MATHEMATICAL APPROACHES FOR IMAGE ENHANCEMENT PROBLEMS

Dongwook Cho

A thesis

in

The Department

of

Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

Concordia University

Montréal, Québec, Canada

April 2012

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: **Mr. Dongwook Cho**

Entitled: **Mathematical Approaches For Image Enhancement Problems**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Computer Science)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
    Dr. C. Wang

_____ External Examiner
    Dr. E. Dubois

_____ External to Program
    Dr. W.-P. Zhu

_____ Examiner
    Dr. T. Fevens

_____ Examiner
    Dr. A. Krzyzak

_____ Thesis Supervisor
    Dr. T. D. Bui

Approved _____
      Dr. V. Haarslev, Graduate Program Director

April 25, 2012 _____
      Dr. Robin Drew
      Dean of the Faculty of Engineering and Computer Science

Abstract

Mathematical Approaches For Image Enhancement Problems

Dongwook Cho, Ph.D.

Concordia University, 2012

This thesis develops novel techniques that can solve some image enhancement problems using theoretically and technically proven and very useful mathematical tools to image processing such as wavelet transforms, partial differential equations, and variational models. Three subtopics are mainly covered. First, color image denoising framework is introduced to achieve high quality denoising results by considering correlations between color components while existing denoising approaches can be plugged in flexibly. Second, a new and efficient framework for image contrast and color enhancement in the compressed wavelet domain is proposed. The proposed approach is capable of enhancing both global and local contrast and brightness as well as preserving color consistency. The framework does not require inverse transform for image enhancement since linear scale factors are directly applied to both scaling and wavelet coefficients in the compressed domain, which results in high computational efficiency. Also contaminated noise in the image can be efficiently reduced by introducing wavelet shrinkage terms adaptively in different scales. The proposed method is able to enhance a wavelet-coded image computationally efficiently with high image quality and less noise or other artifact. The experimental results show that the proposed method produces encouraging results both visually and numerically compared to some existing approaches. Finally, image inpainting problem is discussed. Literature review, psychological analysis, and challenges on image inpainting problem and related topics are described. An inpainting algorithm using energy minimization

and texture mapping is proposed. Mumford-Shah energy minimization model detects and preserves edges in the inpainting domain by detecting both the main structure and the detailed edges. This approach utilizes faster hierarchical level set method and guarantees convergence independent of initial conditions. The estimated segmentation results in the inpainting domain are stored in segmentation map, which is referred by a texture mapping algorithm for filling textured regions. We also propose an inpainting algorithm using wavelet transform that can expect better global structure estimation of the unknown region in addition to shape and texture properties since wavelet transforms have been used for various image analysis problems due to its nice multi-resolution properties and decoupling characteristics.

Acknowledgments

First and most of all, this thesis could not be possible without support of Prof. Tien Bui. His earnest advice coming from various research and life experiences, profound insight on the discipline, and truthful encouragement inspired me to conclude this thesis. Again I truly appreciate his full support, passionate guidance and patience. Also, I would not forget academic atmosphere at Concordia and joyful discussions with friends and colleagues, Shuo Li, Chao Jin, Guangyi Chen, Wumo Pan, Xiaojun Du, Calin Brezeanu, Israat Tanzeena Haque, and other friends. The pleasant and constructive discussion and comments gave me valuable inspiration. I also appreciate thoughtful consideration related to my thesis work and support of colleagues at Genetec and Orus, especially Drs. Jean-Pierre Polonovski and Paule Brodeur. In addition, I could obtain various hand-on research experiences on image processing and computer vision by working with them. Finally, I express my sincere gratitude to my wife Hyojung. Her sacrifice and devotion to the family, patience, and encouragement were indispensable for this outcome. I'm also proud of Daniel, Gabriel, and Angela growing well during those years.

Contents

Acronyms

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
ABME	Absolute mean brightness error
AWBS	Automatic wavelet base selection
BV	Bounded variation
CES	Color enhancement by scaling
DCT	Discrete cosine transform
DWT	Discrete wavelet transform
FWT	Forward wavelet transform
HVS	Human visual system
iid	independently and identically distributed
IWT	Inverse wavelet transform
JPEG	Joint photographic experts group
MAP	Maximum *a posteriori*
MB	Mean brightness
MC	Mean contrast
MCA	Morphological component analysis
MGGD	Multivariate generalized Gaussian distribution
MPEG	Moving picture experts group
MRA	Multiresolution analysis
MRF	Markov random field
NASA	National aeronautics and space administration
PDE	Partial differential equation
PSNR	Peak signal to noise ratio
TV	Total variation
WT	Wavelet transform

List of Figures

List of Tables

Chapter 1

Introduction

In this modern world, we cannot imagine a life without digital images. They are not only limited to store one's memory, please our eyes, and express the world. They depict micro- and macro-spaces, inner material, external world, nature, and history. Images are essential information that can capture the moment of the universe.

However, the images we possess may not be sufficiently good enough. We may lose the information of the scene during the acquisition, conversion, communication, and processing. First, images can be degraded when they are acquired due to lighting condition, sensor resolution and quality, or any limitations of the acquisition system. The degradation could come about when the data compression or transmission to other devices is made. It can also be due to certain processing, transformation, modification, or even intentional changes that may contaminate an image. In addition, any plain 2D discrete image cannot contain all the information for the real scene because the moment of the world is continuous 3D information, and hence loss of information such as object occlusion and low resolution problems necessarily occurs.

Image enhancement is a major and fundamental topic since the time image processing has emerged. The ultimate objective of image enhancement in a broad sense is to improve a degraded image that can express all the information of the scene. This is mathematically an inverse problem and it is impossible to solve in general. However, it is not impossible to obtain *something better* under certain conditions and

the desire to have something better pushes forward advances in image processing research. In this sense, there are practically two main purposes for developing image enhancement algorithms. The first is to satisfy and comfort human visual system (HVS). The second is to estimate the missing information to obtain an image close to the original scene for the analysis either by a human being or an intelligent machine vision system. The former is subjective and aesthetic while the latter is objective and informative. An enhanced image is expected to have better brightness and contrast, good color consistency, reduced noise or defect, less visual artifacts, better resolution, or even contextually meaningful information. Depending on the quality of a given degraded image, each of these improvement factors becomes an important subtopic separately, namely, denoising, contrast enhancement, white balance, deblurring, demosaicking, deblocking, super-resolution, inpainting, sharpening, smoothing, interpolation, gamma correction, chromaticity enhancement, and so forth. Figure 1 shows the examples of image enhancement subtopics. These separate problems can also be considered together and solved in an integrative manner. Various approaches have been used to solve each image enhancement problem as can be seen in figure 2.
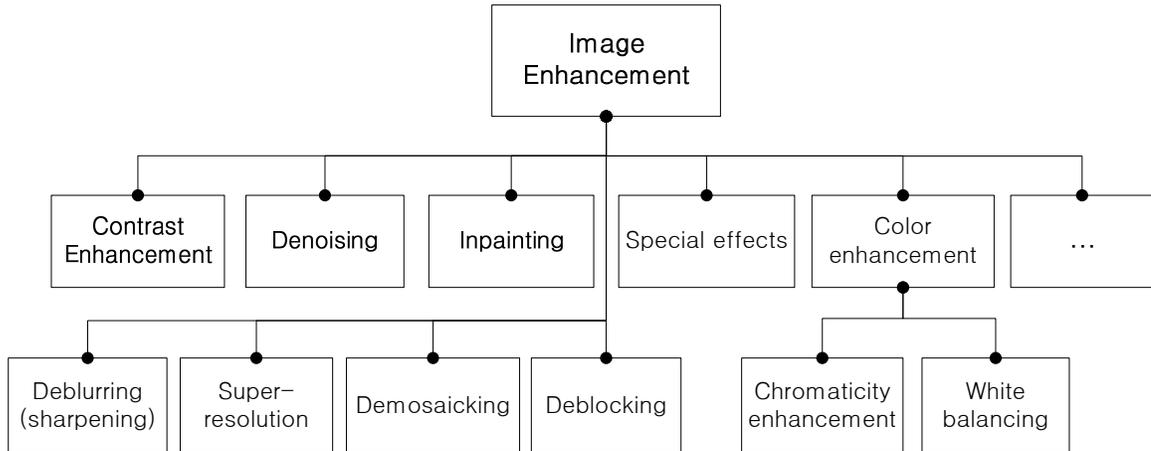
Figure 1: Subtopics of image enhancement

This thesis develops novel techniques based on different mathematical approaches to solve some image enhancement problems. Mathematical tools such as wavelet transforms, partial differential equations (PDE), and variational methods have been
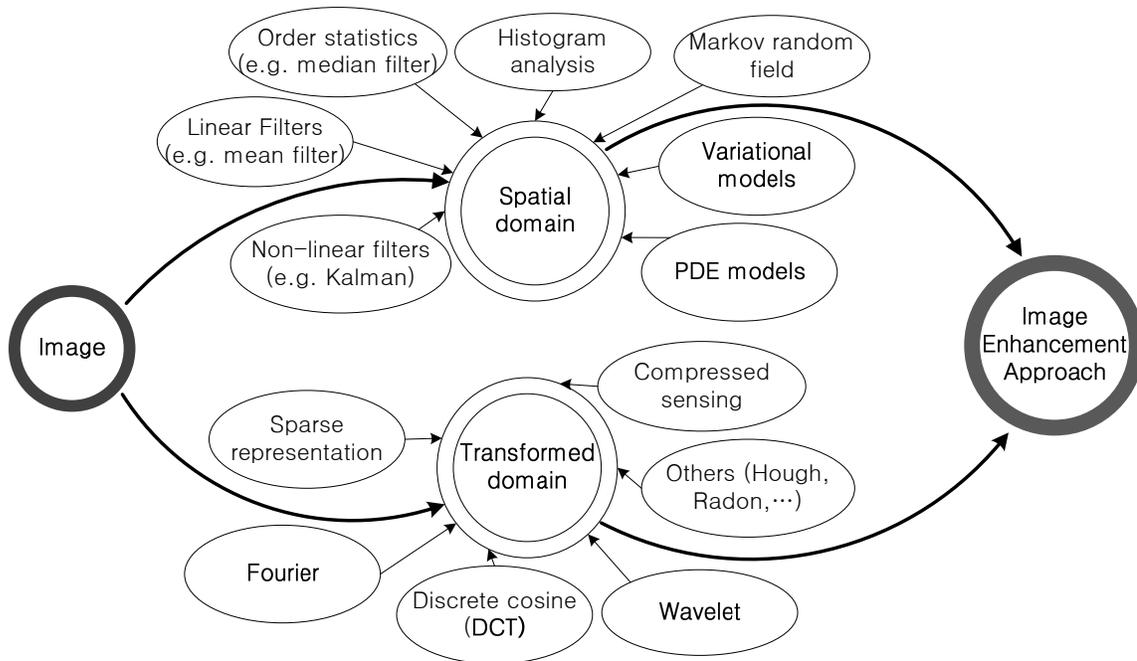
Figure 2: Various approaches for solving image enhancement problems

proven in this thesis to be very powerful in solving many image enhancement problems. Also recently developed mathematical tools such as compressed sensing and sparse representations are leading the new applications in image processing. The theoretical background and review on the key mathematical approaches used in this thesis and the image processing community is introduced in Chapter 2.

Among many image enhancement subtopics depicted in figure 1, this thesis covers mainly three problems, i.e. color image denoising, contrast enhancement, and image inpainting problems. First of all, a color image denoising framework is proposed in Chapter 3. Image denoising is a classical yet still hot topic. It is attractive to many researchers because any imaging system could have random noise and there is no universal solution to the denoising problem. Wavelet transform is very useful to solve this problem and there have been tremendous achievements along this direction because wavelet transform is capable of separating a small number of salient scaling coefficients with low-pass filter and a large number of detailed wavelet coefficients with high-pass filter. Noise is mainly contaminated in the wavelet coefficients. In addition, most of existing denoising approaches using wavelet transform focus on monochrome

images while the majority of images are in color. A color image denoising framework introduced in this thesis achieves high quality denoising results by considering correlations among color components. The proposed approach is flexible enough to allow employing any existing wavelet-based monochrome denoising approaches.

Chapter 4 proposes a new technique for image brightness and contrast enhancement. Comfortable brightness, contrast, and color consistency are major enhancement factors to HVS since photoreceptors in retina (rod and cones) are stimulated differently by lighting intensity and wavelengths, and they transduce the different levels of stimuli to send the analyzed signal to the brain. A contrast enhancement algorithm makes a degraded image visually better perceived. This is a subjective problem. We presents a new and efficient framework for image contrast and color enhancement in the compressed wavelet domain. The proposed approach is capable of enhancing both global and local contrast and brightness as well as preserving color consistency. The framework does not require inverse transform for image enhancement since linear scale factors are directly applied to both scaling and wavelet coefficients in the compressed domain, which results in high computational efficiency. Also contaminated noise in the image can be efficiently reduced by introducing a wavelet shrinkage term adaptively for analysis scale. The proposed method is able to enhance a wavelet-coded image computationally efficiently with high image quality and less noise or other visual artifacts. The experimental results show that the proposed method produces encouraging results both visually and numerically compared to some existing approaches.

Chapter 5 discusses image inpainting problem. *Inpainting* is originally an artistic term describing a procedure to restore a damaged painting or picture such as medieval artwork and old pictures. In computer vision, inpainting is a process to estimate partially unknown regions in an image and make the whole image region complete. Therefore, it is also called image completion and essentially similar to object removal, disocclusion, block recovery, texture synthesis, and image interpolation. Unlike image denoising and contrast enhancement problems, image inpainting

4

is a high-level computer vision problem because inpainting process reflects *perceptual interpolation*, which implies a coherent reaction of both *vision* and *thought* that cannot be separated. Perception implicates the entire process from sensory input to the related mental analysis of the given information. The chapter discusses problem set, literature review, psychological meaning, and challenges on image inpainting problem and its related topics. The existing approaches using various mathematical approaches introduced in Chapter 2 such as PDEs, variational, image decomposition, and compressed sensing are also explained. In Section 5.3, an inpainting algorithm using energy minimization and texture mapping is proposed. The Mumford-Shah energy minimization model detects and preserves edges in the inpainting domain by detecting both the main structure and the detailed edges. The proposed approach utilizes hierarchical level set method that is faster than the standard way and guarantees convergence independent of initial conditions. The estimated segmentation results in the inpainting domain is stored in the *segmentation map.* The segmentation map is referred to by a texture mapping algorithm for filling textured regions. In Section 5.4, we propose an image inpainting algorithm using wavelet transform that can expect better global structure estimation of the unknown region in addition to considering shape and texture properties since wavelet transforms have been used for various image analysis problems due to its powerful multi-resolution and decoupling properties.

The last chapter concludes and summarizes the whole thesis. Future works are also proposed.

Chapter 2

Mathematics in Image Enhancement

This chapter introduces different mathematical approaches used for enhancing images, and their relations to each other, and use in image processing and enhancement problems. In the first overview section, the relations and connectivities between mathematical approaches useful in image enhancement are discussed. More details and existing models on each approach are introduced in the following sections thereafter.

2.1   Overview : PDE to Compressed Sensing

An image is represented as a two dimensional signal with spatial coherence and common properties. Two dimensional signals can be created and with different intensity values on a 2D space but human knows if the 2D signal is *meaningful*. For example, we know a discrete 2D signal produced with independent and identically distributed (*iid*) random variables is not meaningful and we do not consider it as an image. In short, our visual system can analyze, understand, and classify any kind of images, but shows an interest in a meaningful one. Therefore, only meaningful images are considered in this thesis.

Mathematically, an image can be defined as a 2D function $u(x, y)$ that is neither deterministic nor fully random. Based on observations, it can be mathematically modeled by functions with various conditions and special properties. A function u is generally continuous, but it can be discretized just like a digital image represented pixel by pixel.

