scheme is described in detail. Since most of the modules were discussed in Section 5.1, the focus is on the segmentation module and its interaction with the other modules.

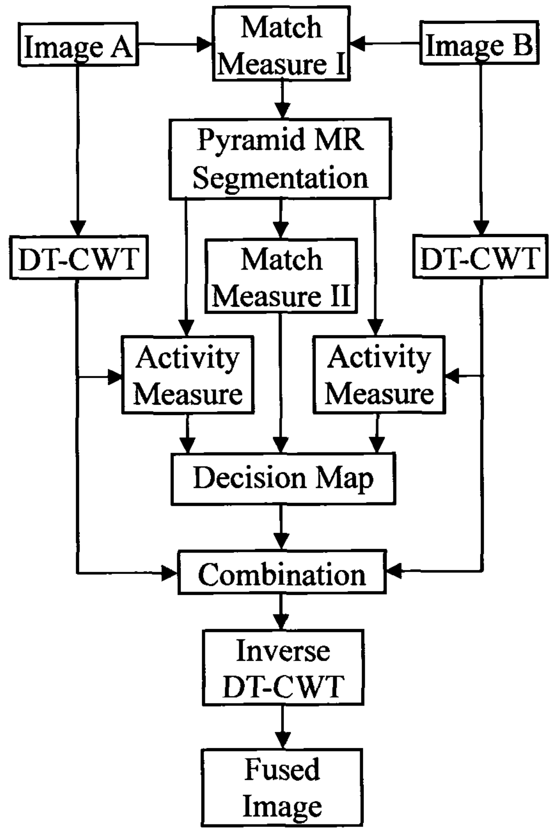


Figure 5.14 Region-based MR fusion scheme.

Match measure I: This module computes the match measure at pixel level and is exactly the same as our proposed match measure module discussed in the pixel-based fusion scheme. The match measure result is then used as the input, which contains information from both sources, for MR segmentation to classify different regions on the source images in a joint way.

Pyramid MR segmentation: In this module, we propose performing multi-resolution segmentation on pixel-level match measure images instead of the source images themselves. Generally, MR segmentation is conducted on source images separately or jointly in either spatial domain or transform domain. Compared to those conventional techniques, the proposed MR segmentation on match measure images has some advantages. First, it is more efficient because it does not segment both source images separately. Match measure images contain the information from both sources. Second, it serves the goal of image fusion better because the fusion of two pixels or two regions is firstly guided by the match measure and secondly by the activity measure. The segmented regions based on match measure have similar match measure values at all pixel locations within a region. Therefore, the pixels in a region can be treated in the same way while the creation of image artifacts can be controlled. However, the current segmentation methods do not consider match measure when segmenting source images, which may create a region that has different match measures at different pixel locations within the region. Thus image artifacts can be created when two regions are combined with all the pixels in a region are governed by the same rule.

The proposed MR segmentation has been implemented using a technique similar to the method of a linked pyramid which was first described by Burt et al. [121]. It consists of a MR decomposition of an image with the bottom level containing the full-resolution image and each successive higher level being a filtered and subsampled version derived from the level below it. The various levels of the pyramid are 'linked' by means of so-called 'child–parent' relations (see Fig. 5.15) between the sample pixels; such child–parent links are established during an iterative processing procedure to be described as follows. First, an approximation pyramid is produced by 2D low-pass filtering and subsampling. Then, child–parent relations are established by linking each pixel in a child level to one of the pixels in the next higher parent level which has the closest gray value (or in some other pixel attribute). The attribute values of the parents are then updated using the arithmetic average values of their children. The process of linking and updating is repeated until convergence occurs or until the preset pyramid level. Finally, the pixels in the top level of the pyramid are labeled as roots. Every root and the pixels which are connected to it induce a tree in the pyramid. The leaves of each tree correspond to pixels which are defined a segmentation region in the full resolution image. Thus, the linked pyramid provides a framework for an iterative process of image segmentation. For example, in Fig. 5.15, pixel Z at level 3 is a root which represents in the level 0 (full resolution image) a segment composed of pixels a through h. There exist many variations on the scheme: in the way the initial pyramid is built, in the manner pixels are linked to each other, in determining when pixels should be declared as roots, in the size of the neighborhood in which children can look for a parent to link to, in the attribute that is being used (e.g., gray value, edge, local texture), etc.

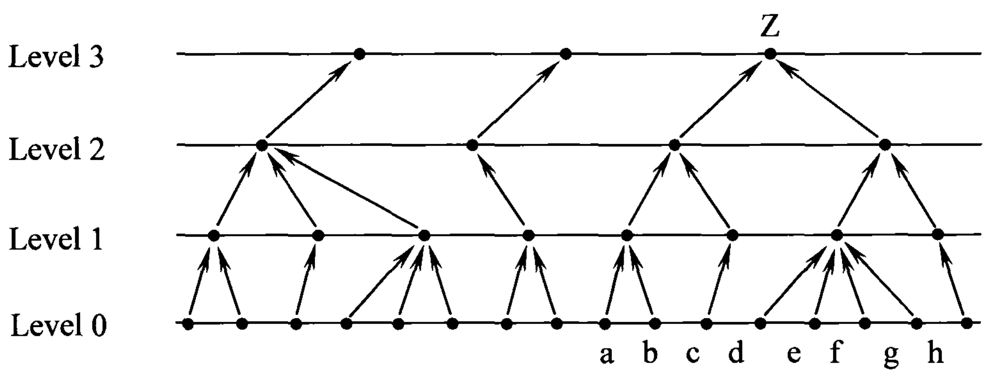


Figure 5.15 A diagram illustrating linked pyramid.

The details of the implementation of the MR segmentation of multilevel match measure images are provided as follows. Our method follows the classical "50% overlapping 4×4" structure. This means that each parent is derived from the pixels in the 4×4 neighborhood immediately below it, and this neighborhood overlaps 50% of that of its 4 neighbors. Thus, each pixel has 16 candidate children and each child up to 4 candidate parents. The bottom of the pyramid corresponds to level zero and, for simplicity, is assumed to be of size $N \times N$ with N being a power of 2. The maximum height of the pyramidal structure is considered to be $K = \log_2 N - 1$. In practice, the highest level of the pyramid is a preset value and generally less than K.

At each level k, the pixels are indexed by the vector $\boldsymbol{v} = (x, y)$, where $x, y = 0, \ldots, N/2^j - 1$. We use $\boldsymbol{C}(\boldsymbol{v})$ to represent the set of candidate children of pixel $\boldsymbol{v}$ at level $k > 0$, which can be expressed as in:

$$C(v) = \{(x', y') \mid x' \in \{2n-1, 2n, 2n+1, 2n+2\},\ \ y' \in \{2m-1, 2m, 2m+1, 2m+2\}\}.$$

Similarly, we use $\boldsymbol{L}(\boldsymbol{v})$ to represent the set of candidate parents of pixel $\boldsymbol{v}$ at level $k < K$:

$$L(v)=\left\{(x',y')\mid x'\in\left\{int\left(\frac{1}{2}(x-1)\right),int\left(\frac{1}{2}x\right),int\left(\frac{1}{2}(x+1)\right)\right\},\right.$$
$$\left.y'\in\left\{int\left(\frac{1}{2}(y-1)\right),int\left(\frac{1}{2}y\right),int\left(\frac{1}{2}(y+1)\right)\right\}\right\},$$

where $int(\cdot)$ represents the integer part of the enclosed value. The set of pixels to which the pixel v is connected at the bottom level is called *receptive field.*

Consider an input image of level j match measure image m_j^{AB}. The pyramid link based MR segmentation algorithm includes three steps.

1. *Initialization*

We associate each pixel $\boldsymbol{v}$ at level zero of the pyramid, $p^0(\boldsymbol{v})$, with the match measure value $m_j^{AB}(\boldsymbol{v})$, and each pixel $\boldsymbol{v}$ at level $k > 0$ with the average value computed from the values of its candidate children as in:

$$p^k(\boldsymbol{v})=\frac{1}{16}\sum_{\boldsymbol{v}'\in \boldsymbol{K}(\boldsymbol{v})}p^{k-1}(\boldsymbol{v}').$$

The area of the receptive field of the pixel $\boldsymbol{v}$ at level k is $S^k(\boldsymbol{v})$.

2. *Linking*

(a) Pixel linking and root labeling.

For each child, a favorable parent is sought among the candidate parents in such a way that it is linked to its most 'similar' parent or it becomes a root if no parent is found. Here, 'similarity' is based on the difference in pixel value. This difference is computed between the child and each of its four candidate parents. A link is established with the parent that minimizes that distance. If more than one candidate parent minimizes it, we arbitrary pick one of them.

In this pyramid segmentation algorithm, the root labeling is conducted within the linking step. That is, when linking to a parent, if the pixel value difference is above some threshold, the link is not established and the pixel is labeled as a root (i.e., it is not considered to be a child any more). One advantage of this method is its speed: a single operation will identify all roots. A disadvantage is that it is not clear beforehand how many roots (and, therefore, how many segments) will be found. Defining a good root labeling threshold is not straightforward. When the threshold is too high, few pixels become roots, whereas many pixels are labeled as root if the threshold is too low. Based on the segmentation experiments, we use a threshold $T = 0.8\Delta p$ where Δp is the standard deviation of the match measure image.

(b) Updating of area $S^k(\boldsymbol{v})$ and pixel value $p^k(\boldsymbol{v})$.

The receptive field area and pixel value of each parent are recomputed using only the children that are linked to it:

$$S^k(\boldsymbol{v}) = \sum_{\boldsymbol{v}' \in \boldsymbol{K}(\boldsymbol{v})} S^{k-1}(\boldsymbol{v}'),$$

$$p^k(\boldsymbol{v}) = \frac{\sum_{\boldsymbol{v}' \in \boldsymbol{K}(\boldsymbol{v})} p^{k-1}(\boldsymbol{v}') S^{k-1}(\boldsymbol{v}')}{S^{k-1}(\boldsymbol{v})},$$

where $S^0(\boldsymbol{v}) = 1$ for all $\boldsymbol{v}$ at level zero of the pyramid.

(c) Iteration of (a) and (b) until convergence.

3. *Classification*

The actual segmentation is obtained by classifying the tree structure of the created links. At each level k, all pixels that are connected to a common root are classified as a single region. In this way, at each level k, we obtain a segmented image $\boldsymbol{R}^k$ which contains all the regions at this level.

An example of the MR segmentation of the multi-level match measure images, shown in Fig. 5.4, is presented in Fig. 5.16. The images are classified into many regions and the grayscale is the same for all pixels within a region. It can be seen that not only the match measure information but also the object information is clearly represented by the segmented regions.

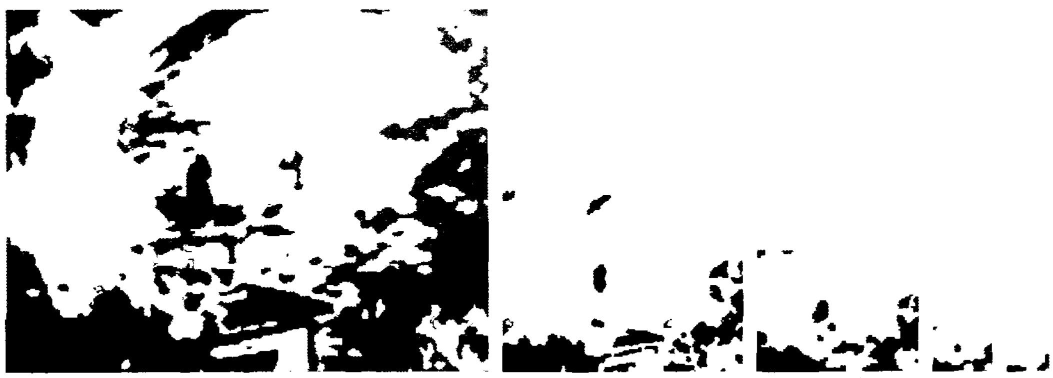


Figure. 5.16 MR segmentation of multi-level match measure images.