Image enhancement is a process to estimate a *good* image $\hat{u}(x, y)$ from a given contaminated or incomplete image $u_0(x, y)$. Now the questions are what a *good* image is and how it can be defined. A good image must satisfy human visual system by allowing good continuity, appropriate smoothness, yet clear edges with some variations. Some mathematical tools are able to model good images and express their properties.

Modeling using partial differential equation (PDE) is one way to define a good image model and behaviors. For instance, let us consider the heat equation, or more generally the diffusion equation as follows:

$$\frac{\partial u}{\partial t} = \nabla \cdot c \nabla u, \tag{1}$$

where c is the diffusion coefficient. Depending on the definition of c, the diffusion equation (1) can be isotropic or anisotropic. This equation models the changes of an image u over time variable t, i.e. image intensity spreads throughout the image space. Equation (1) implicitly shows that it can enhance an image by smoothing the noise while a proper choice of the diffusion coefficient c can preserve the edges. An isotropic diffusion equation can smoothe a noisy image in all directions for both textures and edges. While an anisotropic diffusion equation can preserve edges in preferred directions. In this case, one can design the diffusion coefficient as a tensor and hence edges can be preserved effectively. PDE model essentially covers a continuous image domain, but the practical solution requires discretization by finite difference or multigrid methods.

Variational approach from calculus of variations deals with finding optima (minima or maxima) of a given functional (function of function). It is very useful to define an image model and its behaviors in terms of functionals. A solution of the variational model can be found by associating the Euler-Lagrange equation and converting the variational problem to a PDE formulation. For example, a total variational (TV) model is one of the successful variational models used for image enhancement and it can be defined as

$$\inf_u F(u) \;=\; \inf_u \int_\Omega |u - u_0|^2 \, dxdy + \lambda \int_\Omega |\nabla u| \, dxdy \tag{2}$$

where u_0 is an initial degraded image, λ is a scaling parameter (used for Lagrange multiplier) and $\Omega \subseteq \mathbb{R}^2$ is image domain, which is open and bounded (often assumed to be a Lipschitz domain to express continuity, smoothness, and regularity of image). The associated Euler-Lagrange equation can be obtained as follows:

$$u = u_0 + 2\lambda \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) \tag{3}$$

with Neumann boundary condition from the normal derivative on the boundary surface $\partial\Omega$ (i.e. the derivative of image u along the direction of surface normal), $\frac{\partial u}{\partial \mathbf{n}}\big|_{\partial\Omega} = 0$. It is noticeable that equation (3) is also the case of the diffusion equation (1) when $c(|\nabla u|) = \frac{2\lambda}{|\nabla u|}$. The example of TV model shows strong relationship between PDE and variational approaches.

Bayesian probability and statistical inference also contribute to model an image. We can consider a Bayesian framework using the maximum *a posteriori* (MAP) estimator:

$$\hat{u} = \arg \max_u \ln p(u|u_0) = \arg \max_u [\ln p(u_0|u) + \ln p(u)], \tag{4}$$

where u_0 is a given contaminated image and $\hat{u}$ is an estimated enhanced image. Equation (4) is an optimization problem just like variational formulation. The probability model $p(u|u_0)$ is directly related to the functional in the variational approach, i.e.

maximization problem in equation (4) is equivalent to minimize $-\ln p(u|u_0)$. As will be shown in Section 2.4, equation (2) is interpreted by Bayesian framework.

Another efficient mathematical image analysis tool used in this thesis is wavelet transform. Wavelet-based image representation has been used to take various advantages of wavelet analysis such as multi-scale analysis, compaction, data separations, and perfect reconstruction properties. Wavelet is successfully used especially in the discrete image domain. The construction of wavelets begins with a scaling equation, which is a basic dilation equation from a two-scale difference equation:

$$\phi(x) = \sum_{k=0}^{n} c_k \phi(2x - k), \tag{5}$$

where the coefficients $\{c_k\}$ are given with $k = 0, \cdots, n$ and $\sum_k c_k = 2$ when $\int \phi dx = 1$. Then, the wavelet ψ is obtained by taking the differences as follows:

$$\psi(x) = \sum_{k=1-n}^{1} (-1)^k c_{1-k} \phi(2x - k). \tag{6}$$

Equations (5) and (6) are of special finite difference form. From a fundamental point of view, the numerical solutions of a PDE model of an image require a discretization of the PDE into a finite difference form similar to equations (5) and (6). In a way, the continuous domain represented by PDEs and the discrete domain represented by wavelet transforms are related through the finite difference formulation. As an example, equation (3) can be rewritten as $u_0 = u + (-2\lambda)\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$, where u is the smooth and noiseless component of the image that belongs to the Sobolev space, while $\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$ is the divergence of normalized gradient, which is Laplacian and represents the detailed variations (i.e. edges in image). Similarly, in the wavelet domain the image u_0 can be decomposed by wavelet transform into coarse scaling coefficients having low pass filtered smoothness and detailed wavelet coefficients that have strong magnitudes along the edges. Hence, u and $(-2\lambda)\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$ in equation (3) can be interpreted as scaling and wavelet subband respectively.

Discrete wavelet transform can also be expressed as linear matrix formulation as follows:

$$\mathbf{x} = \mathbf{\Phi\alpha}, \tag{7}$$

where $\mathbf{x}$ is a vector with size m (i.e. 1D signal or reshaped 2D image), $\mathbf{\alpha}$ is transformed coefficient vector with size n (i.e. scaling and wavelet coefficients), and $\mathbf{\Phi}$ is an $m \times n$ inverse transform matrix. If the size of transformed coefficients is the same as that of the input vector, then $n = m$ and we have a complete system. If $n > m$, the system is redundant and we have an over-complete system. Equation (7) can be generalized by sparse linear model, which represents a signal or image as sparse coding and dictionary. In this case, $\mathbf{\Phi}$ is called *dictionary* whose column vectors are basis vectors, and $\mathbf{\alpha}$ is called *sparse code* or *sparse vector* that contains sparse non-zero elements and many zero's with $\mathbf{\alpha} \in \mathbb{R}^n$. Sparse representation is a mathematical formulation that can decompose a signal into elementary atoms and take advantages of over-completeness (or redundancy) and sparseness. Given $\mathbf{x}$, estimation of efficient sparse code $\mathbf{\alpha}$ among many solutions is a key problem, i.e. *sparsity* of $\mathbf{\alpha}$ should be minimized. The sparsity can be measured by L_0 norm $\|\mathbf{\alpha}\|_0^0$, i.e. the number of non-zero components (Hamming distance from zero vector compared to Euclidean distance of L_2 norm). $L - 0$ norm is not F-norm since it is not continuous [41]. Now the problem is to solve the following formulation:

$$\min_{\mathbf{\alpha}} \|\mathbf{\alpha}\|_0^0, \text{ subject to } \|\mathbf{x} - \mathbf{\Phi\alpha}\|_2 \leq \delta. \tag{8}$$

Since L_0 minimization problem is technically difficult (combinatorial and NP-hard [103]), L_1 norm often replaces the problem, i.e. solve $\min_{\mathbf{\alpha}} \|\mathbf{\alpha}\|_1$ instead of equation (8). The equivalent Lagrangian form of equation (8) can be expressed as follows:

$$\min \frac{1}{2}\|\mathbf{x} - \mathbf{\Phi\alpha}\|_2^2 + \lambda\|\mathbf{\alpha}\|_p^p, \tag{9}$$

where p is 0 or 1. This formulation is again similar to the total variational form in
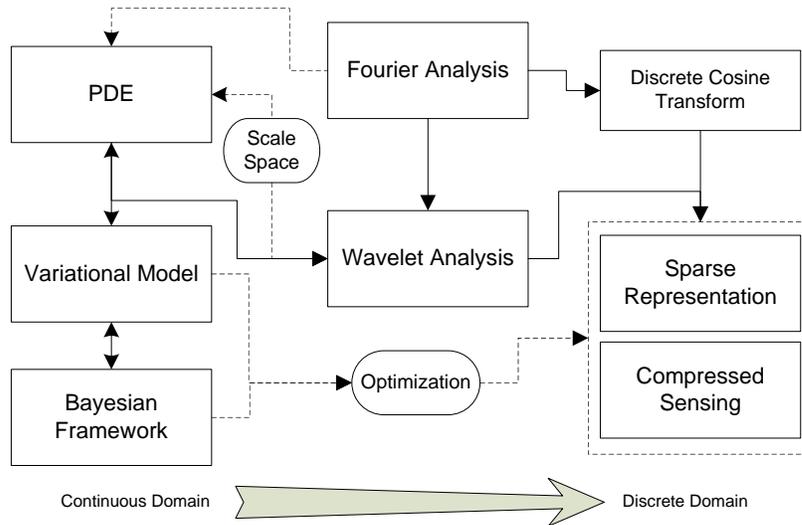
Figure 3: Relations between different mathematical tools

equation (2) when $p = 1$, and also the Bayesian framework in Section 2.4.

Recently developed theory on compressed sensing is also formulated under sparse representation. Using sparse representation in equation (7), a compressed signal $\mathbf{y}$ with size n is defined by

$$\mathbf{y} = \mathbf{M}\mathbf{x} = \mathbf{M}\mathbf{\Phi}\boldsymbol{\alpha} = \mathbf{\Theta}\boldsymbol{\alpha}, \tag{10}$$

where $\mathbf{M}$ is a measurement matrix. When $\mathbf{x}$ is k-sparse, $k \leq n < m$. Since compressed sensing is able to estimate precisely unknown information from partially given signal or image, recent works show promising results for solving image enhancement problems.

In summary, image space can be modeled and analyzed by different mathematical approaches, which are fundamentally closely related to each other as shown in the relation diagram of figure 3. These mathematical tools have been applied to image enhancement problems successfully, and hence we briefly discuss them in the following sections.

2.2 Partial Differential Equation (PDE) Models

PDE models are useful tools to describe and simulate physical behaviors dynamically in image just like many physical phenomena such as heat conduction, solid and fluid dynamics.In this section, we introduce a few useful PDE models for image enhancement problems.

2.2.1 Diffusion Equation

Diffusion equation (1) has some interesting properties for image processing. It can be rewritten for 2D image as

$$\frac{\partial u(x,y,t)}{\partial t} = \nabla \cdot \left[c(x,y,t) \nabla u(x,y,t) \right]. \tag{11}$$

In image enhancement problem, diffusion process would simulate the suppression of noise as the time variable (t) changes. Equation (11) is isotropic if the diffusion coefficient c is constant and anisotropic if c is a symmetric positive definite matrix or a function that can control the rate of diffusion. For example, Perona and Malik's anisotropic diffusion in [111] is efficient to preserve edges in image by defining $c(||\nabla u||) = \exp\left\{-(||\nabla u||/K)^2\right\}$ and $c(||\nabla u||) = \frac{1}{1+(||\nabla u||/K)^2}$, which are functions of image gradient magnitude.

In case of isotropic diffusion, $\frac{\partial u}{\partial t} = c\Delta u$, where Δ is a Laplacian operator. Derivations by 2D Fourier transform lead to the following solution:

$$u(x,y,t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(v_x, v_y, 0) \frac{1}{4\pi ct} \exp\left\{-\frac{(v_x - x)^2 + (v_y - y)^2}{4ct}\right\} dv_x dv_y \tag{12}$$

In other words, equation (12) is a convolution of 2D Gaussian smoothing filter to an initial image u_0 with variance $2ct$ when $u_t(x,y) = u(x,y,t)$. This is also a scale-space representation, where t is a scale parameter that decides a scale level and leads to multi-scale analysis of an image [5].

2.2.2 Navier-Stokes Equation

Incompressible fluid flow can be modeled Navier-Stokes equation:

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{\nabla P}{\rho} + \nu \Delta \mathbf{v}, \text{ and } \nabla \cdot \mathbf{v} = 0 \tag{13}$$

where $\mathbf{v}$ is the velocity vector field of the fluid parcel, P the scalar pressure, ν the kinematic viscosity, and ρ the fluid density. If we assume $w = \nabla \times v$, which is called *vorticity* and curl operator ($\nabla \times$) is applied to equation (13), we obtain

$$\frac{\partial w}{\partial t} + \mathbf{v} \cdot \nabla w = \nu \Delta w \tag{14}$$

since $w_t = -\nabla \times \mathbf{v}_t$, $\mathbf{v} \cdot \nabla w = -\nabla \times (\mathbf{v} \cdot \nabla \mathbf{v})$, and $\nabla \times \nabla P = 0$.

It is interesting to note that the Navier-Stokes equation can also be seen from the probability and statistical point of view. In that the Navier-Stokes equation is a macroscopic form derived from the higher level Boltzmann equation which describes the evolution of the distribution function of particles in a fluid [36].

In 2D space, the velocity $\mathbf{v}$ can be expressed as 90 degrees rotation of the gradient of a stream function. Since we can consider a stream function as an image intensity, u, $\mathbf{v} = \nabla^{\perp} u$ (i.e. isophote direction), therefore, equation (14) implies that the following equation must be satisfied for steady state inviscid flows:

$$\nabla^{\perp} u \cdot \nabla(\Delta u) = 0. \tag{15}$$

Here, the vorticity can be related as $w = \Delta u$ by definition. Hence, equation (15) means the Laplacian of the stream function (i.e. vorticity) has the same level curves as the stream function. In addition, we can apply anisotropic diffusion in equation (11) to a vorticity transport equation (14) as follows:

$$\frac{\partial w}{\partial t} + \mathbf{v} \cdot \nabla w = \nu \nabla \cdot (c(|\nabla w|)\nabla w) \tag{16}$$

which satisfies the Poisson equation $\Delta u = w$ with the Dirichlet boundary condition $u|_{\partial \Gamma} = u_0$.

Similar to the diffusion equation, the Navier-Stokes equation and its variations have been used by many authors in image processing and image enhancement.

2.3 Variational Models

There have been various variational models applied in image processing and computer vision. More general functional form often used in image processing problem is in the following:

$$\inf_u F(u) = \inf_u \int_\Omega f(x, y, u(x, y), \nabla u(x, y)). \tag{17}$$

In this section, we introduce some of the most interesting models used in this thesis.

2.3.1 Total variation (TV) Model

In the earlier section, we have briefly observed total variational model (equation (2)), its equivalent PDE form (equation (3)), and its relations to other mathematical approaches. In image processing, it has been claimed that TV norm, $\int_\Omega |\nabla u|$, is more appropriate than L_2 norm due to its non-linearity. An image u that satisfies TV has bounded variation (BV), which allows not only smoothness but also sharp edges in the mathematical formulation. In [123], an image denoising approach was first proposed to apply TV model to a noisy image. Other image processing applications such as inpainting, segmentation, deblurring, and super-resolution have been successfully developed as we will see some in the following chapters.

2.3.2 Euler Elastica Model

In [101], Mumford introduced Euler's elastica model in computer vision as a prior curve model. The original problem was first proposed and solved by Euler and the

optimized curves were called *elastica*. The model can be generally described by minimizing the following integral:

$$\int_C (\alpha\kappa^2 + \beta)ds, \tag{18}$$

where α and β are constant parameters, κ the curvature of planar curves C, and ds the arc length. The curvature can be given by the divergence of normalized image gradient as follows:

$$\kappa = \nabla \cdot \mathbf{n} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right). \tag{19}$$

We can notice that the curvature term in equation (19) can be plugged into Euler-Lagrange equation of TV model in equation (3), i.e. $u = u_0 + 2\lambda \cdot \kappa(u)$. The endpoints of C are not considered and the energy functional is minimized over the variations of C which preserve the total length. In an intensity image u, which ranges in $[0, 1]$, the curves can be considered as level lines [92], which are boundaries of upper level sets or $X_\lambda = \{x, u(x) \geq \lambda\}$ at each intensity level λ. Analysis of all the level lines in an image can contribute useful image analysis information. The use of level lines in the detection of texts in natural images can be found in [107]. A solution to inpainting problem using elastica model can be found in [17] (see also Variational Approaches in Section 5.2.1).

2.3.3 Mumford-Shah Model

Mumford-Shah (MS) variational model proposed by Mumford and Shah [102] has been deeply studied and used in computer vision due to flexible curve evolution and numerical advantages for implementation. One of the most successful applications using the Mumford-Shah model is image segmentation [21, 143, 60, 149, 70, 151], which enables to analyze noise-invariant global image structure. Mumford-Shah energy functional is defined in the following:

$$F(u, C) = \int_\Omega |u - u_0|^2\, dxdy + \mu \int_{\Omega \backslash C} |\nabla u|^2\, dxdy + \nu\, |C|, \tag{20}$$

where u_0 is the original image, u is the smooth approximation of u_0, C and $|C|$ are segmentation curve and the curve length, Ω and $\Omega \setminus C$ represent respectively image domain and the image domain that excludes segmented curves, and μ and ν are given parameters used to adjust the effects of different terms.

2.3.4   Image Decomposition and Variational Approach

An image can be considered as a combination of image structure (cartoon) and details (texture and noise), i.e. $u = v + w$, where v and w represent a cartoon image and a textured image of u respectively. Meyer's model [96] proposes that a piece-wise-smooth image v can be obtained by TV model in equation (2). The model also characterizes oscillating property of texture image.

Based on Meyer's model, an image decomposition method by Vese and Osher [150] was proposed by optimizing the following functional:

$$
\begin{aligned}
G_p(v, g_1, g_2) &= \int |\nabla v| + \mu \|w\|_* \\
&= \int |\nabla v| + \lambda \int |u - v - \partial_x g_1 - \partial_y g_2|^2 dx dy + \mu \left[\int |g|^p dx dy \right]^{1/p} \quad (21)
\end{aligned}
$$

where the norm $\|w\|_*$ is the lower bound of all L^∞ norms of the functions $|g| = \sqrt{g_1^2 + g_2^2}$, λ, $\mu > 0$ are weight parameters, and $p \to \infty$. In this case, v bounded variation (BV), and $w = \partial_x g_1 + \partial_y g_2$ is in Banach space. In equation (21) u can ensure to be approximation of $v + w$ by the second term. Criterion (21) yields the Euler-Lagrange equation when $p = 1$ as follows:

$$
v = u - \partial_x g_1 - \partial_y g_2 + \frac{1}{2\mu} \nabla \cdot \left(\frac{\nabla v}{|\nabla v|} \right) \quad (22)
$$

with $\frac{\lambda g_1}{|g|} = 2\mu \left[\partial_x(v - u) + \partial_{xx}^2 g_1 + \partial_{xy}^2 g_2 \right]$ and $\frac{\lambda g_2}{|g|} = 2\mu \left[\partial_y(v - u) + \partial_{xy}^2 g_1 + \partial_{yy}^2 g_2 \right]$.

2.4 Bayesian Framework and Variational Approaches

MAP formulation in equation (4) can be translated as a minimization problem instead of maximization. If we assume that u_0 is blurred by a blurring operator K which is linear and contaminated with white Gaussian noise, the following formulation can be considered:

$$\hat{u} = \arg\min_u F(u) = \arg\min_u \left[\frac{1}{2\sigma^2 |\Omega|} \int_\Omega |Ku - u_0|^2 \, dxdy + f(u) \right], \qquad (23)$$

where $F(u) = -\ln p(u|u_0)$, $f(u) = -\ln p(u)$, σ^2 the noise variance, and $|\Omega|$ the Lebesgue measure. We can notice that equation (23) is analogous to other energy functional introduced. For example, if we set $f(u) = \lambda \int_\Omega |\nabla u| \, dxdy$ and there is no blurring, the minimization functional has the same form as equation (3).

2.5 Wavelet Transforms

For the last two decades, wavelet transform has been widely used in image processing due to various properties such as multi-resolution analysis, perfect reconstruction, filtering ability, locality, linearity, separability, sparsity, orthogonality (for orthogonal wavelet), and so on. For example, its efficient separability into dense scaling coefficients and sparse wavelet coefficients lead to good image compression methods. Decomposition ability into image structure (scaling subband) and detailed texture and noise (wavelet subband) makes efficient denoising possible. In this section, basic notions and mathematical properties of wavelet transforms for image enhancement are described. The interesting relationship between wavelet transforms and PDEs has been discussed in Section 2.1. In this thesis, the formulation is presented based on filtering theory to describe discrete wavelet transform in a simple way rather than discussing mathematical properties which can be found in many good books on the subject.

The one-dimensional forward discrete wavelet transform can be expressed as follows:

$$s_{j+1,k} = \sum_{m=-\lfloor \frac{p-1}{2} \rfloor}^{\lfloor p/2 \rfloor} \phi_m s_{j,2k-m} \tag{24}$$

$$w_{j+1,k} = \sum_{m=-\lfloor \frac{q-1}{2} \rfloor}^{\lfloor q/2 \rfloor} \psi_m s_{j,2k-m+1}, \tag{25}$$

where $s_{j,k}$ and $w_{j,k}$ are scaling and wavelet coefficients in j-th level and ϕ_m and ψ_m are the scaling and wavelet filters with length p and q respectively. Also, $\sum \phi_m = \sqrt{2}$ and $\sum \psi_m = 0$. In general, the filters ϕ and ψ should satisfy equations (5) and (6).

The following example defines forward transform of Le Gall 5/3 filter, which is biorthogonal filter that is used for lossless compression standard of JPEG 2000:

$$s_{j+1,k} = \sqrt{2} \left[\frac{3}{4} s_{j,2k} + \frac{1}{4} \left(s_{j,2k-1} + s_{j,2k+1} \right) - \frac{1}{8} \left(s_{j,2k-2} + s_{j,2k+2} \right) \right] \tag{26}$$

$$w_{j+1,k} = \frac{1}{\sqrt{2}} \left[s_{j,2k+1} - \frac{1}{2} \left(s_{j,2k} + s_{j,2k+2} \right) \right]. \tag{27}$$

The inverse wavelet transform for synthesis can be expressed as follows:

$$s_{j,k} = \frac{1}{2} \left[\sum_{m=-\lfloor \frac{p-1}{2} \rfloor}^{\lfloor p/2 \rfloor} \check{\psi}_m s_{j+1,k-m} + \sum_{m=-\lfloor \frac{q-1}{2} \rfloor}^{\lfloor q/2 \rfloor} \check{\phi}_m w_{j+1,k-m} \right], \tag{28}$$

where $\check{\psi}_m$ and $\check{\phi}_m$ are the synthesis filters of the scaling and wavelet functions and the coefficients s_{j+1}, k and w_{j+1}, k are upsampled. In the case of Le Gall 5/3 using equation (28), even and odd sampled scaling coefficients in the j-th level are respectively obtained by

$$s_{j,2k} = \frac{1}{\sqrt{2}} \left[s_{j+1,k} - \frac{1}{2} \left(w_{j,k-1} + w_{j,k} \right) \right] \tag{29}$$

$$s_{j,2k+1} = \sqrt{2} \left[\frac{3}{4} w_{j,k} + \frac{1}{4} \left(s_{j,k} + s_{j,k+1} \right) - \frac{1}{8} \left(w_{j,k-1} + w_{j,k+1} \right) \right]. \tag{30}$$

The shapes of the analysis and synthesis filters for Le Gall 5/3 are illustrated in

18

| (a) scaling analysis filter | (b) wavelet analysis filter |
| (c) scaling synthesis filter | (d) wavelet synthesis filter |
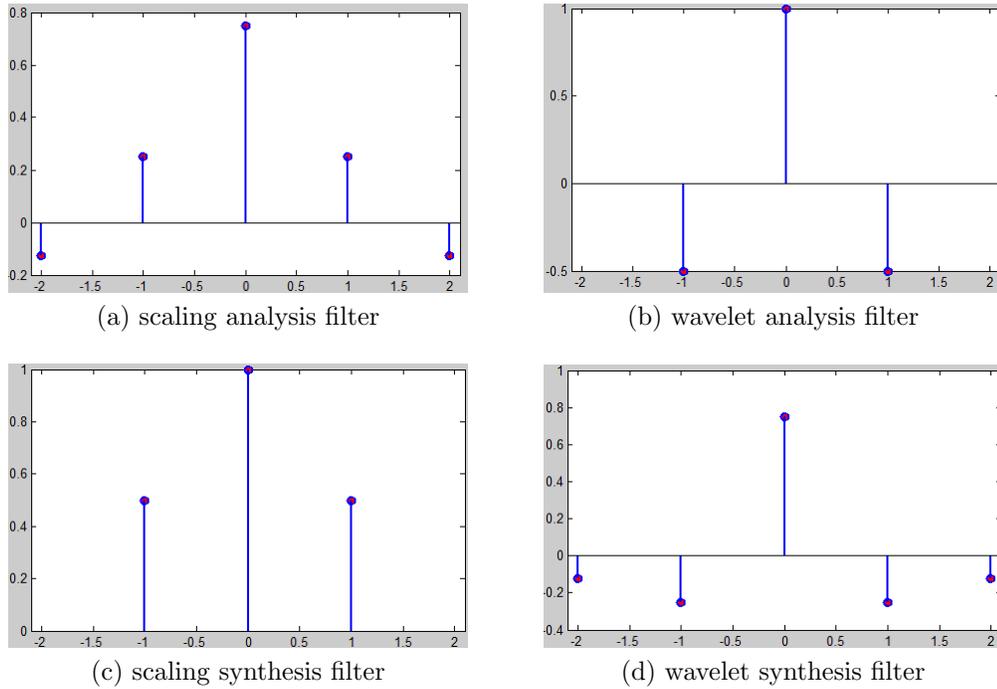
Figure 4: Le Gall 5/3 analysis and synthesis filters

figure 4.

In 2D wavelet transform, j-th level scaling coefficients are located in LL_j subband while wavelet coefficients are separated into HL_j, LH_j, and HH_j subbands. The 2D implementation can be made straightforward by applying 1D transform to rows and then columns of the image or the scaling subband in the previous decomposition level.

2.6 Sparse Representation and Compressed Sensing

Applications to image processing of sparse representation and compressed sensing are relatively new. They have been actively employed in image enhancement problems. Some reviews on sparse representation and compressed sensing can be found in [50, 48, 157, 113, 15, 121, 122, 52, 142, 43]. In this section, we discuss the basics of the theory and useful development as applied to images.

2.6.1 Sparse Representation

Image model in equation (7) is a general form for sparse coding, which includes wavelet packets, discrete cosine transform (DCT), Gabor filters, singular value decomposition (SVD), and many other recently developed filters, i.e. curvelet, ridgelet, bandelet, etc. Narrow meaning of sparse representation takes advantages of sparsity and optimality of the transformed domain and redundancy (or over-completeness). In this case, it is an ill-posed problem since $n > m$. To solve the optimization problem in equation (8) or (9), greedy algorithms such as basis pursuit[24], matching pursuit[91, 109, 105], LASSO [141], and gradient projection [57] have been developed. A dictionary Φ can be obtained directly from sparse filters, but it can be constructed by learning technique [47, 2, 98]. In image enhancement problems, the solution of optimization formulation in equation (9) or its variation can be considered as an enhanced image. In fact, various algorithms for denoising, super-resolution, and deblurring problems have been proposed based on this concept.

2.6.2 Image Decomposition and Sparse Representation

As shown in Section 2.3.4, an image can be decomposed into a piece-wise smooth cartoon image and a texture image. Sparse representation including wavelet transform is also good for image decomposition considering that it shares a variational scheme like equation 9 and both transformed cartoon and texture images can satisfy sparsity conditions.

Morphological component analysis (MCA) approach proposed in [134, 54, 135] uses linear sparse representations with different dictionaries and results in layer separation, i.e. $u = T_v \boldsymbol{\alpha}_v + T_w \boldsymbol{\alpha}_w$, where T_v and T_w are $N \times L$ matrices representing given dictionaries, and $\boldsymbol{\alpha}_v$ and $\boldsymbol{\alpha}_w$ are sparse vector with size L ($L \gg N$ and N is image size). In this case, the main problem requires to estimate the sparse representations, $\boldsymbol{\alpha}_v$ and $\boldsymbol{\alpha}_w$ as follows:

$$F(\boldsymbol{\alpha}_v, \boldsymbol{\alpha}_w) = \|\nabla T_v \boldsymbol{\alpha}_v\|_1 + \lambda \|u - T_v \boldsymbol{\alpha}_v - T_w \boldsymbol{\alpha}_w\|_2^2 + \mu(\|\boldsymbol{\alpha}_v\|_1 + \|\boldsymbol{\alpha}_w\|_1). \qquad (31)$$

The dictionaries can be chosen from known transforms as mentioned in the previous section. For example, local DCT, Gabor, or wavelet packets can be used for T_w and orthogonal wavelet transform, curvelet, ridgelets, and contourlets can be selected to compose T_v.

2.6.3   Compressed Sensing

Recall a compressed sensing formulation in equation (10) compared to equation (7). While sparse representation seeks the *sparsest* representation $\boldsymbol{\alpha}$, compressed sensing reconstructs probabilistically a signal or image $\mathbf{x}$ given a partial observation $\mathbf{y}$ [14]. However, there is very close relationship between sparse representation and compressed sensing theory since $\mathbf{x}$ can be expressed as a linear combination of a dictionary $\boldsymbol{\Phi}$ and sparse vector $\boldsymbol{\alpha}$.

For image enhancement problems, reconstruction of good image $\mathbf{x}$ with size m from a given incomplete data $\mathbf{y}$ with size l . The reconstruction of $\mathbf{x}$ is an ill-posed problem with $m > l$ like equation (7). In addition, there is an issue to design a stable measurement matrix $\mathbf{M}$ not to destroy significant observation and bases to keep k-sparsity. The former problem can be well-posed since we assume that $\mathbf{x}$ is k-sparse and $l \geq k$ with the restricted isometry property (RIP), $(1-\epsilon)\|\mathbf{v}\|_2 \geq \|\boldsymbol{\Theta}\mathbf{v}\|_2 \geq (1+\epsilon)\|\mathbf{v}\|_2$, and its related *incoherence* condition between rows of $\mathbf{M}$ and $\boldsymbol{\Phi}$ [42]. For example, $\mathbf{M}$ could be constructed by *iid* random variables from Gaussian distribution.

2.7   Conclusion

In this chapter we have described some mathematical approaches that have important applications to image enhancement, the topic of this thesis. These mathematical approaches are: PDEs, variational models, wavelet transform, sparse representation and compressed sensing. This chapter also describes the relationship that exists among these powerful methods.

Chapter 3

Image Enhancement : Color Image Denoising

Wavelet transform has been used for denoising problem and various approaches have been proposed for the last two decades since it has been proven to be efficient for removing noise of images as well as for signals. It has been shown that denoising using wavelet transforms produces superb results. This is because wavelet transform has the compaction property of having only a small number of significant coefficients and a large number of detailed coefficients as we discussed in the previous chapter. Therefore, it is possible to suppress the noise in the wavelet domain by killing the detailed coefficients that represent the detailed information as well as the noise.

Most of the image denoising approaches using wavelet transforms have been designed for monochrome images. However in practice more common image type is color image composed of multiple color components. Unlike monochrome images, color images can be expressed as multiple components of monochrome images or a set of pixels represented by vectors. Therefore it is possible to extend most of the existing denoising approaches to color image straightforwardly by denoising each color component independently. In this case, correlation and dependency between color components are ignored.

In order to estimate visually comfortable and meaningful color pixels from noisy

image, it is necessary to take advantage of correlation or dependency of color components. In the spatial domain, there have been many works including vector processing [115, 32, 144], PDE modeling [71, 4], and fuzzy approaches [125, 95]. For vector processing approaches, it is possible to develop a rule to estimate a correlated vector by averaging or utilizing order statistics between the vectors (or pixels).