DT-CWT: This module performs MR decomposition of source images using dual-tree complex wavelet transform, which is the same as the DT-CWT module discussed in Section *5.1.1.*

Match measure II: This module computes the similarity between the corresponding regions on both sources. The computation is based on the MR segmentation of source images and the match measure results at pixel level obtained in module *match measure I.* In our fusion scheme, the match measure between corresponding regions at level j, $m_{j,R}^{AB}$, is defined as in:

$$m_{j,R}^{AB} = \frac{1}{S_{j,R}} \sum_{z \in R} m_j^{AB}(z) \tag{5.4}$$

where vector $z = (x, y)$ is the pixel coordinates and $S_{j,R}$ is the area of region R at level j, which is the number of pixels within the region. In fact, the similarity between corresponding regions defined in Eq. (5.4) is an averaged similarity at pixel level.

Activity measure: This module computes the activity measures of regions at all levels in both wavelet transformed source images. The computation is based on the activity measure at pixel-level and is defined in a similar way to Eq. (5.4):

$$E_{j,R}^{A} = \frac{1}{S_{j,R}} \sum_{z \in R} E_j^A(z), \quad E_{j,R}^{B} = \frac{1}{S_{j,R}} \sum_{z \in R} E_j^B(z) \tag{5.5}$$

where $E_{j,R}^{A}$ and $E_{j,R}^{B}$ are the activity measure of region R at level j in both source images;

$E_j^A(z)$ and $E_j^B(z)$ are pixel level activity measures at level j in both source images, which are defined Eq. (5.2). Thus the activity measure of a region can be considered as the arithmetic average of activity measures of all pixels within the region.

Combination: This module combines MR decomposition coefficients of both sources at region level to obtain the composite MR decomposition coefficients of the fused image. Similar to the coefficient combination at pixel level, the coefficient combination at region level is performed in a linear way as in:

$$F_j^k(z) = C_{j,R}^A \cdot A_j^k(z) + C_{j,R}^B \cdot B_j^k(z), \quad z \in R \tag{5.6}$$

where $A_j^k(z)$ and $B_j^k(z)$ are detail coefficients at pixel location z in region R at level j and orientation k while $F_{j,R}^k(z)$ is the composite coefficient at pixel location z in region R at level j and orientation k. $C_{j,R}^A$ and $C_{j,R}^B$ are weight factors for region R at level j in both sources. The Eq. 5.6 indicates that all pixels within a region will be assigned the same weight factor which is given in the *decision map*.

As for approximation coefficients, the combination is much simpler and identical to the combination methods discussed in pixel-level based fusion scheme. The approximation coefficients can be obtained directly from those of either source image, or by averaging those of both source images, or by choosing the larger coefficient at each pixel location between both sources. However, to use which method is determined by the properties of both source images and the visual quality of fused images produced by those methods.

Decision map: This module determines the values of the weight factors in Eq. (5.6) based on the match and activity measures of regions. The weight factor computation method used in this module is identical to what is proposed in the decision map module of our pixel based fusion scheme except that the pixel-level match and activity measures need to be replaced by region-level ones. So we rewrite those equations as in:

$$\text{Case 1:} \begin{cases} C_j^A = \dfrac{E_j^A}{E_j^A + E_j^B}, \\ C_j^B = \dfrac{E_j^B}{E_j^A + E_j^B}, \end{cases} \quad \textit{if} \quad m_j^{AB} \geq 0.9$$

$$\text{Case 2: } \begin{cases} C_j^A = \dfrac{E_j^A}{E_j^A + E_j^B} \cdot T + (1-T) \cdot W_A, \\ C_j^B = \dfrac{E_j^B}{E_j^A + E_j^B} \cdot T + (1-T) \cdot W_B, \end{cases} \quad if \quad 0.7 < m_j^{AB} < 0.9$$

$$where \quad T = \frac{m_j^{AB} - 0.7}{0.9 - 0.7}, \quad and \quad \begin{cases} W_A = 1 \quad if \quad E_j^A \geq E_j^B, \quad otherwise \quad 0 \\ W_B = 1 \quad if \quad E_j^A < E_j^B, \quad otherwise \quad 0 \end{cases}$$

$$\text{Case 3: } \begin{cases} C_j^A = 1 \quad C_j^B = 0, \quad if \quad m_j^{AB} \leq 0.7 \quad and \quad E_j^A \geq E_j^B \\ C_j^A = 0 \quad C_j^B = 1, \quad if \quad m_j^{AB} \leq 0.7 \quad and \quad E_j^A < E_j^B \end{cases}$$

Examples of the region level decision maps for the level 1 subsampled source images (Fig. 5.1) are shown in Fig. 5.17. The over all grayscale pattern is similar to the pixel-level decision maps shown in Fig. 5.6, and still more image information from IR image is merged in the fused image than the visual image. The difference from the pixel-level decision maps is that there are almost no individual spots on both region-level decision maps. Compared with the source images, it can be observed that the pixels in one region have similar grayscale values and are spatially related. In addition, each region can also represent a whole object, background or a part of either one.

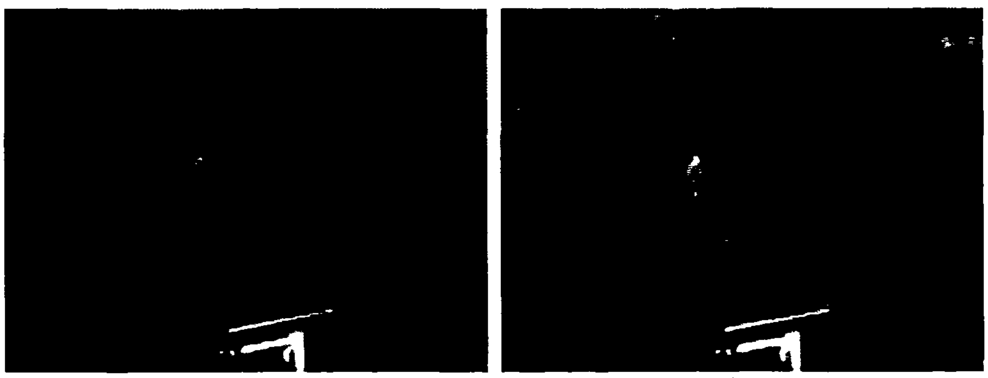
(a) (b)

Figure 5.17 Region-level decision maps for level 1subsampled source images: (a) visible image, and (b) IR image.

Inverse DT-CWT: This module performs inverse wavelet transform on the composite wavelet coefficients to reconstruct the fused image. Following this step, additional image processing steps may be needed to achieve color restoration or image enhancement, which have been discussed in Section *5.1.1*.

5.1.2 Results and Discussion

One example of the fused image, produced by the proposed region-based fusion algorithm, is presented in Fig. 5.18 accompanied by a fused image produced by the pixel-based fusion scheme discussed in Section 5.1. Both fused images are obtained with averaged approximation coefficients from both sources and they are almost identical. The halo effect the person is little less pronounced in the image produced by region based fusion. Fig.5.18(a) is also very similar to those yielded by other region based methods.



(a) (b)

Figure 5.18. Image fusion results produced by: (a) region based scheme and (b) pixel based scheme.

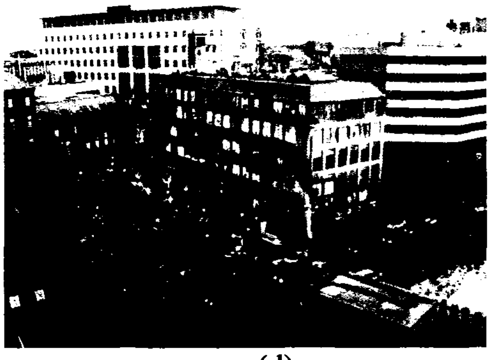
Another image fusion example is shown in Fig. 5.19. The two source images are displayed in Fig. 5.12. The images in the top row are image segmentation results for level 1 subsampled source images while the images in the middle row are decision maps for both source images. The images in bottom row are two fused images: one produced by region based algorithm and the other by pixel based algorithm. Although the source images have many complex patterns and are captured under complex lighting conditions, the quality of the fused image is still equivalent to pixel based algorithm with even less artifacts created

The evaluation of the performance of fusion algorithms is mainly subjective in many applications, and it is difficult to quantitatively assess a fused image because of the lack of the 'ideal' fused image. In the current literatures, quantitative evaluation of fusion algorithms is still an open problem. Despite of the difficulty, a few pixel-level objective methods have been proposed to measure the quality of the fused images. For instance,

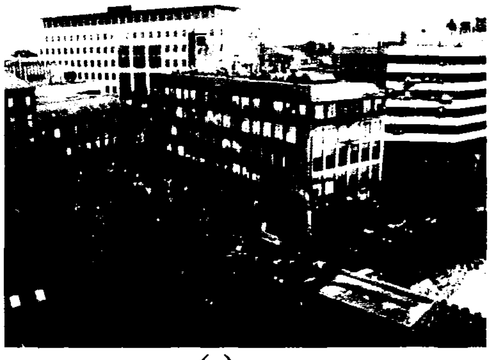
(a)

(b)

(c)



(d)



(e)

Figure 5.19. Image fusion results produced by: (a) segmentation result on match measure image; (b) decision map of visible image; (c) decision map of thermal image; (d) fused image by region based scheme; and (d) fused image by pixel based scheme.

mean squared error based metrics are commonly used for the comparison between the fused image and the sources. One example of those metrics is the root mean square error (RMSE) computed as:

$$RMSE = \left[\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(I_R(m,n) - I_F(m,n) \right)^2 \right]^{1/2} \tag{5.7}$$

where I_R and I_F are the reference image (ideal fused image) and the computed fused image, respectively, and M and N are the dimensions of the images. Due to the ambiguous concept of 'ideal' fused image and the difficulty of obtaining it, we propose an image fusion evaluation method based on a variant of Eq. (5.7) which can be expressed as:

$$RMSE_i = \left[\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(I_S^i(m,n) - I_F(m,n) \right)^2 \right]^{1/2} \tag{5.8}$$

where i is the index of source images, and I_S^i is the i^{th} source image. Obviously, Eq. (5.8) measures the difference between the fused image and each source. Then we use the arithmetic average of all $RMSE_i$:

$$AVG_{RMSE} = \frac{1}{S} \sum_{i}^{S} RMSE_i, \quad S = number \ \ of\, source \ \ images \tag{5.9}$$

as the quality measure of the fused image and a lower AVG_{RMSE} indicates a better fused image.

Information theory based metrics such as mutual information has also been proposed for fusion evaluation [132]. Given two images: the fused image I_F and the reference image I_R, their mutual information MI is defined as:

$$MI(I_R, I_F) = \sum_{a=1}^{L} \sum_{b=1}^{L} h_{R,F}(a,b) \log_2 \frac{h_{R,F}(a,b)}{h_R(a) h_F(b)} \tag{5.10}$$

where h_R and h_F are the normalized gray level histograms of I_R and I_F, respectively; $h_{R,F}$ is the joint gray level histogram of I_R and I_F; and L is the number of bins in the histograms. Based on Eq. (5.10), Qu et al. [133] proposed a non-reference objective fusion performance metric by summation of all the mutual information between the fused image and each source image:

$$MI_{ALL} = MI(I_A, I_F) + MI(I_B, I_F) + \dots \ . \tag{5.11}$$

Therefore, a larger MI_{ALL} indicates that more image information have been transferred from source images to the fused image.