In this chapter, a color image denoising approach using wavelet transform is introduced. Color image denoising problem in the wavelet domain has not been studied deeply. Only a few color image denoising algorithms using wavelet transform such as [63] and [140] have been developedat the time our work was proposed, which was initially published in SPIE Wavelet XI 2005 as an invited paper [28]. Since then, there have been more works on wavelet-based color image denoising approaches such as [81, 82, 86, 67, 164]. Some of recent methods combine a wavelet-based approach and spatial filters or PDE scheme and there have been some improvement. On the other hand, our work presented in this chapter is still comparable to the most recent state-of-the-art. In addition, the major contribution of this work is the color image denoising framework itself that can flexibly employ more recent and efficient wavelet-based denoising algorithm for a single channel in order to reach better performance.

The proposed general framework of color image denoising utilizes color components in the wavelet domain by decorrelating them statistically. In the decorrelated space, the sample space is randomly distributed and this results in more robust estimation for clean image. Before we present the color image denoising framework, some monochrome image denoising approaches using wavelet transform are briefly discussed because basic concepts are shared and any of these methods can be used in our proposed framework.

The organization of this chapter is as follows. In the following section, we briefly review wavelet-based monochrome image denoising approaches including the works originally proposed in [29] and [31]. In Section 3.2, a color denoising framework that performs decorrelation between color components in the wavelet domain is proposed. Experimental results are shown and compared with the existing methods in

Section 3.3. And finally we conclude this chapter in Section 3.4.

3.1 Monochrome Image Denoising Approaches Using Wavelet Transform

Wavelet coefficients are not strongly correlated, but they still have dependency on each other. So many of the recent works have taken into account this dependency in order to obtain better coefficient estimate. Cai and Silverman [13] proposed a simple and effective approach for signal denoising by incorporating the neighboring coefficients. Chen and Bui [22] designed new neighboring threshold for multiwavelet based on *NeighBlock*. Their method considers both multiwavelet properties and neighboring dependency. Mihcak et al. [97] proposed a local variance estimator to get a locally-adaptive shrinkage value. Malfait and Roose proposed an image denoising algorithm using Markov random field image model as *a priori* [88]. Also Pizurica et al.[114] considered a joint inter- and intra-scale statistical model and improved the approach of Malfait and Roose. A parent coefficient in the coarser level was also considered to estimate a threshold by Sendur and Selesnick [127]. They obtained better results when they applied the local variance together with the dual-tree complex wavelet transform (DT CWT) [128]. DT CWT provides better *shift-invariant* features and *directional selectivity* than the usual separable wavelet transform [72]. Portilla et al.[118] presented an image denoising algorithm which is based on a Gaussian scale mixture (GSM) model using an overcomplete multiscale oriented basis. They define a vector using neighboring coefficients and obtain an accurate estimate by the vector operations. These works show that incorporating different information like neighbors and parents is helpful to remove noise and preserve details for natural image denoising.

Also, two different denoising methods for monochrome image were proposed in master's thesis of the author [27]. These methods are described more in detail since they are used in the color image denoising framework as we will see in Section 3.2. In both methods, neighboring coefficients of given noisy image are utilized to estimate

those of clean image accurately. However they have different approaches to obtain the clean estimates. The first method utilizes the L^p-norm of a vector composed of all the related neighboring coefficients by comparing it with universal threshold [40]. On the other hand, the second method generally estimates the clean coefficients using Bayesian statistics based on our multivariate *a priori* model.

We define some common notations first. Let A be a clean natural image with size $N \times N$, B a noisy image which can be expressed as $B = A + \sigma C$, and C zero-mean Gaussian white noise, which is $C \sim N(0, 1)$. σ^2 is noise variance. After performing multiresolution wavelet decomposition on B, we get the wavelet coefficient $y_{j,k}$, which is the k-th wavelet coefficient in j-th level for B. Due to the linearity of the wavelet transform, we have:

$$y_{j,k} = x_{j,k} + \sigma z_{j,k}, \tag{32}$$

where $x_{j,k}$ and $z_{j,k}$ are the wavelet coefficients of A and C respectively in the same location as $y_{j,k}$.

3.1.1 *NeighLevel* : a simple and efficient shrinkage rule

In the wavelet domain, the strong dependency between parent and child coefficients has been widely realized in image coding and denoising since *zerotrees* were introduced by Shapiro [129]. Among them, Cai and Silverman [13] proposed a simple and effective approach for a signal denoising. The method, called *NeighBlock*, takes the neighboring coefficients into account and obtains a threshold by comparing the sum of squared neighboring coefficients with Donoho's universal threshold. In addition, parent and child coefficients have inter-dependency similar to neighbors. Therefore, if we can properly utilize neighbors spread both vertically and horizontally as shown in figure 5a, a better performance can be expected.

Based on these ideas, we proposed an efficient image denoising approach called *NeighLevel* for a monochrome image in [31]. This can be briefly described in the

(a) Choosing context for *NeighLevel*.

(b) Performance changes as function θ in $\lambda^* = \theta\lambda$ for different kinds and sizes of images when *NeighLevel* ($M = 5$) is applied using DT CWT.
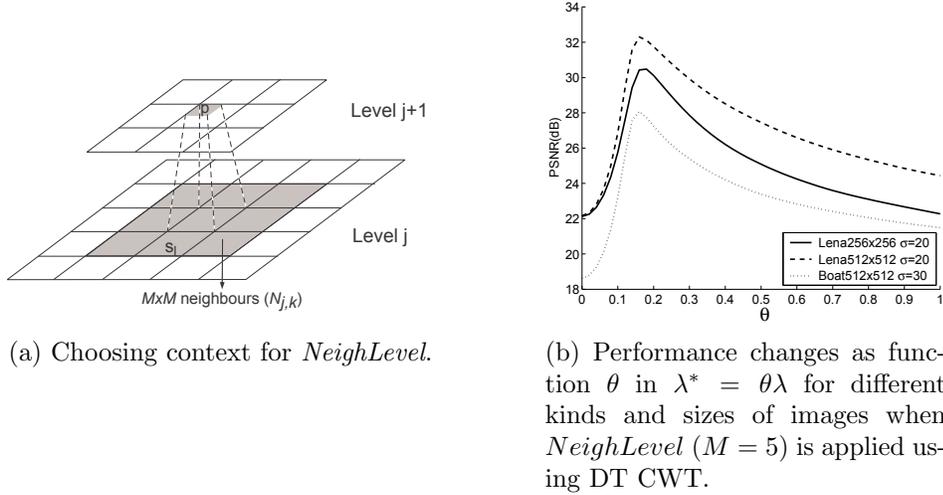
Figure 5: Main concept and analysis of *NeighLevel* approach.

following equation:

$$\hat{x}_{j,k} = y_{j,k} \left[1 - (M^2 + 1)\lambda^{*2} \middle/ \left(\sum_{s_l \in N_{j,k}} s_l^2 + p^2 \right) \right]_+$$

or in general,

$$\hat{x}_{j,k} = y_{j,k} \left(1 - d \left| \frac{\lambda^*}{\|\mathbf{y}\|_r} \right|^r \right)_+, \tag{33}$$

where s_l denotes the coefficient to be thresholded and its neighbors in an $M \times M$ window, and p is a corresponding parent of the coefficient to be thresholded, which is a coefficient matched in the coarser level (see figure 5a). $\mathbf{y}$ is a d-dimensional vector composed of all the related coefficients including the coefficient to be thresholded, neighbors and parent and $\|\mathbf{y}\|_r$ denotes L^r-norm (we mostly set r to 2). $\lambda^* = \theta\lambda$, where λ is the universal threshold $\lambda = \sqrt{2\sigma^2 \log N^2}$. θ is given as a parameter which satisfies $0 \leq \theta \leq 1$.

In equation (33), it should be noted that a normalized factor, d or $M^2 + 1$, is used which is the number of correlated elements in the context. By this rule, the effect of the local variance from the parent level is considered as well as from the current level.

In this method, it is important to choose an appropriate paramter θ. The universal

threshold is designed for smoothness rather than for minimizing the errors. So λ is more meaningful when the signal is sufficiently smooth or the length of the signal is close to infinity. Natural images, however, are usually neither sufficiently smooth nor composed of infinite number of pixels. In fact, if we suppose that an optimal threshold that minimizes mean square error (MSE) (or maximizes peak signal to noise ratio (PSNR)), θ is always much less than 1.0 for natural images as shown in figure 5b. Especially we got very similar θ value for different kinds and sizes of images when we applied soft thresholding rule. It might vary depending on the wavelet filter, but the appropriate range is similar for different images and noise level we have tested in our experiments.

3.1.2 Bayesian Estimation for Multivariate Statistical Model

Another method for monochrome image denoising proposed in [29] is to use the statistical model of clean wavelet coefficients in addition to taking advantage of neighboring coefficients. This drives us to use multivariate statistical model. The method uses a general estimation rule in the wavelet domain to obtain the denoised coefficients from the noisy image based on the multivariate statistical theory and Bayesian estimator. We briefly review the estimator in the following.

Let $\mathbf{x}$ be a d-dimensional wavelet coefficient vector, $\mathbf{x} = (x_1, x_2, \cdots, x_d)^t$, where x_1 is the wavelet coefficient under consideration and x_i $(i = 2, \cdots, d)$ are the related coefficients to be taken into consideration, e.g. neighbors, parent and offsprings. Here for simplicity, we replace the double subscripts in $x_{j,k}$ by a single subscript x_i. The corresponding vectors $\mathbf{y}$ and $\mathbf{z}$ can be similarly defined for the noisy image B and the noise C. We assume that x_i, y_i and z_i correspond to each other in both decomposition level and location. Therefore,

$$\mathbf{y} = \mathbf{x} + \sigma\mathbf{z}. \tag{34}$$

For the sake of simplicity, we omit subscripts j, k in equation (34) and the rest of the chapter.

Our concern lies mainly in obtaining the estimate of a clean wavelet coefficient vector, $\hat{\mathbf{x}}$. $\hat{\mathbf{x}}$ should be obtained only from $\mathbf{y}$, a wavelet coefficient vector of the noisy image B. One of the ways to estimate $\hat{\mathbf{x}}$ is to use MAP estimator to maximize $p(\mathbf{x}|\mathbf{y})$. MAP estimator for $\hat{\mathbf{x}}$ can be obtained as follows:

$$
\begin{aligned}
\hat{\mathbf{x}} &= \arg\max_{\mathbf{x}\in\mathbb{R}^d} \left[\ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x})\right] \\
&= \arg\max_{\mathbf{x}\in\mathbb{R}^d} F(\mathbf{x}),
\end{aligned}
\tag{35}
$$

where $F(\mathbf{x})$ represents the term inside $\arg\max$. This means that the optimal value $\hat{\mathbf{x}}$ with minimum probability error can be estimated by $p(\mathbf{y}|\mathbf{x})$ and $p(\mathbf{x})$.

From equation (34), $p(\mathbf{y}|\mathbf{x})$ is the multivariate Gaussian distribution with $N(\mathbf{0}, \Sigma_z = \sigma^2 \mathbf{I})$ since Gaussian noise is independently and identically distributed for each element of the vector. Hence,

$$
\ln p(\mathbf{y}|\mathbf{x}) = -\frac{d}{2}\ln\left(2\pi\sigma^2\right) - \frac{(\mathbf{y}-\mathbf{x})^t(\mathbf{y}-\mathbf{x})}{2\sigma^2}.
\tag{36}
$$

We assume that $p(\mathbf{x})$ is known. $p(\mathbf{x})$ might vary depending on the type of sample images. Also suppose that $g(\mathbf{x}) = \ln p(\mathbf{x})$ and there exists $\hat{\mathbf{x}}$ which satisfies $F(\hat{\mathbf{x}}) > \lim_{x_i \to \pm\infty} F(\mathbf{x})$. From equations (35) and (36), equation (35) is equivalent to the solution of the following equation:

$$
\begin{aligned}
\nabla F(\hat{\mathbf{x}}) = -\frac{\hat{\mathbf{x}} - \mathbf{y}}{\sigma^2} + \nabla g(\hat{\mathbf{x}}) = 0 \\
\Leftrightarrow \hat{\mathbf{x}} = \mathbf{y} + \sigma^2 \nabla g(\hat{\mathbf{x}}).
\end{aligned}
\tag{37}
$$

Therefore the estimate of $\mathbf{x}$ highly depends on the probability density of clean wavelet coefficients, $p(\mathbf{x})$.

The existing models for wavelet denoising are usually based on univariate statistical model whereas $p(\mathbf{x})$ is a multivariate pdf in our model. There are several multivariate functions which are symmetric spherically like multivariate Gaussian model. We use extended generalized Gaussian distribution (GGD) model [89] for

its simple form and to achieve good fitting results. We call this model multivariate generalized Gaussian distribution (MGGD):

$$p(\mathbf{x}) = \gamma \exp \left\{ - \left(\frac{(\mathbf{x} - \mu)^t \Sigma_{\mathbf{x}}^{-1} (\mathbf{x} - \mu)}{\alpha} \right)^{\beta} \right\}, \tag{38}$$

where α and β are parameters which can represent the spherical shape of the model and γ indicates a normalized constant defined by α, β and the covariance matrix $\Sigma_{\mathbf{x}}$.

When the dimension of $\mathbf{x}$ is one (scalar), the MGGD is still applicable and is denoted by univariate generalized Gaussian distribution (UGGD). MGGD is a particular case of the v-spherical distribution defined by Fernández [56]. Using MGGD model, we can derive more specific forms of equation (37). Since we can assume that $\mu = \mathbf{0}$,

$$\nabla g(\mathbf{x}) = -\frac{2\beta}{\alpha^{\beta}} (\mathbf{x}^t \Sigma_{\mathbf{x}}^{-1} \mathbf{x})^{\beta-1} \Sigma_{\mathbf{x}}^{-1} \mathbf{x}. \tag{39}$$

From equations (37) and (39),

$$\hat{\mathbf{x}} = \left(\Sigma_{\hat{\mathbf{x}}} + \frac{2\sigma^2 \beta}{\alpha^{\beta}} (\hat{\mathbf{x}}^t \Sigma_{\hat{\mathbf{x}}}^{-1} \hat{\mathbf{x}})^{\beta-1} I \right)^{-1} \Sigma_{\hat{\mathbf{x}}} \mathbf{y}. \tag{40}$$

To simplify equation (40), we define $q(\hat{\mathbf{x}}) = \hat{\mathbf{x}}^t \Sigma_{\hat{\mathbf{x}}}^{-1} \hat{\mathbf{x}}$. Hence :

$$q(\hat{\mathbf{x}}) = \mathbf{y}^t \left(\Sigma_{\hat{\mathbf{x}}} + \frac{2\sigma^2 \beta \{q(\hat{\mathbf{x}})\}^{\beta-1}}{\alpha^{\beta}} I \right)^{-2} \Sigma_{\hat{\mathbf{x}}} \mathbf{y}. \tag{41}$$

Equations (40) and (41) allow us to solve for $\hat{\mathbf{x}}$.

However, there is no general solution for equation (41). To overcome this problem, we can define a particular condition for α, β and $\Sigma_{\hat{\mathbf{x}}}$ or use a numerical method. In our case, we simply use Newton's method.

3.2 Proposed Color Image Denoising Framework

A color image can be described as a set of multiple image components. The color image representation is based on the trichromatic theory, which reflects the fact that human has three kinds of cones and they absorb light of different wavelengths. This leads to the separation of most of color images into three color components. In this chapter, we consider RGB color space that is an additive color model using three primary colors, i.e. red, green, and blue. We assume color images are composed of these three color channels. In this case we assume that a given noisy image contains randomly distributed Gaussian additive noise for each channel of RGB components. . Unlike trichromatic color models such as RGB space, color models based on luminance and chrominance such as HSI, YUV, YIQ, L*a*b, and so on, have some advantages. For example, HVS is more sensitive to the change of luminance than chrominance. This enables us to decorrelate color components and limit the bandwidth of chrominance channels. Many of these color spaces can separate two uncorrelated parts, luminance and chrominance[1] while RGB color components are highly correlated. By considering these properties of different color models, some works for color image denoising have been proposed [115].

The main difficulty in applying a monochrome image denoising algorithm to color image is the fact that a color image is multi-channel and the channels are correlated to each other. The simplest and straightforward way to apply the denoising algorithms discussed in Section 3.1 to color images is to consider each color channel as a single monochrome image and denoise each channel separately. As one can expect, this approach does not take advantage of any color information and channel correlation.

Another straightforward approach is to take advantages of a single color space that considers HVS such as YUV or YIQ color space. As shown in figure 6 , a simple framework that utilizes YUV color space system can be considered. Since YUV conversion can be linearly performed and there is less correlation between Y

---

[1]Luminance part substitutes intensity, lightness, or brightness depending on the color system. Chrominance information includes hue and saturation.
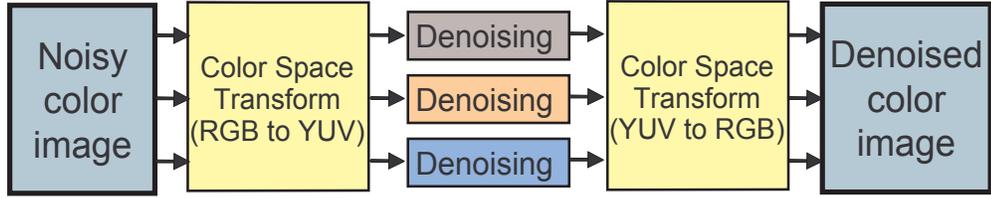
Figure 6: Color image denoising framework using YUV color space conversion

and UV channels, better performance can be achieved. We discuss more details in the following section.

Finally, our proposed color denoising scheme is depicted in figure 7. The framework estimates clean wavelet coefficients by different analyses which can make the color channels decorrelated. The analyses generate multiple estimates and construct probable ranges of clean coefficients in the color vector space. In the limited ranges of the vector space, we can decide an estimated coefficient vector which has the highest *a posteriori*.

3.2.1   Color Space Conversion for Human Visual System

Any physical wavelength of visible light is recognized by three different types of cone cells in human retina. A color that represents light energy of a certain wavelength can be expressed by combining three primary color components such as red, green, and blue. Physically any optical systems including human optic nerves and charge-coupled devices (CCD) acquire these primary colors from the reflected light of an object. Human visual system is more sensitive to luminance information than hue and saturation. Among many color spaces that consider HVS, we mainly use a linear YIQ color conversion from RGB space defined in the following:

$$
\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix},
\tag{42}
$$

31

where $(R, G, B)^t$ denote a color vector of RGB components that are normalized gamma-corrected values and $(Y, I, Q)^t$ is a YIQ color coordinate. Similarly we can use YUV color conversion that is defined in the following:

$$
\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & -0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{43}
$$

where $(Y, U, V)^t$ denotes vector of YUV color channels. YIQ is a standard of NTSC system while PAL uses YUV color space. Both YIQ and YUV color spaces can be converted linearly from RGB space. Due to the linear properties of both the wavelet transform and the conversion to YIQ color space, the conversion to YIQ can therefore be performed in the wavelet domain. We prefer the linear conversion in order to keep the noise properties. In equation (32), we can notice that the noise is still zero-mean Gaussian after a linear transformation. However, the variance for each color channel may vary. In our case, the noise variances for Y, I, Q components become

$$
\sigma_Y^2 = (0.299\sigma_R)^2 + (0.587\sigma_G)^2 + (0.114\sigma_B)^2
$$

$$
\sigma_I^2 = (0.596\sigma_R)^2 + (0.275\sigma_G)^2 + (0.321\sigma_B)^2
$$

$$
\sigma_Q^2 = (0.212\sigma_R)^2 + (0.523\sigma_G)^2 + (0.311\sigma_B)^2
$$

respectively, where σ_R, σ_G, and σ_B are given noise standard deviations for each RGB channel.
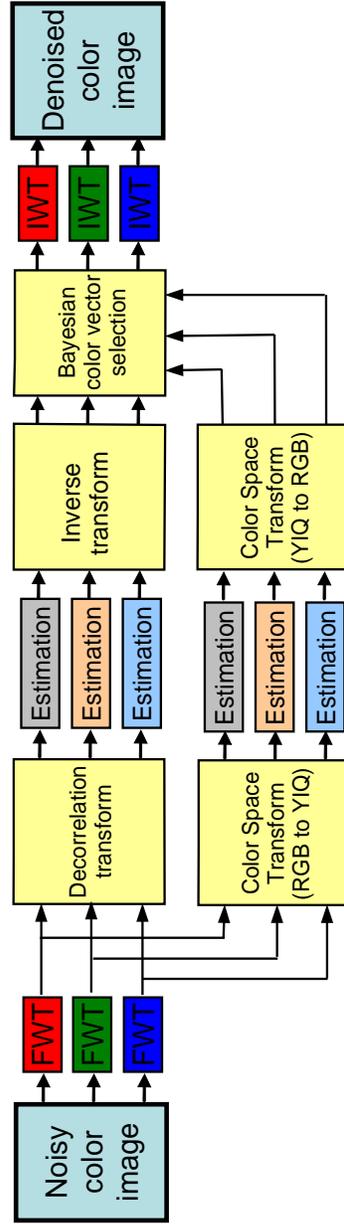
Figure 7: Proposed color image denoising framework using decorrelation-based selective estimation

3.2.2 Decorrelation Approach in the Wavelet Domain

Since RGB color space is strongly correlated for natural images, one can assume that uncorrelated space can be helpful for image denoising. In fact, one of the important inspirations in denoising methods using wavelet transform is drawn from weak correlation property between wavelet coefficients. Rao and Jones developed denoising method for a multisensor [120]. They show the similarity between Karhunen-Loeve (KL) and wavelet transforms and utilize the uncorrelation properties in spatio-temporal manner. KL transform is able to decorrelate strongly correlated multichannel data perfectly. It has been used for image analysis and coding widely. From equation (32), we define one color pixel as a vector, $\mathbf{v} = (v_r, v_g, v_b)^t$. Hence, we can denote a wavelet vector composed of wavelet coefficients from each color component as follows:

$$\mathbf{y} = \mathbf{x} + \Sigma_z^{1/2}\mathbf{z}, \tag{44}$$

where $\Sigma_z^{1/2} = \mathtt{diag}(\sigma_R, \sigma_G, \sigma_B)$. In order to decorrelate the given wavelet vectors, we obtain the KL transform matrix Φ, which satisfies $\Phi^t\Sigma_\mathbf{y}\Phi = \Lambda$, where $\Sigma_\mathbf{y}$ is a covariance matrix of the given wavelet vectors $\mathbf{y}$, Λ is a diagonal matrix containing the eigenvalues of $\Sigma_\mathbf{y}$, and Φ is a eigenmatrix of $\Sigma_\mathbf{y}$ expressed as $\Phi = [a_{ij}]$, where a_{ij} denotes the element of i-th row and j-th column in a 3×3 matrix Φ. They are all 3×3 matrices. Then the transformed vectors

$$\mathbf{y}_{KL} = \Phi^t\mathbf{y} \tag{45}$$

become ideally uncorrelated. In this case, our denoising problem is to estimate $\mathbf{x}_{KL} = \Phi^t\mathbf{x}$. From equations (44) and (45), $\mathbf{y}_{KL} = \mathbf{x}_{KL} + \Phi^t\Sigma_z^{1/2}\mathbf{z}$. Therefore,

$$\sigma_j^2 = (a_{1j}\sigma_R)^2 + (a_{2j}\sigma_G)^2 + (a_{3j}\sigma_B)^2,$$

where σ_j denotes noise standard deviation for a transformed channel j and $j = 1, 2, 3$. Since KL transform is linear like YIQ color conversion, the noise model in the

transformed domain is also Gaussian. Then we can apply a monochrome denoising algorithm without modification for each transformed channel. Since Φ is orthogonal, inverse KL transform can be obtained from equation (45) as follows:

$$\mathbf{y} = \Phi \mathbf{y}_{KL} \tag{46}$$

The main advantage of KL transform is the ideal decorrelation depending on the given data. In our framework, KL transform is performed for each subband in the wavelet domain in order to take maximum advantages of multiresolution analysis property of wavelets. It should be noted that any denoising algorithm using wavelet transform can be applied to each of the estimation procedures for the denoised wavelet coefficients.

3.2.3 Bayesian Selection of Candidate Vectors

Previous sections imply that there might be many ways to estimate the clean wavelet coefficients. We can compose multiple candidate vectors from the estimates and confine the vector space including them. In this section, we follow the approach described in Section 3.1.2 where we assume that the clean wavelet coefficient is an estimate that maximizes the *a posteriori* $p(\mathbf{x}|\mathbf{y})$. If $p(\mathbf{x}|\mathbf{y})$ can be modeled accurately, it is possible to estimate the clean coefficient in a locally confined vector space with less computational cost. Suppose that the power set $2^D = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_m\}$ is constructed from the estimated vector set D obtained by the two different ways at the end of step 4 (estimation blocks) of the framework in figure 7. Then in order to estimate the most probable vector values, we evaluate the following equation for each element vector of

2^D:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in 2^D} \ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x})$$

$$= \arg\max_{\mathbf{x}\in 2^D} \left[-\frac{1}{2}\ln (2\pi)^3 |\Sigma_z| + \gamma - \frac{(\mathbf{y}-\mathbf{x})^t \Sigma_z^{-1}(\mathbf{y}-\mathbf{x})}{2\sigma^2} - \left(\frac{(\mathbf{x}-\mu)^t \Sigma_{\mathbf{x}}^{-1}(\mathbf{x}-\mu)}{\alpha} \right)^\beta \right].$$

$$(47)$$

3.3 Experimental Results

In the experiments, we have used mainly color images with 512×512 sizes from USC-SIPI image database and other public sources. Clean images are assumed to be gamma corrected to satisfy HVS. The noise model we assume is zero-mean additive white Gaussian noise (AWGN). In addition, we also used color images taken by consumer digital cameras to see how the denoising algorithms can enhance the practically used images which include naturally generated noise.

The wavelet filter we have used is dual-tree complex wavelet transform (DT CWT) suggested in [72]. DT CWT has helpful properties for the denoising such as redundancy and directionality. We also have chosen Daubechies' length 8 wavelet filter (DAUB. 8) which is one of the most common mother wavelets for denoising for comparison.

In order to evaluate our method, we show the results using *Wiener*, a straightforward extension in RGB space using *NeighLevel*, and our proposed framework described in the previous sections. Wiener filter is the optimal minimum mean squared error estimator which considers neighboring information. We simply used Matlab implementation of pixelwise adaptive Wiener filtering, `wiener2()` function with given noise (σ) and 5×5 or 7×7 local window for measuring mean and standard deviation. The preliminary estimator for clean wavelet coefficients is *NeighLevel* presented in Section 3.1, in order to investigate how the correlation between color components affects the denoising results. We also implemented a method described in [85], which

takes advantage of chromatic filters with anisotropic diffusion for luminance channel. Tables 1 and 2 shows the denoising results with additive white noise $N(0, 20^2)$ and $N(0, 30^2)$ respectively in terms of peak signal to noise ratio (PSNR). PSNR for color image is defined as follows:

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right),\tag{48}$$

and

$$MSE = \frac{1}{3nm}\sum_{i \in R,G,B}\sum_{y=0}^{n-1}\sum_{x=0}^{m-1}|u(x,y,i) - \hat{u}(x,y,i)|^2,\tag{49}$$

where $u(x,y,i)$ is a clean image pixel on (x,y) spatial coordinate of i-th channel of RGB color components and $\hat{u}$ is an estimated image with size $n \times m$. In figures 8, 9 and 12, some of the results are displayed. Figure 8 shows one example denoising results for different approaches and figure 9 present its zoomed region. The compared denoising results in figure 9 are interesting since figure 9(c) has rainbow-colored artifact while figure 9(d) looks more consistent with the white mountains. Another example is *pepper* image in figure 11 and its cropped one figure 10. This example gives the smallest difference between *NeighLevel* RGB and the proposed (10e and 10f). It's mainly because the original image has relatively good decorrelation between RGB channels. Therefore, independent denoising works better than the other examples. We also tried to enhance an image without *artificial* Gaussian noise in figure 12. As can be seen, our proposed wavelet approach can remove the noise in the image without too much blurring. These results show that we can achieve high-quality color image denoising by decorrelation of color components for most of natural images. However, the denoising performance depends on a given image itself and its statistical properties. For example, *Baboon* image in figure 13 has more random noise-like texture, which is natural and not supposed to be removed. In this case, it's difficult to distinguish texture and noise. Therefore, none of methods give good PSNR although visual impression is acceptable.

(a) Clean image

(b) Noisy image (18.59dB)

(c) Wiener (27.50dB)

(d) Chromatic filters (27.49dB)

(e) *NeighLevel* RGB (29.62dB)

(f) *Proposed* (30.99dB)

Figure 8: Denoising results for *F-16* image with Gaussian noise $N(0, 30^2)$.

Image	Noisy image (dB)	Wiener filter (dB)	Chromatic filters[85] (dB)	NeighLevel RGB (dB)	Proposed
Lena	22.11	29.70	29.54	31.50	32.53
F-16	22.11	29.60	28.91	31.47	32.89
Lake	22.11	27.27	26.94	28.40	29.09
Baboon	22.11	22.99	23.22	25.41	26.87
Pepper	22.11	29.65	28.68	30.61	31.06

Table 1: PSNR values of denoised 512×512 color images with AWGN $N(0, 20^2)$ from USC SIPI database

Image	Noisy image (dB)	Wiener filter (dB)	Chromatic filters[85] (dB)	NeighLevel RGB (dB)	Proposed
Lena	18.59	27.63	27.84	30.01	31.09
F-16	18.59	27.50	27.49	29.62	30.99
Lake	18.59	25.79	25.94	26.79	27.56
Baboon	18.59	22.74	22.90	23.81	24.98
Pepper	18.59	27.67	27.15	29.48	29.72

Table 2: PSNR values of denoised 512×512 color images with AWGN $N(0, 30^2)$ from USC SIPI database

(a) Clean image

(b) Noisy image (18.59dB)

(c) Wiener (27.50dB)

(d) Chromatic filt. (27.49dB)

(e) *NeighLevel* RGB (29.62dB)

(f) *Proposed* (30.99dB)
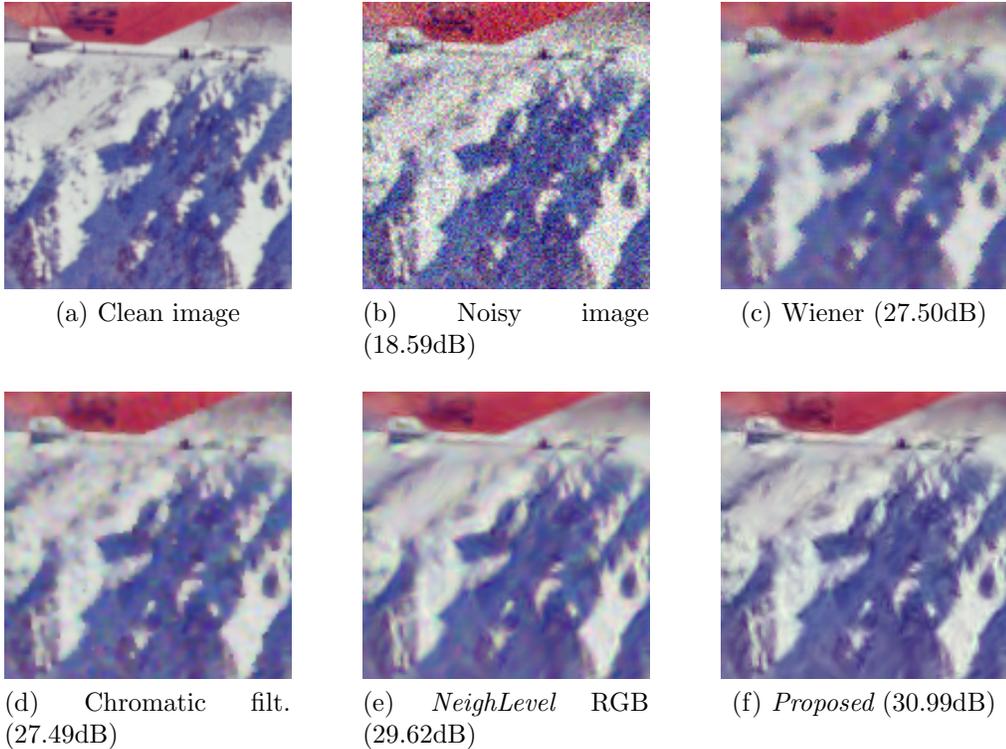
Figure 9: Denoising results for cropped *F-16* image with Gaussian noise $N(0, 30^2)$.