To evaluate the performance of our fusion algorithm, the fusion metrics proposed in Eq. (5.9) and the referred evaluation defined in Eq. (5.11) are applied to analyze the fused images shown in Figs. 5.7 and 5.18.

Table 5.1 Fusion evaluation between region-based and pixel-based algorithms.

Fusion Algorithms	AVG_{RMSE}	MI_{ALL}
Proposed Region-Based	9.5863	3.2409
Proposed Pixel-Based	9.7016	3.0274
Burt [92]	9.7877	2.8892

The results are presented in Table 5.1 with comparison to the classical algorithm (pixel-based maximum selection) by Burt in [92]. It can be observed that the region-based algorithm performs the best in terms of both evaluation criteria, which means that the region-based algorithm incorporates the most information from both source images.

CHAPTER VI

APPLICATIONS OF VISIBILITY ENHANCEMENT TECHNIQUES

In this chapter, we discuss the prototype application systems which have been developed based on the techniques proposed in this dissertation. These application systems include: PC based software package for image and video enhancement; image enhancement tool in mobile device; software package for wavelet transform based MR image fusion; and a prototype driver visibility improvement (DVI) system that is based on multi-sensor image fusion.

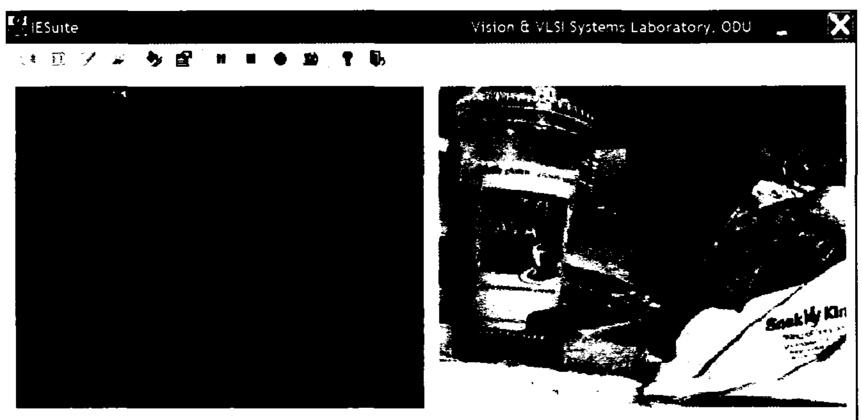
6.1 Software for Image and Video Enhancement

6.1.1 PC Based Software Package

A commercial software package "IESuite" has been developed based on the proposed nonlinear enhancement algorithms discussed in Chapter III for enhancing still images, recorded video, or real-time video streams from digital video camera. This software package is implemented in C++ in Windows XP environment. OpenCV libraries are used to provide Intel processor optimized program routines for basic mathematical operations while Microsoft DirectX libraries are used for digital video input and output. 2D spatial convolution is realized in frequency domain with FFTW libraries.

IESuite supports various image file formats, including JPEG, BMP, PNG, TIFF, and PBM. The original image and the enhanced result are displayed in two windows with best fit size simultaneously. The enhanced image can also be displayed in its original size by clicking the 'view' button. 'IESuite' includes 6 enhancement algorithms. It can either process single images or batch process and save large number of still images automatically without human interaction. A screen capture of IESuite processing a still image is shown in Fig. 6.1(a).

(a)



(b)

Figure 6.1 Screen captures of image enhancement software package IESuite: (a) still image enhancement; (b) video enhancement.

Real-time enhancement of a digital video stream has also been realized in IESuite using I-R model based algorithm on a desktop or notebook PC. A processing speed of 26 frames per second has been achieved on a PC with a 3.2GHz Intel Pentium 4 processor and 1GB 400 MHz DDR SDRAM memory. Sony DCR-HC85 digital video camera is used to capture the video stream with a frame size of 360 × 240 pixels. The digital video stream is transferred from the video camera into the computer via the IEEE 1394 port. Fig. 6.1(b) shows a screen capture of the interface of the video enhancement program. The video enhancement can be controlled with various playback commands including the 'snapshot' command which allows capturing of any individual enhanced frame. The enhanced digital video stream can be split and recorded in multi-clips at anytime by the user according to his/her interest, and all the clips finally are automatically saved into one

AVI video file with the selected encoder/compressor (e.g. Microsoft MPEG-4 video codec V2 is usually selected). In addition, the recording rate can be manually set. Besides the live video streams from the video camera, all the enhancement functions and controls are also applicable to enhancing the digital video streams from the video files stored in the computer.

The processing times needed for enhancing color images of different sizes are compared between the proposed AINDANE algorithm and MSRCR. The results are provided in Table 6.1. The computing platform is the same the one used for real time video enhancement. A commercial digital image processing software, PhotoFlair® version 2.0 (TruView Imaging Company), is used to implement MSRCR algorithm. Generally, the processing time of our algorithm is approximately 9% ~ 13% of that of MSRCR. One reason is because MSRCR processes all three spectral bands of color images while our algorithm only processes the intensity information, which is therefore suitable for fast processing.

Table 6.1 Comparison of processing time between AINDANE and MSRCR.

Image resolution (pixels)	AINDANE (seconds)	MSRCR (seconds)
360 × 240	0.10	1.21
640 × 480	0.42	4.02
1024 × 768	1.01	8.00
2000 × 1312	2.21	18.02

Finally, some image enhancement results produced by AINDANE implemented in IESuite are shown in Fig. 6.2 in which the original images are captured under various types of lighting and media conditions. The original image in the top row shows a very low illumination across the entire scene area while the original image in middle row exhibits a non-uniform illumination distribution over the entire scene. The original image in the bottom row is degraded by the turbulent light media. The visibility of all three images has been largely improved by the image enhancement.

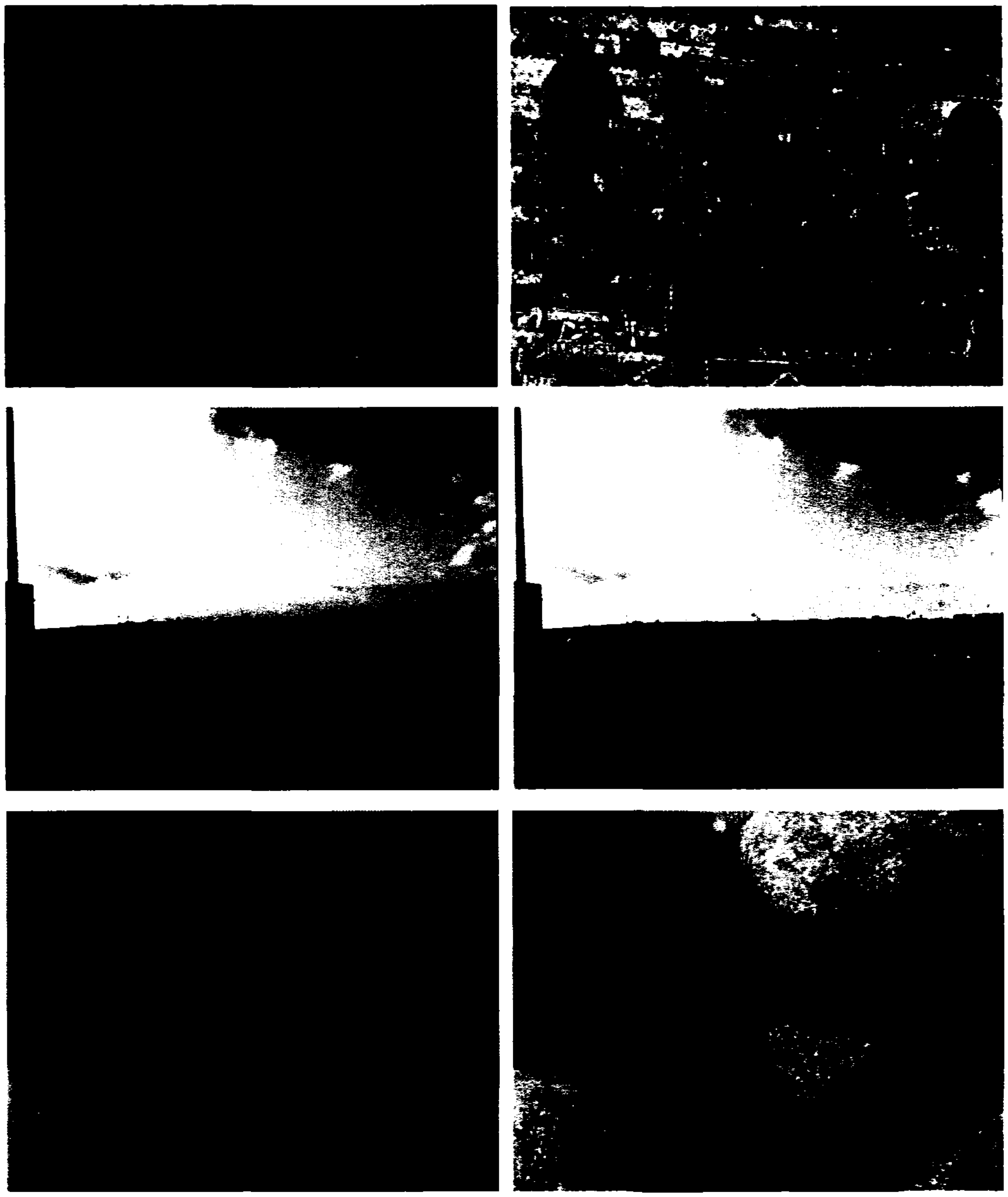


Figure 6.2 Image enhancement results of some sample images.

6.1.2 Embedded Enhancement Application on Pocket PC

A software package for still image enhancement has also been implemented on a pocket PC. Fig. 6.3 shows two pictures of a HP iPAQ H5555 pocket PC displaying an image before and after enhancement. The pocket PC running in Windows CE environment is equipped with a 400MHz Intel XScale processor and 128 Mbytes memory. The processing time of a 320×240 24-bit color image is about 8-10 seconds. The Windows bitmap and JPEG file format are supported. The images can be either captured by the built-in/add-on camera or imported from another source, e.g., an email or a PC that connects to the pocket PC. Due to the limited computing ability of the processor, the image size that is larger than a predefined threshold is scaled down to a preset size prior to the enhancement and restored to the original size after enhancement. The coding and compiling of the algorithm programs as well as the pocket PC simulation are realized with the eMbedded Visual C++ 4.0

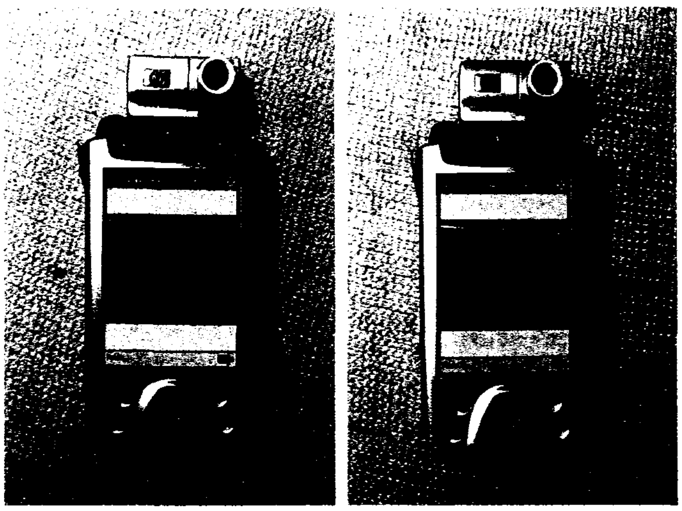


Figure 6.3 Image enhancement application on pocket PC with an add-on digital image camera: original image shown on the screen (left) and enhanced image shown on the screen (right).