(a) Clean image
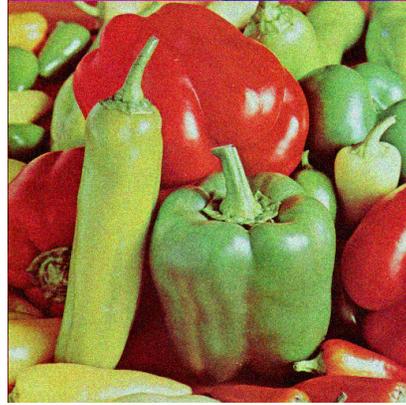
(b) Noisy image (22.11dB)
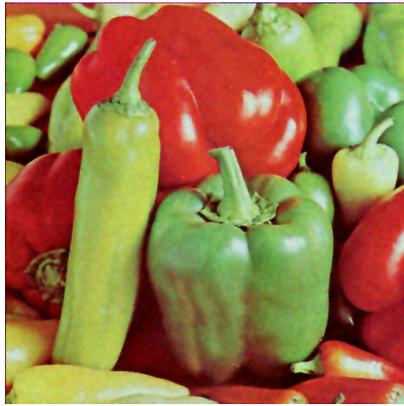
(c) Wiener (29.65dB)

(d) Chromatic filters (28.68dB)

(e) *NeighLevel* RGB (30.61dB)

(f) *Proposed* (31.06dB)

Figure 10: Denoising results for *pepper* image with Gaussian noise $N(0, 20^2)$.

(a) Clean image

(b) Noisy image (22.11dB)

(c) Wiener (29.65dB)

(d) Chromatic filt. (28.68dB)

(e) *NeighLevel* RGB (30.61dB)

(f) *Proposed* (31.06dB)
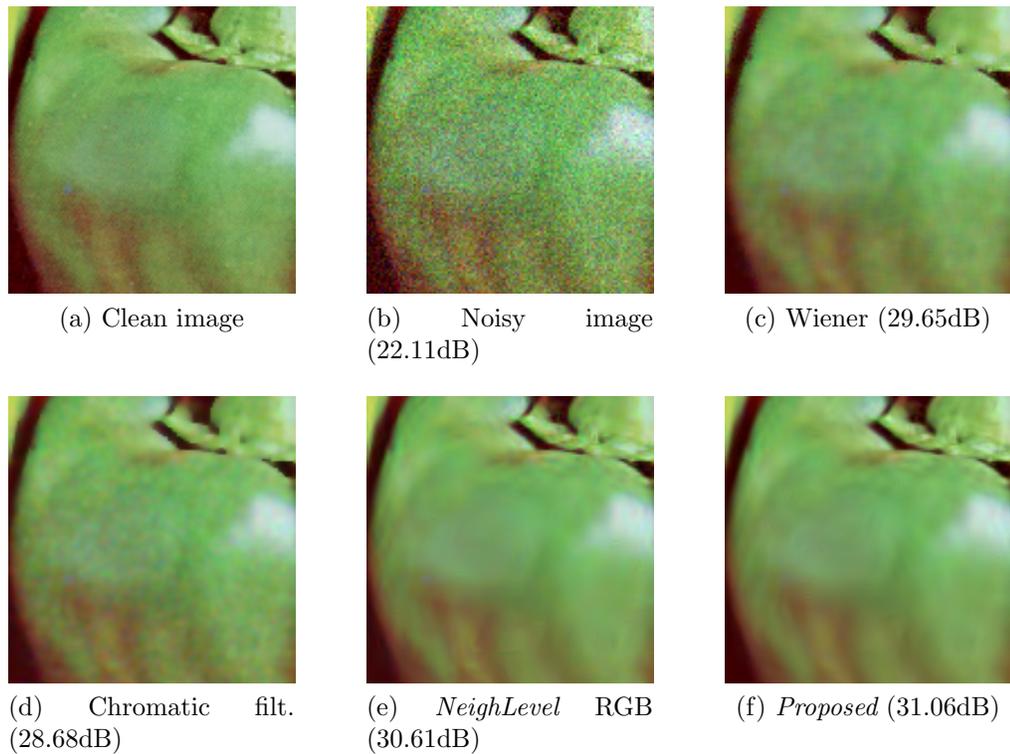
Figure 11: Denoising results for *pepper* image with Gaussian noise $N(0, 20^2)$.

(a) Captured test frame using Sony PD-100

(b) Zoom of (a); original (top-left), Wiener (top-right), chromatic filter [85] (bottom-left), and proposed (bottom-right)
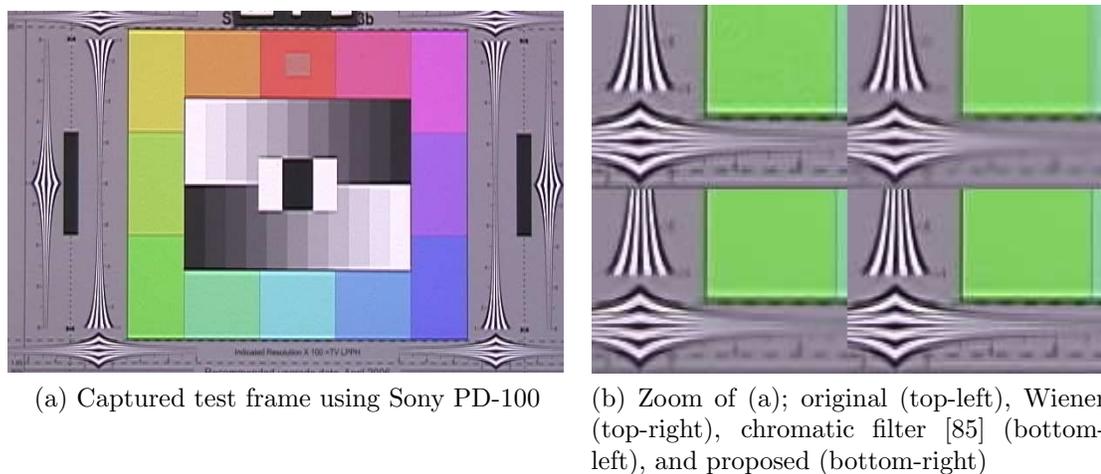
Figure 12: Denoising results of a naturally corrupted image frame captured by consumer digital camcorder
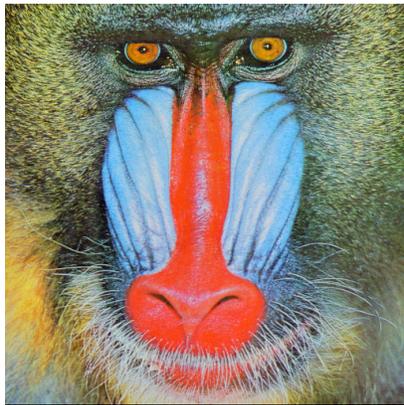
(a) Clean image
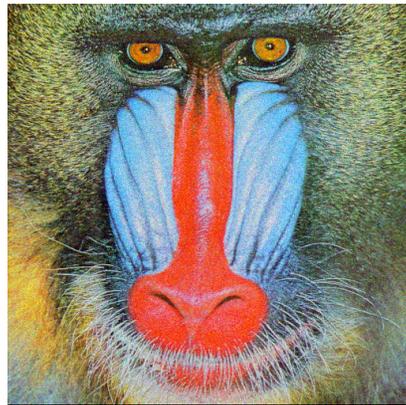
(b) Noisy image (22.11dB)
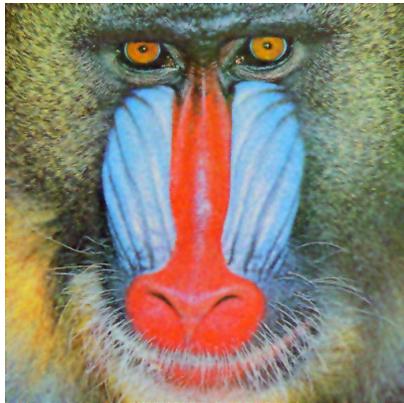
(c) Wiener (22.99dB)
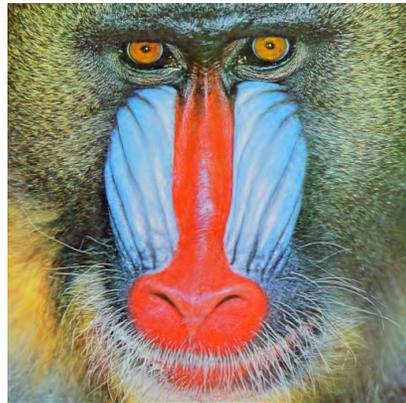
(d) *Proposed* (26.87dB)

Figure 13: Denoising results for *baboon* image with Gaussian noise $N(0, 20^2)$.

3.4 Conclusion

In this chapter, color image denoising framework using wavelet transform has been presented. The correlation between color components must be maximally utilized for an efficient denoising algorithm. The proposed framework is capable of employing any wavelet shrinkage algorithms for monochrome image. In the experiments, we use wavelet shrinkage algorithms that take advantages of intra-dependency of the wavelet coefficients since the dependency between neighboring wavelet coefficients is critical information for image denoising. Our experiments show that the proposed framework is practically flexible and efficient.

Chapter 4

Image Enhancement : Contrast, Brightness, And Color

Natural images can be degraded when they are acquired due to lighting condition, sensor resolution and quality, or limitation or noise of optical system. In addition, most of natural images are compressed with a certain degree of data loss for more efficient storage and communications. An image enhancement algorithm makes such degraded images visually better perceived. It is a fundamental problem in image processing field and somehow subjective since the quality is decided by human visual system (HVS). The enhanced image is expected to have better brightness and contrast, good color consistency, reduced noise or defect, less visual artifacts, or better resolution. Depending on the quality of a given degraded image, each of these improvement factors became an important topic separately in the area such as denoising, contrast enhancement, white balance, deblurring, demosaicking, deblocking, super-resolution, inpainting, and so forth. Among them, comfortable brightness, contrast, and color consistency are major enhancement factors to HVS since photoreceptors in retina (rod and cones) are mainly stimulated differently by light strength levels and wavelengths, and transduce different levels of stimuli to send the analyzed signal to the brain. In this chapter, we will be dealing with image contrast enhancement in the compressed wavelet domain. This to our knowledge is the first attempt in the

direction.

There have been various works along this line. In the spatial domain, there have been many good methods that can be easily found in introductory image processing books.

Retinex-based approaches, one of the successful methods, express HVS-adaptive dynamic ranges with different lighting conditions in the spatial domain [64, 65, 94]. In the compressed discrete cosine transform (DCT) domain that the most popular JPEG image format employs, there have recently been some efficient image enhancement approaches such as [78, 100]. Since they can be applied directly to the encoded data, it is not required to decode the image for image enhancement. These methods also consider the degradation caused by the compression approach such as blocking artifact or related noise.

Wavelet transforms have been used for image contrast enhancement to take advantage of the multi-scale properties in the wavelet domain [147]. Curvelet is also used to preserve better the multi-scale edges[136]. The method proposed in [25] automatically chooses one discrete wavelet filter that can produce the best enhancement result for a given image. These methods are designed to use wavelet transforms to achieve a better image enhancement quality. Therefore, computational efficiency regarding transformation or data transmission was not considered. On the other hand, our method puts an emphasis on the image data processing in the wavelet compressed domain itself. Figure 14 shows a conceptual diagram for the communication process of compressed image data. In the process, image enhancement can be implemented directly in the compressed domain.

JPEG 2000 is a representative standard which can replace the DCT-based JPEG [133]. In JPEG 2000, symmetric filters with biorthogonal properties such as Le Gall 5/3 and CDF 9/7 have been chosen. Unlike orthonormal wavelet filters with some nice mathematical properties, biorthogonal wavelet causes a difference between the scaling and wavelet filters in their filter lengths; However, biorthogonal wavelet makes it possible for simple and fast implementation. More details on wavelet transform have
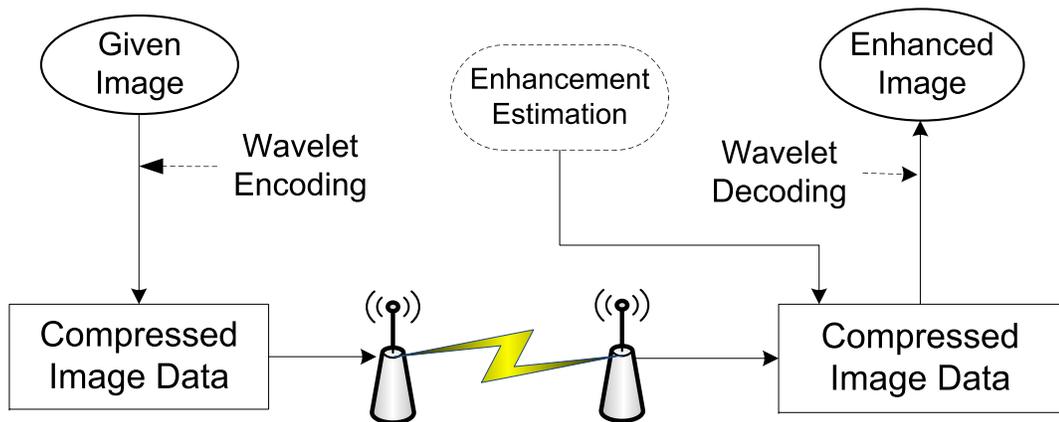
Figure 14: Transmission of image data

been introduced in Section 2.5.

In this chapter, we propose a new, fast image enhancement framework in the compressed wavelet domain, especially for JPEG 2000. The framework is required to estimate the enhanced transformed image data by applying scale factors of scaling and wavelet coefficients. Since each meaningful wavelet coefficient is simply scaled by a proposed formulation, the framework is computationally efficient. In addition, the framework is flexible since the resulting scaling factor for wavelet coefficients is a combination of global scaling enhancement factor for a scaling coefficient, shrinkage factor for noise suppression, and wavelet enhancement factor. We derive the scaling factors by taking into account the multi-scale property and the low- and high-pass representation of the wavelet transforms.

This chapter is organized as follows. In the next section, we review the discrete wavelet transform and its properties for image enhancement problem. Proposed algorithm is described in Section 4.1. Then experimental results and performance evaluation are demonstrated before concluding this work.

4.1 Proposed Wavelet Framework for Image Enhancement

Our proposed approach is performed in the compressed wavelet domain. In other words, the approach does not require any information in the spatial domain and benefits from the wavelet properties. Therefore, some advantages are expected as follows:

- First, computational cost can be tremendously reduced when the image is compressed by the wavelet transform (e.g. JPEG 2000). Actual enhancement process is performed only in the compressed domain.

- Secondly, both low and high frequency information are treated separately. A scaling subband that contains the compact information describes the brightness of the image and each coefficient represents the local brightness. Wavelet subbands are represented as a small number of significant coefficients and a large number of trivial coefficients. In other words, significant coefficients mainly preserve the edge structure while trivial coefficients represent noise or detailed texture. Therefore, it is possible to develop a technique for emphasizing important edges while suppressing noise.