6.1.3 Image Enhancement for the Improvement of Face Detection

The enhancement of the visual quality of digital images is usually applied to improve the performance of computer vision algorithms. Inspired by this relation, the proposed image enhancement technique is tested as an image preprocessor for face detection. Fig. 6.4 illustrates the effectiveness of the enhancement technique by applying Viola & Jones's face detection algorithm [123] on those images before and after enhancement. We chose the 8 sample images from FRGC databases [108] which are captured under complex lighting conditions.

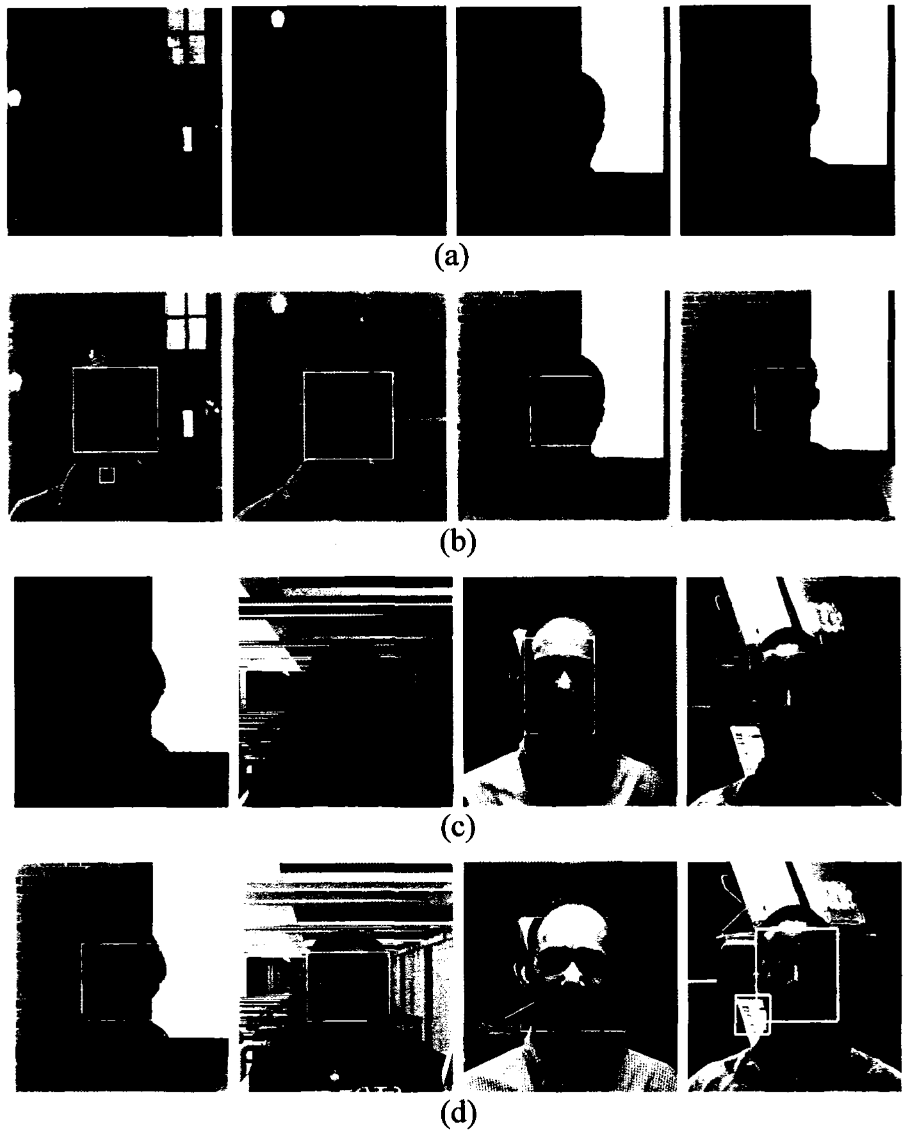


Figure 6.4 Face detection improved by image enhancement: (a) and (c): Face detection on FRGC sample images captured under complex lighting conditions; (b) and (d): Face detection on the same set of images after enhancement by AINDANE.

After applying face detection without enhancement, there are 6 failures in face detection (out of 8) of this image set. However after enhancement on those original images, it is possible to detect all the faces in the enhanced image set. In order to provide a statistical report of the face detection improved by image enhancement, we choose 2156 'difficult' face images (e.g., dark face with a bright background or a low luminance across the entire image in which human faces are hard to detect.) to be processed by the proposed I-R model enhancement algorithm. Then face detection is performed on the original and the enhanced face images to evaluate the effect of the image enhancement

Table 6.2 Face detection results of 2156 'difficult' FRGC face images before and after image enhancement by AINDANE, HE, Retinex and MSRCR

Image enhancement Methods	True positive	False positive
No image enhancement	1284	302
HE	1677	189
Retinex	1756	153
MSRCR	1787	162
AINDANE	2049	52

technique. In addition, histogram equalization (HE), the original Retinex, and MSRCR are also applied to process those 2156 images for face detection. The face detection results of those experiments are provided in Table 6.2 and the corresponding ROC curves are shown in Fig. 6.5. Both the proposed enhancement algorithm and the referred techniques can make improvement to the face detection rate because they all can increase

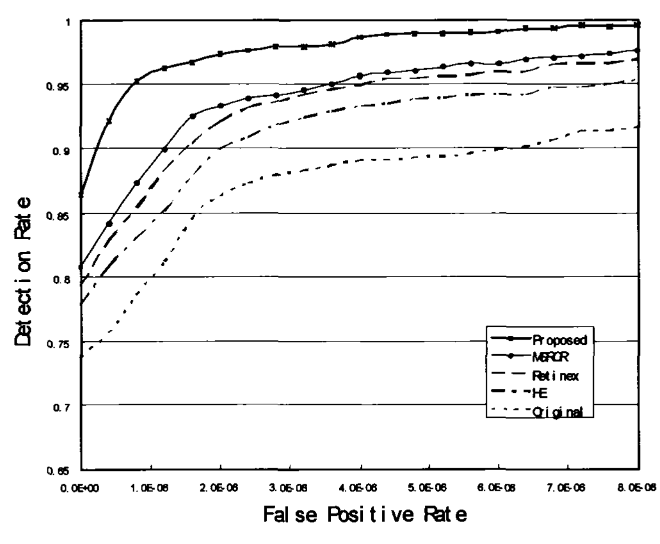


Figure 6.5 Overall ROC results for face detection of the 2156 FRGC face images with various preprocessing: no pre-processing (original), enhanced by HE, enhanced by Retinex, enhanced by MSRCR, and enhanced by AINDANE.

the visual quality of those face images to some extent. However, it is also found that the proposed algorithm performs considerably better than the other methods. This is because the proposed algorithm is able to adaptively process the images with well- balanced luminance enhancement and contrast enhancement to properly bring out all important image features which then result in improved performance of the face detection algorithm that is essentially image feature based method.

The detection rate is defined as the ratio between the number of successful detections and the number of faces. The false positive rate is the ratio between the number of false positive detections and the number of scanned windows. The ROC curves presented in Fig. 6.5 exhibit the relationship of detection rate versus false positive rate. The variation of the detection rate and false positive rate is caused by the change of certain threshold that determines whether a scan window is a human face or not. As the threshold decreases, both the false positive rate and detection rate will increase monotonously.

Since the tested face detection algorithm is able to conduct real time face detection at a high speed (about 50 frames per second for a frame size of 320×240), a prototype software has been developed to combine real time image enhancement and real time face detection together for improved face detection performance. The preliminary result is encouraging, and it is promising for real applications after some further improvement is made.

6.2 Image Fusion Based Visibility Improvement for Images and Videos

Two prototype application systems have been developed based on the proposed image fusion algorithms, which include an image fusion software package for image and video fusion as well as a prototype DVI system that combines video images from a CCD camera and a long wavelength infrared (LWIR) camera to provide the driver with a better view of the road condition at night.

6.2.1 Image Fusion Software Package

A prototype image fusion software package, named 'Image Fusion', has been implemented on a PC platform using C++ programming. The software, based on our pixel-level image fusion algorithm, is able to fuse either still image pairs or video frame pairs (the realization of real time video fusion is still in progress). A screen capture of the software interface is presented in Fig. 6.6 where the two input images are displayed in the top row while the fused image is shown in the bottom row. The input images can be either RGB color images or monochromatic images. The package can be applied to conduct image fusion of multi-sensor, multi-focus, multi-exposure, and multi-illuminance images.

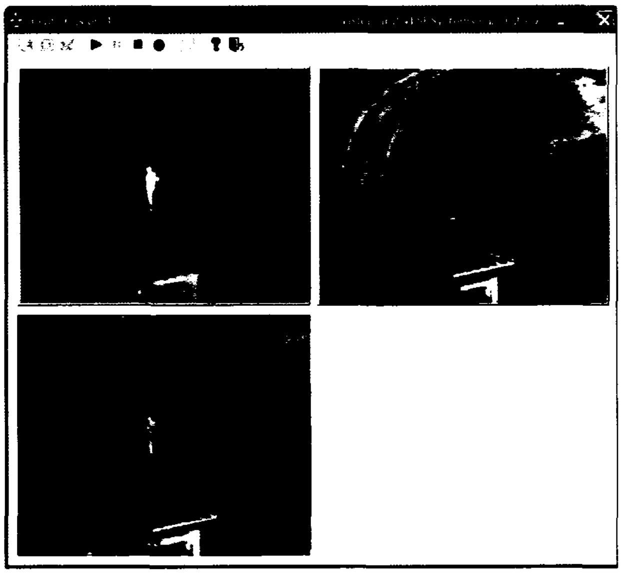


Figure 6.6 Interface of the software package for image fusion.

Two image fusion results produced by the software package are presented in Fig. 6.7 where the source images are shown in the top and middle rows and the fused image is provided in the bottom row. The left side images demonstrate the fusion of multi-sensor images. Red circles are marked in the left side images to point out the two persons who are missing in the visible image while appearing in the thermal image. The right side images are showing the fusion of multi-focus images. In the bottom image, all image features become clear and sharp while the farther objects look blurred in the top image and the near objects look unclear in the middle image.

6.2.2 Driver Visibility Improvement System Based on Image Fusion

Driver visibility improvement (DVI) system is an effective measure for improving road transportation safety. Driver's visibility can be severely weakened due to poor lighting conditions (e.g., night driving and bad weather) as well as impaired human vision caused

Figure 6.7 Left column is the fusion of multi-sensor images: CCD (visible) image (top); LWIR (thermal) image (middle); fused image (bottom). Right column is the fusion of multi-focus images: short focal length image (top); long focal length image (middle); fused image (bottom).

by aging or illness. A DVI system augments the driver's ability to see objects in the vehicle path by using on-board imaging sensors and digital image processing unit to create road images that are displayed to the driver on a head-up display (HUD) or a LCD screen. The studies of road traffic safety [125-127] indicate that a significant portion of the road accidents are caused by low visibility of the road condition due to poor lighting at night and aging-induced impaired vision. Therefore, researchers have been trying to develop appropriate vision enhancement systems which can be incorporated in vehicles for driver use [128-131]. Commercial products are available for certain vehicles. Current DVI systems acquire night vision capability by using infrared cameras operating at either long wavelength infrared (LWIR) range (8-12 μm) or short wavelength infrared (SWIR) range (0.8-1.2 μm). In addition, a combination of SWIR and CMOS cameras for DVI system has been reported. LWIR camera detects LWIR radiation which is mainly dependent on objects' temperature while SWIR camera is sensitive to both visible and near infrared light so it produces appearances of objects similar to those captured by cameras operating in visible range. Compared to SWIR cameras, LWIR cameras have the advantage of a longer range of sensing objects and better contrast. Although IR cameras are able to capture images without illumination, the images contain no chromatic information and thus look like monochrome (or grayscale) images.

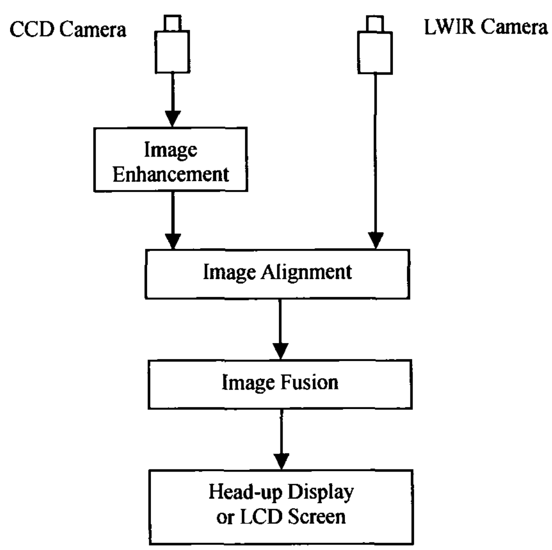


Figure 6.8 Structure of the proposed DVI system.

The prototype DVI system proposed in this dissertation is based on the image enhancement and image fusion of video sequences captured by CCD and LWIR cameras. The proposed system applies nonlinear neighborhood dependent image enhancement to improve the visibility of images captured by CCD camera, and then uses wavelet transform based multiresolution image fusion method proposed in this dissertation to combine the enhanced visible images and thermal images which are aligned before fusion. Fused images can provide more road scene information than what a human driver can perceive. The structure of the proposed DVI system is illustrated in Fig. 6.8.

In order to properly combine the visible and thermal images, image alignment is realized in two steps to achieve high-precision alignment of the input image pairs. The first step is camera alignment which conducts the alignment between the CCD and thermal cameras. Mechanical alignment between the two cameras is first performed, but it is not enough for this objective. Further camera-to-camera alignment is performed with both CCD and thermal video cameras by synchronizing, focusing and imaging on the same object at various distances to the lenses of the cameras. Images captured in these experiments are analyzed so that the orientation of both cameras can be further adjusted and aligned to obtain optimal overlapping of the fields of view (FOV) of both cameras. More importantly, the second step, image registration, is conducted manually to obtain image transformation and cropping schemes which are used for image alignment. The details of the image registration step are presented in the following paragraph.

Video frames captured by the CCD camera of Sony DCR-HC85 have a frame size of 720×480 while video frames captured by the LWIR camera of Thermal-Eye 250D (8-12μm) have a frame size of 640×480. The CCD camera has a larger FOV than the IR camera, thus the visible images must be transformed and cropped to align with thermal images. In the manual image registration experiment, 4 control point pairs are manually selected in each pair of visible and thermal image, and a 'projective' image transformation is assumed. In the 'projective' transformation of an image, besides translation, rotation and scaling, the scene appears tilted, which means that straight lines remain straight, but parallel lines converge toward vanishing points that might or might not fall within the image. In the manual image registration experiments, the control point pairs are selected from certain pixel locations which are uniformly-distributed across the image and are generally the corner pixels of the salient features or objects in the images. The image registration results indicate that the obtained 'projective' transformation parameters are able to provide reliable and consistent image alignment with high precision.

The objective of adopting image fusion in the system is to integrate scene information obtained from both cameras. With insufficient illumination, important road scene information may be severely degraded or even lost in the images captured by the CCD camera. Moreover, the information may still appear corrupted or lost even after image enhancement. However, the thermal images captured by LWIR camera may contain the information which is lost in the visible images. On the other hand, visible images look more natural to human viewers for they may contain objects' color information and generally have more details and higher contrast than thermal images. Therefore, the fusion of both types of images is able to provide the driver with a more complete knowledge of the road condition.

The DT-CWT based pixel-level MR image fusion method proposed in this dissertation is used to implement the fusion of CCD and thermal images. Compared to region based image fusion technique, which makes a multiresolution segmentation based on all source images and uses this segmentation to guide the fusion process, pixel based image fusion is able to provide more complete scene information on fused images while region based method may create loss of image features. In addition, pixel based method is also more robust and easier to implement than region based method.

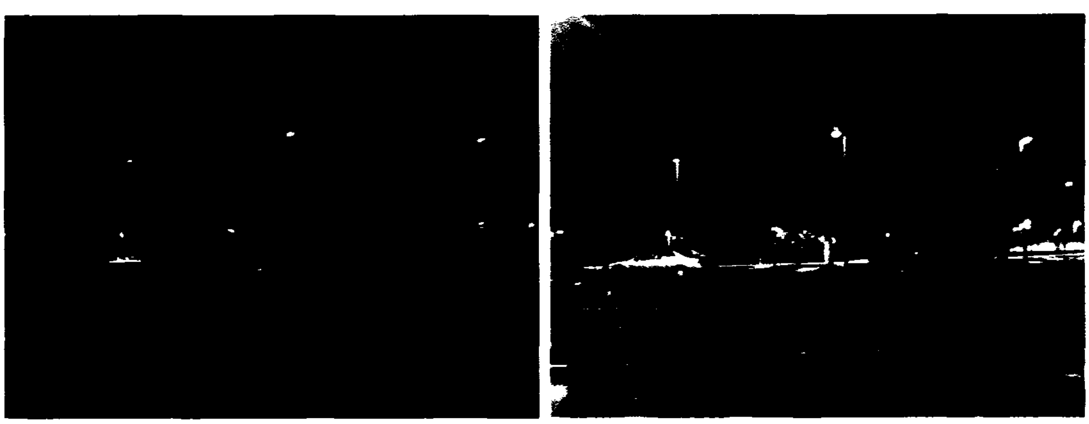


Figure 6.9 Nonlinear image enhancement of visible image. Left: original image; right: enhanced image

Some experimental results of the proposed DVI system are presented in the following part of this subsection. Fig. 6.9 illustrates the effectiveness of the nonlinear

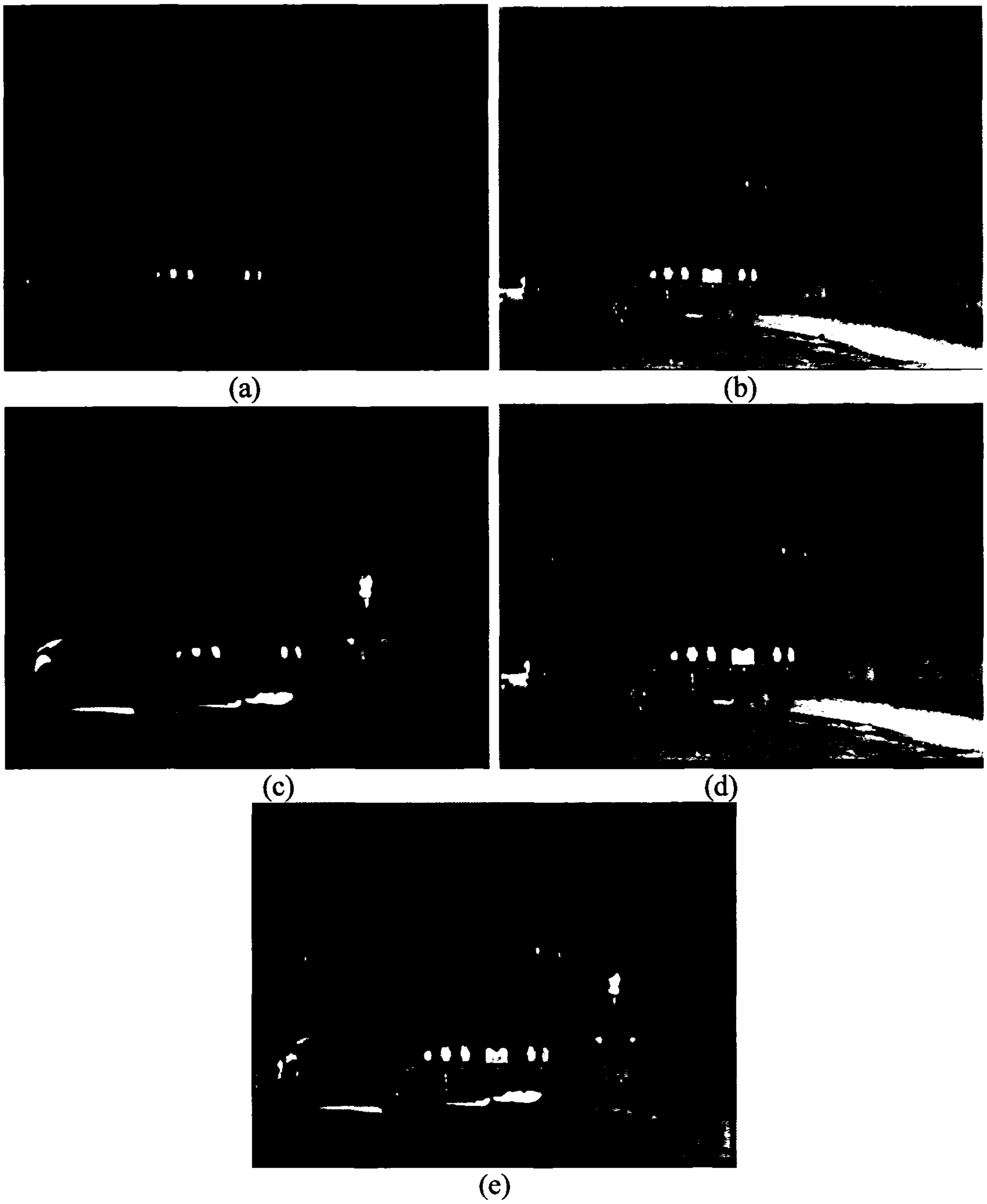


Figure 6.10 Results of image enhancement, image alignment and image fusion: (a) original visible image; (b) enhanced visible image; (c) original thermal image; (d) 'projective' transformed visible image aligned with thermal image; (e) fused image with color restoration.

image enhancement algorithm for improving visibility of the road scene image recorded by CCD camera. The original images exhibit dominant dark regions created by low illumination and limited dynamic range effect. However, the enhanced image exhibits a much more visible and clear view which is similar to what human viewers perceive. The enhanced images also illustrate the effectiveness of the contrast enhancement process for producing sharp images with compressed dynamic range.

The image alignment and image fusion results are presented in Fig. 6.10. Since the visible image has a larger FOV than that of thermal image, the visible image is transformed and cropped to align with thermal image. The red dots in the visible image represent the control points which can be chosen for image registration. The corresponding control points in the thermal image are marked as green dots. The result shows good image alignment between Figure 6.9(c) and 6.9(d). It can be observed that the fused image covers all important scene information from both visible and thermal images and looks similar to natural images with color information obtained from the visible image. Moreover, the fused image does not exhibit any severe image distortion or artifacts, and the noise level is modest.