- Finally, the multi-resolution property of wavelet transform can help analyzing images in different scale. Multi-resolutaiton analysis (MRA) makes possible to catch both global and local characteristics of images.

The proposed approach consists mainly of the estimation of scale factors for scaling coefficients and wavelet coefficients of luminance component and their application to chrominance color components. The overview block diagram of the proposed framework is shown in figure 15.
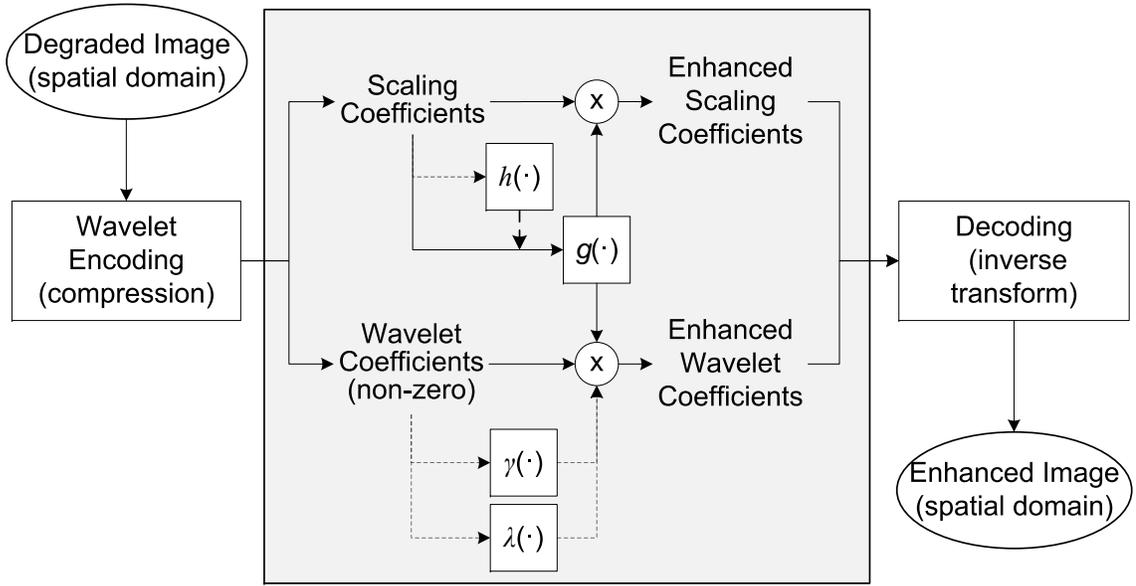
Figure 15: Block diagram of wavelet-based enhancement framework.

4.1.1 Estimation of scaling coefficients

Scaling subband represents overall image structure since it includes a relatively small number of coefficients by a few iterations of decimation. Therefore, this information is closely related to global brightness and contrast of the image. Scaling coefficients do not have exactly the same properties as the intensity values in the spatial domain. For example, statistical indicators such as mean and standard deviation are scaled by applying linear filter iteratively. Bounded set is also changed in the same manner. Let's denote the mean and standard deviation of an image by μ_{im} and σ_{im}. Then their relation to the mean and standard deviation for the L-th level scaling subband, μ_{s_L} and σ_{s_L}, becomes

$$\mu_{s_L} \cong 2^L \mu_{im} \text{ and } \sigma_{s_L} \cong 2^L \sigma_{im} \tag{50}$$

The bound of scaling coefficient is larger than the one in the spatial domain. For instance, the bound by Le Gall 5/3 filter becomes $2^L I_{\max} \left[-\frac{L \cdot 5^{L-1}+1}{4^L}, \frac{5^L+5L}{4^L} \right]$, where $I_{\max}$ is a maximum image intensity value. In practice, most of coefficients belong to the bound $[0, 2^L I_{\max}]$ because there exist weak tails on both sides of the bounds $[\inf, 0]$ and $[2^L I_{\max}, \sup]$. Figure 16 shows the difference between histograms of an image

(a) original image   (b) scaling subband
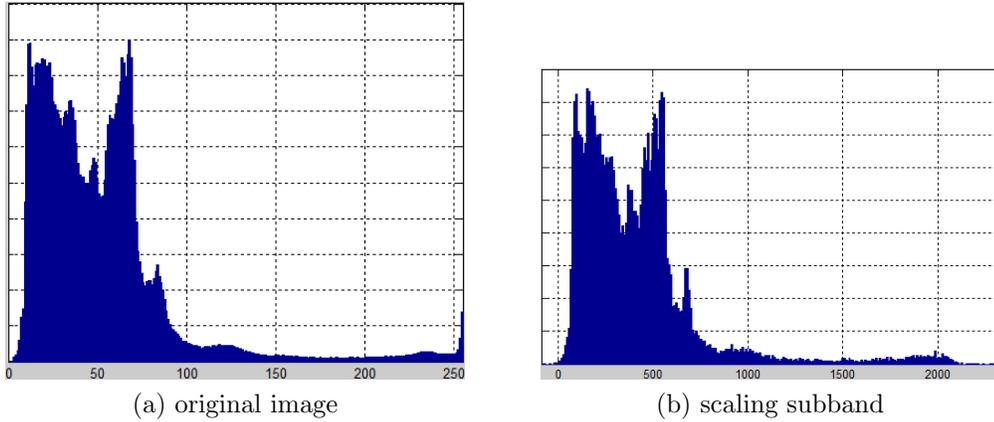
Figure 16: Distributions of image intensity values and its transformed scaling coefficients after three level decompositions with image no. 15 of NASA database.

and its scaling subband. In this case, infimum and supremum should be -2422.5 and 4462.5 respectively after three levels of decompositions. However, more than 99% of scaling coefficients lie between 0 and 2040 (i.e. $2^L I_{\max}$).

Based on the above observations, we can develop contrast enhancement criteria as follows:

$$\hat{s} = 2^L I_{\max} f\left(\frac{s}{2^L I_{\max}}\right) = 2^L I_{\max} f(\bar{s}), \tag{51}$$

where $f(\cdot)$ is a mapping function for a scaling coefficient s and $\bar{s} = \frac{s}{2^L I_{\max}}$. $f(\cdot)$ should be able to adjust local image brightness adaptively from global image structure. In this chapter, two possible ways to decide a mapping function $f(\bar{s})$ for the enhanced scaling coefficient estimation are proposed. The first one is a global mapping function and the other one is a flattening function that levels off the image brightness all over the image regions.

Flexible mapping function

An enhanced scaling coefficient is obtained by multiplying a given coefficient by a proper factor. In other words, equation (51) can be rewritten using the scale factor κ_s as follows:

$$\hat{s} = \kappa_s s = g(\bar{s})s, \tag{52}$$

49

where $g(\cdot)$ is a scaled mapping function, which can be obtained by dividing $f(\cdot)$ by the normalized scaling coefficient s.

The scaled mapping function $g(\bar{s})$ is defined with the following conditions:

1. $\bar{s}$ ranges between 0 and 1.

2. $f(\bar{s}) = \bar{s}g(\bar{s})$

3. $0 \leq f(\bar{s}) \leq 1$.

4. $f(\bar{s})$ is monotonically increasing, i.e. $g(\bar{s}) + \bar{s}g'(\bar{s}) \geq 0$.

If $g(\bar{s})$ is greater than 1, the coefficient s is scaled-up. On the other hand, s becomes scaled-down if $g(\bar{s})$ is less than 1. The mapping functions introduced in [78, 100] has the form of $f(\bar{s})$ and satisfy these conditions. However, it is possible to design a maaping function with great flexibility. In this chapter, we propose the following scaling mapping function:

$$
g(\bar{s}) = \begin{cases} 1 + c_1(m - \bar{s}) \exp\left(-\frac{|\bar{s}-m|^2}{\sigma_1^2}\right), & 0 \leq \bar{s} \leq m \\ 1 + c_2(m - \bar{s}) \exp\left(-\frac{|\bar{s}-m|^2}{\sigma_2^2}\right), & \text{otherwise} \end{cases}, \tag{53}
$$

with the assumption that the above-mentioned conditions are satisfied for the mapping function $g(\bar{s})$; and the parameters should be chosen accordingly. The proposed scaling function is continuous and its shape can be flexibly decided by the parameters. The *balance control parameter*, m, is between 0 and 1. It provides a balanced position for an equal state, i.e. $g(\bar{s}) = 1$. σ_1 and σ_2 are *shape parameters* that determine the shapes of scale-up or down. c_1 and c_2 are *amplitude constants* that decide the maximum and minimum scale factors. The shape flexibility plays an important role in dealing with different statistical properties of images and some examples of equation (53) with different parameters are presented in figure 17. It is important to choose the right parameters depending on the type of images. The parameters could be chosen manually, but we use automatic parametrization based on image statistics, which is described in the following section.
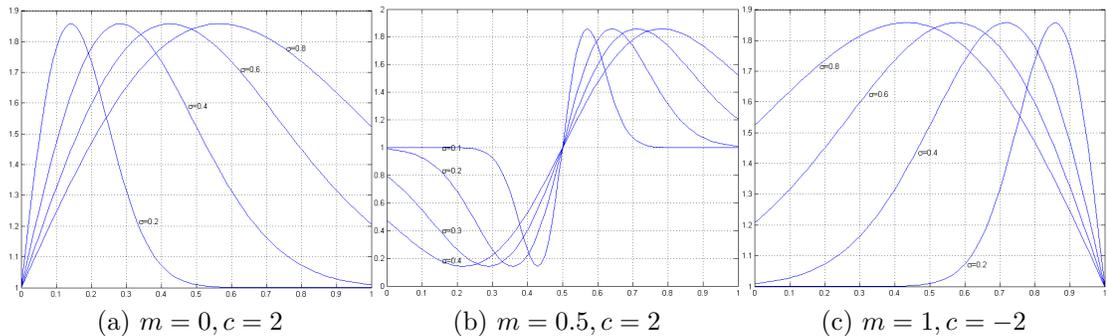
Figure 17: Proposed scale mapping function g.

Automatic parametrization

Natural images can be visually and statistically very different from each other. Figure 18 shows different image statistics on different images. An image could be globally very bright (figure 18a) or very dark (figure 18b). Some regions of image could be very bright while other local regions are dark or just about right (figure 18c). Or an image like figure 18d has narrow dynamic range. Statistically balanced image with higher dynamic range is desirable in many cases. One popular way to achieve this idea is to use histogram equalization. It aims to statistically distribute the image intensities equally through all the ranges by spreading out more frequent intensity values and clustering less frequent ones. However, general histogram equalization is not good for some cases. For example, the brightness of background and foreground objects is very different; or some background noise can be too much emphasized, etc. Although there are some extreme cases, histogram equalization is still useful for many situations.

As shown in figure 16, scaling coefficients have similar statistical distribution as in the spatial domain. In order to get a good global mapping function described in equation (53), we estimate a good parameter set that fit into the ideal statistical model. In this chapter, we assume that the ideal statistical model for scaling coefficients after applying a mapping function is simply a uniform distribution like histogram equalization approach. Parameters are estimated by simulated annealing

(a) Overexposed image

(b) Underexposed image

(c) Locally over- or under-
exposed

(d) Narrow dynamic range
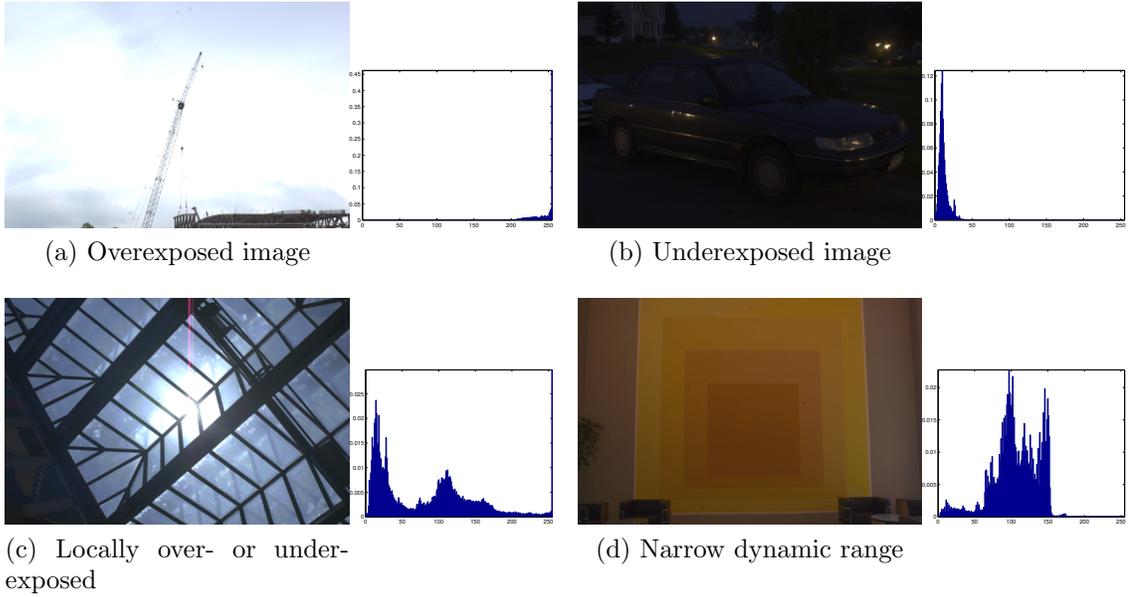
Figure 18: Statistics of global image intensity with different images from RIT MCSL High Dynamic Range Image Database [77]

approach to avoid local minima solution.

Flattening function

The scaling coefficients in the wavelet domain represent the local brightness of an image. HVS feels more comfortable when a local region is neither too dark nor too bright. A local region could have well-balanced brightness by modifying the scaling coefficients to a proper brightness that is comfortable to HVS. Let's assume that the proper mean brightness is a mid intensity value, $\frac{I_{\max}}{2}$, when the image intensity ranges from 0 to $I_{\max}$. In this case, a mapping function is defined by $f(\bar{s}) = 1/2$ in equation (51) and called a *flattening function*. Or a scaled mapping function of $f(\bar{s})$ becomes

$$g(\bar{s}) = \frac{1}{2\bar{s}} \text{ for } 0 < \bar{s} \leq 1. \tag{54}$$

The flattening process strongly affects the overall quality of a reconstructed image since scaling coefficients are condensed information with high entropy. The locality of a scaling coefficient is related to the downsampling factor and the analysis scale and

the decomposition level especially plays an important role to yield visually acceptable balanced image. For example, figure 19 shows visually different results when the number of decomposition levels is changed. The scaling coefficients are scaled by equation (54) and the wavelet coefficients are simply scaled based on the upsampled scales of the scaling coefficients. In this example, the low decomposition level like figure 19a gives high brightness level over all the regions, but large objects are not distinguishable well (see for example the plane versus its background or the sky outside versus the floor). On the other hand, high decomposition levels produce higher contrast as seen in figure 19f where the sky is too bright and the floor is relatively dark.

Discontinuities of downsampled scaling coefficients

Unlike spatial domain, changes in the scaling coefficients profoundly affect all the other coefficients and overall image brightness since the scaling coefficients are obtained by downsampling spatial image data by 2 iteratively. When L decompositions are applied to an image, $1/4^L$ coefficients of input data are preserved in the scaling subband. Therefore, there is much less spatial continuity between neighboring coefficients. In addition, a small number of scaling coefficients contains significant information of an image. The wrong estimation of scaling coefficients may result in serious image degradation.