The results from the prototype system are encouraging. However, the algorithms are implemented on a PC platform, which is not suitable for the real world application. The hardware design and implementation of the proposed enhancement and fusion algorithms are currently in progress in our research lab. The entire DVI system when completed would help the drivers to clearly see the objects on the road in order to have enough time to respond to these obstacles. The preliminary results demonstrate the capability of the proposed system which is promising for applications in reality.

CHAPTER VII

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the major contributions of the image enhancement techniques developed in this dissertation and possible future work is suggested for further improvement of the proposed algorithms or development of new ones.

7.1 Conclusions

In this dissertation we have investigated three types of image enhancement techniques for visibility improvement of digital images. The first type of enhancement technique is based on the spatial domain image processing methods. The second type is wavelet transform based techniques. The last type uses image fusion based techniques for visibility improvement.

In spatial domain techniques, I-R model based algorithm and AINDANE were presented. Both algorithms employ image dependent nonlinear intensity transfer function for adaptive and effective dynamic range compression. The proposed neighborhood dependent local contrast enhancement technique is successfully used in a single-scale or multi-scale manner to enhance image features of various scales. The linear color restoration process is able to produce reliable results with good color consistency for color image enhancement. In our experiments, both algorithms yield high quality enhanced images while AINDANE performs better than the I-R model based algorithm because of its multi-scale processing. However, since the high frequency components (reflectance) are not included in the dynamic range compression, only a single-scale contrast enhancement is used in the I-R based algorithm which is still able to produce good enhancement results as well as fast computation speed. Our algorithms compare favorably with other typical spatial domain techniques in terms of algorithm capability, flexibility, robustness, noise level, and balance between luminance and contrast enhancement.

Wavelet transform based image denoising and contrast enhancement algorithms are developed based on the modification of the wavelet coefficients. In image denoising, a bivariate pdf model is introduced to explore the interlevel dependency among the wavelet coefficients. In addition, an approximate solution to the MAP estimation problem is proposed to avoid complex iterative computations to find a numerical solution. The image denoising algorithm implemented with DTCWT produces high quality denoised images equivalent to the results of the best algorithms in literature which however are much more complex than our proposed method. New wavelet transform based contrast enhancement technique was developed based on the systematic studies of the correlation between the image quality change and wavelet coefficient change. Appropriate nonlinear and linear transfer functions were designed for the modification of wavelet coefficients. Furthermore, dynamic range compression of the approximation images is proposed to provide luminance enhancement to the wavelet based enhancement technique. The performance of the proposed technique is similar to that of the curvelet transform based method which is claimed to be the most suitable method for contrast enhancement conducted in transformed domain. However, curvelet transform is much slower than wavelet transform.

Both pixel-based and region-based image fusion schemes were studied for the purpose of image visibility improvement. New wavelet transform based MR fusion schemes were proposed using some new techniques developed in this dissertation: (1) match measure obtained in spatial domain instead of in transform domain is applied to guide the fusion process; (2) three cases are considered when combining corresponding pixels or regions for fusing as much information as possible; (3) MR segmentation was conducted on match measure images obtained in the spatial domain. The purpose of developing those new techniques was to accurately evaluate the similarity between the corresponding pixels and regions, which is essential for determining the weight factor of each pixel or region in the linear combination process to yield the fused image. Compared to other fusion schemes, the proposed fusion techniques have higher processing speed and produce equivalent or slightly improved fusion results.

Prototype application systems were also successfully developed using the visibility improvement algorithms proposed in the dissertation. These application systems include: a Windows software package based on spatial domain algorithms for still image and real time video enhancement; a Windows software package for still image enhancement on a mobile device, like a pocket PC or PDA; a Windows software package of wavelet transform based pixel-level image fusion for fusion of still image pairs or video file pairs;

a PC based conceptual DVI system for night or bad weather driving using image fusion of visible and thermal images. Those prototype systems exhibit encouraging results and are promising for real world applications.

7.2 Future Work

For the purpose of improving the currently proposed algorithms as well as developing new image enhancement algorithms, some meaningful work or suggestions are introduced for future investigation.

(1) The core issues of enhancing images captured in high dynamic range scenes are dynamic range compression of scene illuminance and enhancement of scene reflectance. From another point of view, they are problems of the dynamic range compression of low frequency components and the enhancement of high frequency components. Therefore, it is desirable to develop an algorithm which can make an optimal balance between those two problems and provide high and robust adaptivity for automatic enhancement or provide a highly flexible control over the enhancement process so that the user can tune the enhancement process to obtain optimized results to his/her interest.

(2) Both interlevel and intralevel dependency among wavelet coefficients need to be considered simultaneously in the coefficient thresholding to better prevent the formation of image artifacts induced by improper coefficient thresholding. In other words, neighborhood dependent coefficient thresholding needs to be developed. For example, the bivariate pdf used in our proposed denoising algorithm is an appropriate model to describe the interband dependency of wavelet coefficients. However, it is not neighborhood dependent. The neighborhood dependency could be added to the shrinkage function by estimating the noise content at each pixel location with the information obtained in the neighborhood. In addition, we also suggest introducing the interlevel and intralevel dependency into the wavelet transform based image enhancement algorithms for better quality enhanced images.

(3) To improve the performance of the proposed region-based image fusion algorithm, we suggest that more advanced MR segmentation techniques need to be incorporated and

tested in our algorithm, such as the segmentation techniques based on texture analysis, gradient, watershed method and statistical method.

REFERENCES

[1] E. Land and J. McCann, "Lightness and Retinex theory," *Journal of the Optical Society of America*, vol. 61, pp.1-11,1971.

[2] E. Land, "Recent advances in Retinex theory and some implications for cortical computations," *Proc. Nat. Acad. Sci.*, vol. 80, pp. 5163-5169, 1983.

[3] E. Land, "Recent advances in Retinex theory," *Vision Research*, vol. 26, pp. 7-21, 1986.

[4] J. McCann, S. McKee, and T. Taylor, "Quantitative studies in Retinex theory, a comparison between theoretical predictions and observer responses to the color mondrian experiments," *Vision Research,* vol. 16, pp. 445-458, 1976.

[5] J. Frankle, and J. McCann, "Method and apparatus for lightness imaging," *US Patent #4,384,336,* 1983.

[6] E. Land, "An alternative technique for the computation of the designator in the retinex theory of color vision," *Proc. of the National Academy of Science USA,* vol. 83, pp. 2078-3080, 1986.

[7] J. McCann, "Lessons learned from mondrians applied to real images and color gamuts," *Proc. IS&T/SID Seventh Color Imaging Conference,* pp. 1-8, 1999.

[8] R. Sobol, "Improving the Retinex algorithm for rendering wide dynamic range photographs," *Proc. SPIE* 4662, pp. 341–348, 2002.

[9] A. Rizzi, C. Gatta, and D. Marini, "From Retinex to ACE: Issues in developing a new algorithm for unsupervised color equalization," *Journal of Electronic Imaging*, vol. 13, pp. 75-84, 2004.

[10] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround Retinex," *IEEE Transaction on Image Processing,* vol. 6, pp. 451-462, 1997.

[11] Z. Rahman, D. Jobson, and G. Woodell, "Multiscale Retinex for color image enhancement," *Proceedings of the IEEE International Conference on Image Processing,* Lausanne, Switzerland, vol. 3, pp. 1003-1006, 1996.

[12] Z. Rahman, D. Jobson, and G. Woodell, "Multiscale Retinex for color rendition and

dynamic range compression," *Applications of Digital Image Processing XIX,* Denver, Colorado, pp. 9-17, 1996.

[13] D. Jobson, Z. Rahman, and G. Woodell, "A multi-scale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing: Special Issue on Color Processing*, vol. 6, pp. 965-976, 1997.

[14] Z. Rahman, G. Woodell, and D. Jobson, "A comparison of the multiscale Retinex with other image enhancement techniques," *Proceedings of the IS&T 50th Anniversary Conference*, IS&T, pp. 426-431, 1997.

[15] T. Watanabe, Y. Kuwahara, A. Kojima, and T. Kurosawa, "Improvement of color quality with modified linear multi-scale Retinex," *Proceedings of the 15th SPIE Symposium on Electronic Imaging*, Santa Clara, CA, pp. 59-69, 2003.

[16] K. Barnard and B. Funt, "Analysis and improvement of multi-scale Retinex," *IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, pp. 221-226, 1997.

[17] K. Barnard, G. Finlayson, and B. Funt, "Color constancy for scenes with varying illumination," *Proceedings of the 4th European Conference on Computer Vision,* pp. II: 1-15, 1996.

[18] L. Tao and V. K. Asari, "Modified luminance based MSRCR for fast and efficient Image enhancement," *IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2003, Washington DC, USA,* pp.174-179, 2003.

[19] S. M. Pizer, J. B. Zimmerman, and E. Staab, "Adaptive grey level assignment in CT scan display," *Journal of Computer Assistant Tomography*, vol. 8, pp. 300-305, 1984.

[20] J. B. Zimmerman, S. B. Cousins, K. M. Hartzell, M. E. Frisse, and M. G. Kahn, "A psychophysical comparison of two methods for adaptive histogram equalization," *Journal of Digital Imaging*, vol. 2, pp.82-91, 1989.

[21] S. M. Pizer and E. P. Amburn, "Adaptive histogram equalization and its variations," *Computer Vision, Grpahics, and Image Processing*, vol. 39, pp. 355-368, 1987.

[22] K. Rehm and W. J. Dallas, "Artifact suppression in digital chest radiographs enhanced with adaptive histogram equalization," *SPIE Conference on Medical Imaging III*, 1090, pp. 290-300, 1989.

[23] Y. Jin, L. M. Fayad, and A. F. Laine, "Contrast enhancement by multiscale adaptive

histogram equalization," *Proc. SPIE*, vol. 4478, pp. 206-213, 2001.

[24] G. W. Larson, H. Reshmeier and C. Piatko, "Visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, pp. 291-306, 1997.

[25] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang and K. Zimmerman, "Spatially non-uniform scaling functions for high contrast images," *Proceedings of Graphics Interface*, pp. 182-191, 1993.

[26] C. Schlick, "Quantization techniques for visualization of high dynamic range pictures," *Photorealistic Rendering Techniques, Poceedings of the 5th Eurographics Rendering Workshop*, pp.7-20, 1994.

[27] J. Tumblin and G. Turk, "LCIS: A boundary hierarchy for detail-preserving contrast reduction", *SIGGRAPH 99 Conference Proceedings, Computer Graphics Annual Conference Series*, pp.83-90, 1999.

[28] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild and D. P. Greenberg, "A multiscale model of adaptation and spatial vision for realistic image display", *SIGGRAPH 98 Conference Proceedings, Computer Graphics Annual Conference Series*, pp. 287-298, 1998.

[29] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images", *Proc. of SIGGRAPH2002*, pp. 267-277, 2002.

[30] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Recog. and Mach. Intellig.*, vol. 11, pp. 674-693, 1989.

[31] J. Woods and S. O'Neil, "Subband coding of images," *IEEE Trans. Acous. Speech Signal Proc.*, vol. 35, pp. 1278-1288, 1986.

[32] P. Vaidyanathan and P. Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction filter banks," *IEEE Trans. Acous. Speech Signal Proc.*, vol. 36, pp. 81-94, 1988.

[33] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Comm.*, vol. 31, pp. 532-540, 1983.

[34] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukhah, and A. Willsky,

"Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 766-784, 1992.

[35] J. C. Brailean and Aggelos K. Katsaggelos, "Simultaneous recursive displacement estimation and restoration of noisy-blurred image sequences," *IEEE Trans. Image Processing*, vol. 4, pp. 1236-1251, 1995.

[36] S. G. Chang, Y. Bin, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. on Image Proc.*, vol. 9, pp. 1532-1546, 2000.

[37] J. M. Shapiro, "Embedded image coding using zero trees of wavelet coefficients," *IEEE Image Proc.*, vol. 41, pp. 3445-3462, 1993.

[38] K. Kinebuchi, D. Muresan, and T. Parks, "Image interpolation using wavelet-based hidden markov trees," in *IEEE Int. Conf. Acoust. Speech Signal Proc.*, USA, pp. 7-11, 2001.

[39] N. Lee, Q. Huynh, and S. Schwarz, "New methods of linear time-frequency analysis for signal detection," in *IEEE Int. Symp. Time-Freq. Time-Scale Anal.*, USA, pp. 26-29, 1996.

[40] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, pp. 910-916, 1992.

[41] D. Taubman and M. Marcellin, *JPEG2000: Image compression fundamentals*, Kluwer, USA, 2002.

[42] T. Ebrahimi and F. Pereira, *The MPEG book*, Prentice Hall, USA, 2002.

[43] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Recog. and Mach. Intellig.*, vol. 6, pp. 373-383, 1985.

[44] J. S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Recog. and Mach. Intellig.*, vol. 2, pp. 165-168, 1980.

[45] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Proc.*, vol. 12, pp. 1338-1351, 2003.

[46] M. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical

modeling of wavelet image coefficients and its application to denoising," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, USA, pp. 3253-3256, 1999.

[47] H. Chipman, E. Kolaczyk, and R. McCulloch, "Adaptive Bayesian wavelet shrinkage," *J. Amer. Stat. Assoc.*, vol. 92, pp. 1430-1442, 1997.

[48] F. Abramovich and T. Sapatinas, "Wavelet thresholding via a Bayesian approach," *Journal of Royal Stat. Soc., Ser. B*, vol. 60, pp. 725-749, 1998.

[49] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrica*, vol. 81, pp. 425-455, 1994.

[50] J. Lee, "Digital image smoothing and the sigma filter," *Computer Vision, Graphics and Image Processing*, vol. 24, pp. 255-269, 1983.

[51] J. L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. on Image Proc.*, vol. 11, pp. 670-684, 2002.

[52] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," *IEEE Trans. on Signal Proc.*, vol. 46, pp. 886-902, 1998.

[53] S. LoPresto, K. Ramchandran, and M. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework," in *Proc. Data Compression Conf.*, USA, pp. 321-230,, 1997.

[54] E. Simoncelli and E. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. IEEE Int. Conf. Image Proc.*, Switzerland, pp. 379-382, 1996.

[55] L. Zhang, P. Bao, and X. Wu, "Hybrid inter- and intra-wavelet scale image restoration," *Pattern Recognition*, vol. 36, pp. 1737-1746, 2003.

[56] M. Banham and A. Katsaggelos, "Spatially adaptive wavelet-based multiscale image restoration," *IEEE Trans. Image Proc.*, vol. 5, pp. 619-634, 1996.

[57] J. Liu and P. Moulin, "Image denoising based on scale-space mixture modeling for wavelet coefficients," in *Proc. IEEE Int. Conf. Image Proc.*, Japan, pp. 386-390, 1996.

[58] J. Scharcanski, C. Jung, and R. Clarke, "Adaptive image denoising using scale and space consistency," *IEEE Trans. Image Proc.*, vol. 11, pp. 1092-1101, 2002.

[59] M. Malfait and D. Roose, "Wavelet-based image denoising using a markov random field a prior model," *IEEE Trans. on Image Proc.*, vol. 6, pp. 545-557, 1997.

[60] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for bayesian wavelet based image denoising," *IEEE Trans. on Image Proc.*, vol. 11, pp. 724-735, 2002.

[61] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden markov models," *IEEE Trans. Image Proc.*, vol. 10, pp. 1056-1068, 2001.

[62] G. Fan and X. Xia, "Improved hidden markov models in the wavelet domain," *IEEE Trans. Signal Proc.*, vol. 49, pp. 115-120, 2001.

[63] M. Jansen and A. Bultheel, "Empirical bayes approach to improve wavelet thresholding for image noise reduction," *J. Amer. Statist. Assoc.*, vol. 96, pp. 629-639, 2001.

[64] F. Jin, P. Fieguth, and L. Winger, "Image denoising using complex wavelet and markov prior models," *Image Analysis and Recognition, Second International Conference, ICIAR 2005, Proc.*, Toronto, Canada, pp. 73-80, September, 2005.

[65] G. Fan and X. Xia, "Wavelet-based texture analysis and synthesis using hidden markov models," *IEEE Trans. Circuits and Systems I: Fundamental theory and applications*, vol. 50, pp. 106-120, 2003.

[66] N. G. Kingsbury, "The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters," *IEEE Digital Signal Processing Workshop, DSP 98*, Bryce Canyon, paper no 86, 1998.

[67] N. G. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," *Proc. European Signal Processing Conference, EUSIPCO 98*, Rhodes, pp 319-322, 1998.

[68] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Appl. and Comp. Harmon. Analy*," vol. 10, pp. 234-253, 2001.

[69] H. Choi, J. Romberg, R. Baraniuk, and N. G. Kingsbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proc. ICASSP 2000*, Istanbul, Turkey, vol. 1, pp. 133–136, 2000.

[70] P. Hill and D. Bull, and C. Canagarajah, "Rotationally invariant texture features

using the dual-tree complex wavelet transform," in *Proc. ICIP 2000*, Vancouver, Canada, vol. 3, pp. 901-904, 2000.

[71] P. F. C. de Rivaz and N. G. Kingsbury, "Fast segmentation using level set curves of complex wavelet surfaces," in *Proc. ICIP 2000*, Vancouver, Canada, vol. 3, pp. 592-595, 2000.

[72] R. Gonzalez and R. Woods, *Digital image processing*, Addison-Wesley, USA, 1992.

[73] A. Jain, *Fundamentals of digital Image Processing*, Prentice-Hall, USA, 1989.

[74] G. McLean and E. Jernigan, "Indicator functions for adaptive image processing," *Optical Engineering*, vol. 29, pp. 141-157, 1991.

[75] M. Sundareshan, T. Ji, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties," *IEEE Trans. Med. Imag.*, vol. 13, pp. 168-178, 1994.

[76] S. Thurnohofer and S. Mitra, "Edge-enganced image zooming," *Optical Engineering*, vol. 35, pp. 1862-1870, 1996.

[77] S. Mitra, H. Li, I. Lin, and T. Yu, "A new class of nonlinear filters for image enhancement," in *Proc. IEEE Int. Conf. Acoust. Speech, signal Proc.*, USA, pp. 2525-2528, 1991.

[78] H. Greenspan, C. Anderson, and S. Akber, "Image enhancement by nonlinear extrapolation in frequency space," *IEEE Trans. Image Processing*, vol. 9, pp. 1035-1047, 2000.

[79] P. Burt and R. Kolczybski, "Enhanced image capture through fusion," in *Proc. Int. Conf. Computer Vision*, Germany, pp. 173-182,, 1993.

[80] D. Woo, I. Eom, and Y. Kim, "Image interpolation based on inter-scale dependency in wavelet domain," in *ICIP 2004*, Singapore, October 2004, vol. 3, pp. 1687-1690, 2004.

[81] W. Carey, D. Chuang, Sheila, and S. Hemami, "Regularity-preserving image interpolation," *IEEE Trans. Image Processing*, vol. 8, pp. 1293-1297, 1999.

[82] K. V. Velde, "Multi-scale color image enhancement," in *Proc. Int. Conf. Image Processing*, vol. 3, pp. 584-587, 1999.

[83] J.-L. Starck, F. Murtagh, E. J. Candes, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Trans. Image Processing*, vol. 12, pp. 706-717, 2003.

[84] C. Pohl, J.L. Genderen, "Multisensor image fusion in remotesensing: concepts, methods and applications," *International Journal of Remote Sensing 19 (5)* pp. 823–854, 1998.

[85] R.C. Luo, M.G. Kay, "Multisensor integration and fusion for intelligent machines and systems," Ablex Publishing Corporation, 1995.

[86] P.J. Burt, R.J. Kolczynski, "Enhanced image capture through fusion," *Proceedings of the 4th International Conference on Computer Vision*, Berlin, Germany, pp. 173–182, 1993.

[87] M.M. Daniel, A.S. Willsky, "A multiresolution methodology for signal-level fusion and data assimilation with application to remote sensing," *Proceedings of the IEEE 85, 1)*, pp.164–180, 1997.

[88] B.V. Dasarathy, "Fuzzy evidential reasoning approach to target identity and state fusion in multisensor environments," *Optical Engineering 36 (3)*, pp. 683–699, 1997.

[89] M. Dubuisson, A.K. Jain, "Contour extraction of moving objects in complex outdoor scenes," *International Journal of Computer Vision 14*, pp. 83–105, 1997.

[90] B.V. Dasarathy, "Decision fusion", *IEEE Computer Society Press*, Los Alamitos, California, 1994.

[91] B. Jeon, D.A. Landgrebe, "Decision fusion approach for multitemporal classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp.1227–1233, 1999.

[92] P.J. Burt, "The pyramid as a structure for efficient computation, in: A. Rosenfeld (Ed.)," *Multiresolution Image Processing and Analysis*, Springer-Verlag, Berlin, Germany, pp. 6–35, 1984.

[93] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition 9* pp.245–253, 1989.

[94] G.K. Matsopoulos, S. Marshall, "Application of morphological pyramids: fusion of MR and CT phantoms," *Journal of Visual Communication and Image Representation*, vol. 6, pp. 196–207, 1995.

[95] A. Toet, "A morphological pyramidal image decomposition," *Pattern Recognition Letters*, vol. 9, pp. 255–261, 1989.

[96] T. Ranchin, L. Wald, "The wavelet transform for the analysis of remotely sensed images," *International Journal of Remote Sensing*, vol. 14, pp. 615–619, 1993.

[97] H. Li, B.S. Manjunath, S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, pp. 235–245, 1995.

[98] T.A. Wilson, S.K. Rogers, L.R. Meyers, "Perceptual based hyperspectral image fusion using multiresolution analysis," *Optical Engineering*, vol. 34, pp. 3154–3164, 1995.

[99] I. Koren, A. Laine, F. Taylor, "Image fusion using steerable dyadic wavelet transforms," *Proceedings of the IEEE International Conference on Image Processing*, Washington DC, pp. 232–235, 1995..

[100] Z. Liu, K. Tsukada, K. Hanasaki, Y.K. Ho, Y.P. Dai, "Image fusion by using steerable pyramid," *Pattern Recognition Letters*, vol. 22, pp. 929–939, 2001.

[101] O. Rockinger, "Pixel-level fusion of image sequences using wavelet frames," *Proceedings of the 16th Leeds Applied Shape Research Workshop*, Leeds University Press,, 1996.

[102] T. Pu, G. Ni, "Contrast-based image fusion using the discrete wavelet transform," *Optical Engineering*, vol. 39, pp. 2075–2082, 2000.

[103] S.T. Li, Y.N. Wang, "Multisensor image fusion using discrete multiwavelet transform," *Proceedings of the 3rd International Conference on Visual Computing*, Mexico City, Mexico,, 2000.

[104] P. Scheunders, Multiscale edge representation applied to image fusion, *Proceedings of Wavelet Applications in Signal and Image Processing VIII*, SPIE, San Diego, USA, pp. 894-901, 2000.

[105] S.G. Mallat, "A Wavelet Tour of Signal Processing," *Academic Press*, San Diego, California, 1998.

[106] L. Tao and V. K. Asari, "Modified luminance based MSRCR for fast and efficient image enhancement," *IEEE AIPR Proc IEEE,* pp. 174-179, 2003.

[107] Horn, B.K.P. "Determining Lightness from an image," *Computer Graphics and Image Processing,* vol. 3, pp. 277-299, 1974.

[108] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 947-954, 2005.

[109] B. Funt, F. Ciurea, and J. McCann, "Retinex in Matlab", *Proc. CIC'8 Eighth Color Imaging Conference, (Imaging Science \& Technology Society)*, Scottsdale, Arizona, pp. 112-121, 2000.

[110] L. Tao and V. K. Asari, "An integrated neighborhood dependent approach for nonlinear enhancement of color images," *Proc. IEEE Computer Society International Conference on Information Technology: Coding and Computing – ITCC 2004*, vol. 2, pp.138-139, 2004.

[111] L. Sendur, and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency,' *IEEE Tran. Sig. Proc.*, vol. 50, pp. 2744-2756, 2002.

[112] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, 1995.

[113] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[114] S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Processing*, vol. 9, pp. 1532–1546, 2000.

[115] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Processing*, vol. 9, pp. 1522–1531, 2000.

[116] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 46, pp. 886–902,, 1998.

[117] M. K. Mihcak, I.Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Lett.*, vol. 6, pp. 300–303, 1999.

[118] X. Li and M. T. Orchard, "Spatially adaptive image denoising under overcomplete

expansion," in *Proc. ICIP 2000*, Vancouver, Canada, vol. 3, pp. 300-303,, 2000.

[119] H. Choi, J. K. Romberg, R. G. Baraniuk, and N. G. Kingsbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proc. ICASSP*, Istanbul, Turkey, vol. 1, pp. 133–136, 2000.

[120] J. L. Starck, F. Murtagh, E. J. Candes, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform", *IEEE Trans. Image Proc.*, vol. 12, pp. 706-717, 2003.

[121] P.J. Burt, T.H. Hong, A. Rosenfeld, "Segmentation and estimation of image region properties through cooperative hierarchical computation," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, pp. 802–809, 1981.

[122] G. D. Hines, Z. Rahman, D. J. Jobson, G. A. Woodell, "DSP implementation of the multiscale retinex image enhancement algorithm," *Visual Information Processing XIII*, Proc. SPIE 5438, pp. 13-24, 2004.

[123] P Viola and M Jones, " Robust real-time face detection," *International Journal of Computer Vision,* vol. 57, pp. 137-154, 2004.

[124] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 947-954, 2005.

[125] "Driving at night," *National Safety Council*, Nov. 15, 2004. Available at: http://www.nsc.org/library/facts/nightdr.htm (accessed date: Feb, 2005)

[126] "Newspaper article: Nighttime visibility," *U.S. Department of Transportation Federal Highway Administration.* May 5, 2003. Available at: http://safety.fhwa.dot.gov/pedcampaign/press/02.htm (accessed date: Dec, 2004)

[127] "Bringing the nighttime road to life" *Road Management & Engineering Journal* Available at: http://www.usroads.com/journals/rmej/0109/rm010902.htm (accessed date: Nov, 2004)

[128] D. Koob, F. Bellotti, C. Bellotti and L. Andreone, "Enhanced driver's perception in poor visibility," Available at http://www.edel-eu.org/ (accessed date: Jan, 2005)

[129] J. C. Egelhaaf, and P. M. Knoll, "The 'night sensitive' vehicle, " Available at:

http://www.edel-eu.org/ (accessed date: Jan, 2005)

[130] W. Uhler, H. Mathony, and P. M. Knoll, "Driver assistance systems for safety and comfort," Available at http://www.edel-eu.org/ (accessed date: Oct, 2004)

[131] P. M. Knoll, "A NIR based system for night vision improvement," Available at http://www.edel-eu.org/ (accessed date: Feb, 2005)

[132] Z. Zhang, R. Blum, "A categorization of multiscale-decomposition based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, pp. 1315–1326, 1999.

[133] G.H. Qu, D.L. Zhang, and P.F. Yan, "Information measure for performance of image fusion," *Electronic Letters*, vol. 38, no. 7, pp. 313–315, 2002.

VITA

(By December, 2005)

Li Tao

Homepage: www.litao.info

EDUCATION

B.S. Electronics and Computer Engineering, Sichuan University, Chengdu, China, 1998
M.S. Electronics and Computer Engineering, Sichuan University, Chengdu, China, 2001
Ph.D Electrical and Computer Engineering, Old Dominion University, Fall, 2005

OFFICIAL PATENT FILED

Visibility Improvement in Color Video Stream
(*Patent fees paid by Vision Innovations Corporation. Technology licensed to Vision Innovations Corporation*)

RESEARCH PUBLICATIONS

Journal Papers:

1. **Li Tao** and K. Vijayan Asari, "An adaptive and integrated neighborhood dependent approach for nonlinear enhancement of color images," ***SPIE Journal of Electronic Imaging-JEI***, vol. 14, no. 4, pp. 1.1-1.14, October 2005.

2. Deepthi P. Valaparla, **Li Tao**, Ming-Jung Seow and K. Vijayan Asari, "Neural network model for image enhancement using histogram matching," ***Multimedia Cyberscape Journal: Special Issue on Multimedia Data Processing and Compression***, vol. 3, no. 4, pp. 19-24, November 2005.

3. **Li Tao** and Jiliu Zhou, "The shifing-bit rules of block cipher DES and the implementation of information encryption", ***Journal of Sichuan University (Natural Science Edition)***, China, vol.38, no.3, pp 350-354, 2001.

4. **Li Tao**, Jiliu Zhou and Xiaohai He, "Discussion of strategy for calculating the perimeter and area of objects", ***Transaction of CCEIC National New Computer Science and Technology and Computer Continuing Education***, China, vol. 25, no. 5, pp55-58, 2000.

Book Chapter:

5. **Li Tao** and K. Vijayan Asari, "A robust image enhancement technique for improving image visual quality in shadowed scenes," ***Lecture Notes in Computer Science***, Published by Springer-Verlag Heidelberg (ISSN: 0302-9743), ***Image and Video Retrieval***, Edited by W.-K. Leow, M.S. Lew, T.-S. Chua, W.-Y. Ma, L. Chaisorn, E.M. Bakker: ***Proceedings of the International Conference on Image and Video Retrieval 2005 – CIVR 2005***: (ISBN: 3-540-27858-3), vol. 3568/2005, pp. 395-404, July 2005

Refereed Conference Proceedings:

6. **Li Tao** and Vijayan K. Asari, "Modified luminance based MSR for fast and efficient image enhancement", ***IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2003***, Washington DC, pp. 174-179,October 15 - 17, 2003.

7. **Li Tao** and K. Vijayan Asari, "An integrated neighborhood dependent approach for nonlinear enhancement of color images", ***Proceedings of the IEEE Computer Society International Conference on Information Technology: Coding and Computing – ITCC 2004***, Las Vegas, Nevada, vol. 2, pp. 138-139, April 5-7, 2004.

8. Hau Ngo, **Li Tao** and K. Vijayan Asari, "Design of an efficient architecture for real-time image enhancement based on a luma-dependent nonlinear approach", ***Proceedings of the IEEE Computer Society International Conference on Information Technology: Coding and Computing – ITCC 2004***, Las Vegas, Nevada, vol. 1, pp. 656-660, April 5-7, 2004

9. **Li Tao** and K. Vijayan Asari, "Enhancement of images captured in an extremely low and non-uniform lighting environment", ***SensorsGov Expo and Conference - SensorsGov-2004***, Virginia Beach, VA, USA, September 13-15, 2004.

10. **Li Tao** and K. Vijayan Asari, "A novel technique for the extraction of depth information by gradient analysis on grayscale images," ***IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2004***, Washington DC, USA, ISBN 0-7695-2250-5, pp. 223-228, October 13 - 15, 2004.

11. Hau T. Ngo, **Li Tao**, and K. Vijayan Asari, "A nonlinear technique for enhancement of color images: an architectural perspective for real-time applications," ***IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2004***, Washington DC, USA, ISBN 0-7695-2250-5, pp. 124-129, October 13 - 15, 2004.

12. Satyanadh Gundimada, **Li Tao**, and K. Vijayan Asari, "Face detection technique based on intensity and skin color distribution" ***Proceedings of the IEEE International Conference on Image Processing – ICIP 2004***, Singapore, ISBN 0-7803-8555-1, pp. 1413-1416, October 24-27, 2004.

13. **Li Tao**, Richard Cortland Tompkins, and K. Vijayan Asari, "An illuminance-reflectance model for nonlinear enhancement of color images," ***IEEE Computer Society Proceedings of the Workshop on Face Recognition Grand Challenge Experiments – FRGC 2005*** in conjunction with the ***IEEE Conference on Computer Vision and Pattern Recognition – CVPR 2005***, San Diego, CA, June 20 – 25, 2005.

14. Adam Livingston, Hau Ngo, Ming Zhang, **Li Tao**, and Vijayan K. Asari, "Design of a real time system for nonlinear enhancement of video streams by an integrated neighborhood dependent approach", ***IEEE Computer Society Proceedings of the International Symposium on VLSI – ISVLSI 2005***, Tampa, Florida, pp. 301-302, May 11 – 12, 2005.

15. Hau Ngo, **Li Tao**, Adam Livingston, Ming Zhang, and K. Vijayan Asari, "A visibility improvement system for low vision drivers by nonlinear enhancement of fused visible and infrared video," ***IEEE Computer Society Proceedings of the 1st Workshop on Computer Vision Applications for the Visually Impaired – CVAVI 2005*** in conjunction with the ***IEEE Conference on Computer Vision and Pattern Recognition – CVPR 2005***, San Diego, CA, pp. 1-8, June 20 – 25, 2005.

16. **Li Tao**, Richard Cortland Tompkins, and K. Vijayan Asari, "An illuminance-reflectance model for nonlinear enhancement of video stream for homeland security applications," ***IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2005***, Washington DC, October 19 - 21, 2005.

17. **Li Tao**, Hau T. Ngo, Ming Zhang, Adam Livingston and Vijayan K. Asari, "Multi-sensor image fusion and enhancement system for assisting drivers in poor lighting conditions", ***IEEE International Workshop on Applied Imagery and Pattern Recognition, AIPR - 2005***, Washington DC, October 19 - 21, 2005.

18. **Li Tao**, Ming Jung Seow and Vijayan K. Asari, "Nonlinear color image enhancement to improve face detection in complex lighting environment", ***IS&T/SPIE Symposium on Electronic Imaging: Algorithms and Systems V***, San Jose, California, 15-19 January 2006.

19. **Li Tao** and K. Vijayan Asari, "A heuristic approach for the extraction of region and boundary of mammalian cells in bio-electric images," ***IS&T/SPIE Symposium on Electronic Imaging: Image Processing: Algorithms and Systems V***, San Jose, CA, January 15-19, 2006.

20. **Li Tao** and K. Vijayan Asari, "An efficient illuminance-reflectance nonlinear video stream enhancement model," ***IS&T/SPIE Symposium on Electronic Imaging: Real-Time Image Processing III***, San Jose, CA, January 15-19, 2006.

21. **Li Tao** and K. Vijayan Asari, "Region based multi-resolution image fusion for visibility improvement," ***30th International Congress of Imaging Science - ICIS'06***, Rochester, New York, May 7-11, 2006 *(accepted).*