Direct multiplication of the enhanced scale factor to the downsampled discrete wavelet coefficients could result in spatial distortion and ringing artifacts after inverse transform. In fact, independent estimation of each scaling coefficient by global mapping function without considering neighbors in the scaling subband easily produces stains on the image. Such visual artifacts can be suppressed by guaranteeing smooth intradependency. We propose to consider neighboring dependency of scaling coefficients and reduce discontinuity by applying smoothing operation. equation (52) can be rewritten as follows:

$$\hat{s} = g\left(\frac{h(s)}{2^L I_{\max}}\right) s, \tag{55}$$

53
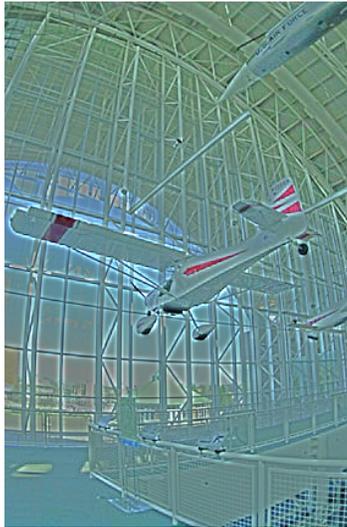
(a) Level 1   (b) Level 2   (c) Level 3

(d) Level 4   (e) Level 5   (f) Level 7

Figure 19: Flattening mapping function and decomposition levels of wavelet transform.

| (a) original image | (b) Non-smooth | (c) Smooth |

Figure 20: Behavior of smoothing filter h.

where $h(s)$ is a smoothing convolution filter such as Gaussian. Figure 20 shows example enhancement results with and without $h(s)$. The example without any smoothing operation shows more severe ringing artifacts, especially around the edges such as borders between lighthouse and sky.

4.1.2 Estimation of wavelet coefficients

While global image brightness and contrast are improved by properly estimating scale factors of scaling coefficients as described in the previous section, it is also important to estimate appropriate enhanced wavelet coefficients to obtain clear details such as edges and textures for better local contrast and sharpness of image. One way is to apply the scale factor of scaling coefficient to its spatially corresponding wavelet coefficient (see equation (55)). In this case, image edges, details, and dynamic ranges are emphasized or diminished in the same way as the scale factor. However, they could be too coarse or easily distorted especially where the number of decomposition levels is increasing because the number of scale factors for the scaling coefficients is much less than the number of wavelet coefficients. Therefore a scale factor for each wavelet coefficient needs to be considered independently. In addition, noise suppression could be considered simultaneously during the estimation since wavelet domain contains most of the noise in an image. Our proposed scale factor for wavelet coefficient is derived using all the above-mentioned elements.

Wavelet scale factor

A wavelet transform used in JPEG 2000 produces one scaling subband in the coarsest level and three wavelet subbands for each level, i.e. there exist $3L$ subbands for L-level decompositions. Each coarse-to-fine level is related to the retinex with different scale (see Section 4.1.2 for more discussion). Since wavelet transform preserves locality, each wavelet coefficient contains edge and detailed information of the corresponding spatial location. Therefore, enhancement of wavelet coefficients can yield a spatially enhanced image with better contrast and sharpness.

The proposed wavelet coefficient estimation is obtained by a scale factor (κ_w) that is composed of mainly three sub-factor terms defined as follows:

$$\hat{w} = \kappa_w w = g_w \gamma_w \lambda_w w, \tag{56}$$

where g_w is a scale factor for the scaling coefficient, γ_w is a locality factor, and λ_w is a shrinkage factor.

A scale factor for scaling coefficient g_w is computed based on equation (55). This term represents that the energy of the details (wavelet coefficient) is proportional to the brightness level (scaling coefficient). Unbalanced modification of scaling and wavelet coefficients causes necessarily unbalanced enhancement of brightness and contrast as well as annoying visual artifacts. Since scaling coefficient is usually down-sampled in discrete wavelet transform and accordingly very coarse, hence g_w is obtained by spatial interpolation of κ_s or $g(\bar{s})$ in equation (51).

Second, a locality factor γ_w is obtained from wavelet coefficients. This factor directly contributes to deblurring or contrast enhancement. Since scale factor for scaling coefficient is obtained from decimated information, it is not enough to reflect a local information for a specific wavelet coefficient. Also, high-frequency information including edges, textures and other details in image is only available in the wavelet domain. γ_w can be defined by a mapping function in equation (53) used for scaling coefficient estimation. Other examples of choice include the mapping functions

proposed in [136] and [25].

The third term is a shrinkage factor λ_w for noise removal. Wavelet domain mainly preserves detailed information of an image such as edges, texture, and even noise. Noise is mainly expressed in small wavelet coefficients. The shrinkage term is helpful to suppress noise emphasized by the scale factor g_w as well as originally prevalent noise. Efficiency of noise reduction in the wavelet domain has already been proven. In fact, shrinkage rule of wavelet coefficients has been proven to be efficient for denoising problem. λ_w should be between 0 and 1. In this chapter, we adopt a technique used in [29].

Retinex and wavelet

Retinex formulation in [64] is defined as $R = \log I(x, y) - \log [F(x, y) * I(x, y)]$, where $F(x, y)$ is a surround function with upper-concave shape such as Gaussian filter. Since the formulation is a special case of the difference between two logged convolved images, it can be rewritten as

$$R = \log [F_1(x, y) * I(x, y)] - \log [F_2(x, y) * I(x, y)], \qquad (57)$$

where F_1 is Dirac delta function. On the other hand, wavelet function is often derived from the Laplacian. For example, Mexican hat wavelet is equivalent to Laplacian of Gaussian and can be approximated as Difference of Gaussian (DoG), i.e.

$$\Delta G * I(x, y) \approx G_{\sigma_1}(x, y) * I(x, y) - G_{\sigma_2}(x, y) * I(x, y). \qquad (58)$$

It can be easily noticed that the only difference between equations (57) and (58) is the logarithmic function that is considered for shunting inhibition in Retinex formulation. This reflects that wavelet coefficients contain information related to locally adaptive contrast level for HVS stimulation. In other words, efficient enhancement in the wavelet domain will result in visual improvement that can satisfy HVS.

57

4.1.3 Color treatment

Wavelet transform of JPEG 2000 is applied to each color channel directly. Color space used in JPEG 2000 is usually either RGB or YCbCr. While RGB channels are strongly correlated to each other, luminance channel is less correlated to chrominance channels in YCbCr space. In practice, chrominance channels are obtained by conversion matrix from RGB channels and preserve the edge structures [100]. Instead of computing scale factors for each channel, we apply the scale factors of luminance channel computed in equation (52) and (56) to the corresponding coefficients in the chromatic channels for better color consistency.

4.2 Experimental Results and Performance Evaluation

In this section, the proposed image enhancement approach is applied to various images and we observe how it behaves. The experiments are performed based on the following consideration:

- First, the influence and behavior of scale factors introduced in the proposed approach are observed.
- The effects of different parameters are related to the enhanced image quality.
- All the experimental results are evaluated by visual observation and some numerical measurements.
- Finally, the proposed method is compared to the other state-of-the-art approaches.

Measures   Unlike some other image processing problems such as compression and denoising, there is no standard measure to quantify the degree of enhanced image quality because a good enhanced image is a subjective matter and fully depends on

HVS. However, there exist some quantitative measures that present image brightness, edge contrast, or statistical distribution. These measures should be used only for reference since they usually reflect one aspect of image characteristics rather than showing satisfaction of HVS. Final quality judgment should be performed visually by human observer because the goal of image enhancement is to satisfy HVS. In this chapter, image quality is quantified by calculating mean brightness, mean contrast [93] and absolute mean brightness error (AMBE) [23], and AME [1]. Mean brightness (MB) is an indication of overall image brightness and defined as a global mean intensity. The desirable MB value is a mid-intensity value, $(I_{min} + I_{max})/2$ (e.g. 127.5 for 8-bit gray scale image). However, MB is mainly regarded as an indicator to identify whether the image is too dark or too bright rather than using it as a reliable quality measure because it depends on the image content and whether HVS is efficiently adapted to brightness. Mean contrast (MC) is defined as follows:

$$MC = \sum \frac{u(x,y) - E(x,y)}{u(x,y) + E(x,y)}, \tag{59}$$

where $E(x,y) = \frac{\sum_{k \in N(x,y)} e_k u_k}{\sum_{k \in N(x,y)} e_k}$, $u(x,y)$ is a pixel intensity, $e(x,y)$ is an edge magnitude that can be obtained by Sobel operator, and $N(x,y)$ is neighboring pixels around (x,y) pixel coordinate. Usually if the value is higher, the image quality is better. MC represents image contrast and clear edge quality relatively well. But it should be noted that too much contrast and sharpness could be uncomfortable to HVS. Absolute mean brightness error (AMBE) describes brightness preservation. It is defined as $AMBE = |E(X) - E(Y)|$, where X is a given image and Y is output image by histogram equalization. $E(Y) = (X_m + X_G)/2$ if X_m is mean of image X and $X_G = (I_{\min} + I_{\max})/2$, where $I_{\min}$ and $I_{\max}$ are lowest and highest luminance values respectively. Lower AMBE means better brightness preservation. AME is defined based on HVS by combining Weber's Contrast Law and modulation-based Michelson's Contrast Law.

The AME formulation is denoted by

$$AME = \underbrace{\max}_{\Phi,\alpha} AME_{\alpha,k_1,k_2}(\Phi). \tag{60}$$

$AME_{\alpha,k_1,k_2}(\Phi) = \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} \alpha \left[\frac{I_{\max;k,l} - I_{\min;k,l}}{I_{\max;k,l} + I_{\min;k,l} + c} \right]^{\alpha} \times \log \frac{I_{\max;k,l} - I_{\min;k,l}}{I_{\max;k,l} + I_{\min;k,l} + c}$, where Φ is a given transform from class of fast unitary transform, $I_{\min}$ and $I_{\max}$ are minimum and maximum luminance values respectively when an image I is split into $k_1 \times k_2$ blocks.

Example images Our experiments have been mainly done with NASA image database [11] since these images are widely used by many researchers for comparison.

Evaluation In order to evaluate the proposed algorithm, several other approaches are employed for comparative studies. Among many existing approaches, we have chosen McCann's retinex-based algorithm [94], color enhancement by scaling (CES) [100], automatic wavelet base selection (AWBS) [25], and Curvelet-based approaches [136] have been considered. A retinex-based approach is one of the most popular image enhancement methods in the spatial domain. We use Matlab implementation for McCann's algorithm that was implemented and contained in [58]. CES algorithm is performed in the compressed DCT domain used in JPEG image format and the enhancement estimation is done by scaling the DCT coefficients. Also some works using wavelet transforms have been considered, i.e. AWBS and Curvelet.

We have used Le Gall 5/3 filters to produce compressed scaling and wavelet domains.

Figure 21 shows one of the experimental results. It can be seen that the proposed image in figure 21c is visually comparable to figure 21d produced by CES method. Contrast level can be changed by mapping function (e.g. constant parameter c in equation (53)). Unlike DCT-based approaches, a wavelet-based approach does not

	MB	MC	AMBE	Time(s)
Original	83.41	.0565	12.39	
Proposed	123.56	**.1007**	**4.76**	**0.44**
Retinex McCann[94]	173.69	.0277	14.78	0.49
CES[100]	**124.47**	.0532	5.64	2.79
Curvelet[136]	83.25	.0662	12.10	34.44
Wave. Sel.[25]	82.85	.0929	11.61	181.00

Table 3: Quality measures for *image10*

	MB	MC	AMBE	Time(s)
Original	54.96	.0611	31.19	
Proposed	**111.59**	**.1240**	**2.08**	**0.14**
Retinex McCann[94]	204.05	.0158	23.86	0.34
CES[100]	94.27	.0606	9.72	1.72
Curvelet[136]	54.98	.0712	25.51	34.41
Wave. Sel.[25]	55.13	.1065	30.37	169.01

Table 4: Quality measures for *image18*

cause blocking artifacts although there could be some ringing artifacts caused by wrongly estimated wavelet coefficients.

Figure 22 shows enhancement results obtained by the proposed approach and other existing approaches. It can be shown that the proposed algorithm preserves local details comparatively better in addition to improving the overall contrast and brightness. Quality measure and elapsed time for the experiments can be found in Tables 3 and 4. Table 3 is for *image 10* and Table 4 for *image 18* of the NASA database. The elapsed time is only for reference purpose since implementation is made with Matlab which does not reflect the optimization of the codes.

Noisy image enhancement Figure 23 presents the role of shrinkage term γ_w in equation (56). One simple shrinkage function can be considered as follows:

$$\gamma_w = \frac{\sigma^2}{\sigma^2 + \sigma_x^2},\tag{61}$$

where σ is noise standard deviation and σ_x is standard deviation of the clean image. σ can be estimated by using $\hat{\sigma} = median(|HH_1|)/0.6745$ as proposed in [39].

(a) original image

(b) McCann[94]

(c) Proposed

(d) TW-CES-BLK[100]

Figure 21: Image enhancement results (*image15*).

(a) original image   (b) McCann[94]   (c) TW-CES-BLK[100]

(d) Curvelet[136]   (e) Wave. Sel.[25]   (f) Proposed

Figure 22: Image enhancement results (*image10*).

(a) original image

(b) without denoising term

(c) with denoising term

Figure 23: Influence of denoising term γ_w.

If we assume independency of noise and signal, $\hat{\sigma}_x = \sqrt{\max\left(\sigma_y^2 - \sigma^2, 0\right)}$, where $median(|HH_1|)$ denotes the median of the absolute values in the finest wavelet subband in diagonal direction and σ_y is standard deviation from a given noisy image. In figure 23, Gaussian random noise with $\sigma = 5$ has been added in order to see how the shrinkage term affects image enhancement result. The magnified results of the top-right region of the image show clear difference of applying the shrinkage term (images in the right column of figure 23).

Computational complexity One of the main advantages of the proposed framework is to make fast computations possible especially when the image is encoded using wavelet transform. Assume that a given image size is $N \times M$, the size of scaling subband is $N_1 \times M_1$, and there exist N_2 non-zero wavelet coefficients. Then it takes $O((k_1^2 + 1)N_1M_1)$ for the estimation and application of scale factor to a scaling subband since it takes $O(k_1^2 N_1 M_1)$ for the smoothing operation and $O(N_1 M_1)$ when $k_1 \times k_1$ is a local window size for a smoothing convolution $h(\cdot)$. On the other hand, computation of the enhanced wavelet coefficients requires $O((k_2^2 + 5)N_2)$ since g_w, γ_w, and λ_w can be computed in $O(4N_2)$, $O(N_2)$, and $O(k_2^2 N_2)$ when the bilinear interpolation is applied to get g_w and locally adaptive shrinkage estimation for λ_w is employed. The computational cost could vary since a different approach for each element of the scale factors could be applied. Therefore, overall computational cost requires $O((k_1^2+1)N_1M_1+(k_2^2+4)N_2) \approx O(N_1M_1+N_2)$ when k_1 and k_2 are constants. Considering that $NM > N_1M_1 + N2$, the overall computation is basically linear to the size of the compressed coefficients instead of a full image size.

4.3 Conclusion

In this chapter, a novel image enhancement algorithm in the compressed wavelet domain has been presented. The proposed approach is simple and computationally efficient since the estimation is only performed in the compressed wavelet domain and the estimated coefficients are linearly scaled. The enhanced images produced by the proposed approach are both visually and numerically encouraging.

Chapter 5

Image Enhancement : Inpainting

Inpainting is an artistic term describing a procedure to restore a damaged painting or picture such as medieval artwork and old pictures. It is also called *retouching*. This delicate artistic activity motivated researchers to imitate the procedure automatically. Image inpainting as an image enhancement problem can be defined in general to estimate and fill automatically predefined unknown region (i.e. normally spatially connected pixels) in an incomplete or damaged image. It is also called *image completion*, *disocclusion*, *object removal*, or some other names depending on the application and its purposes. Unlike other image enhancement problems such as image denoising and contrast enhancement problems we have discussed in the previous chapters, image inpainting is a high-level computer vision problem since the solution should be meaningful and satisfactory to human perception.

In this chapter, we exploit automatic image inpainting problem. In the first section, psychological meaning of image inpainting is discussed. Literature review on inpainting and related topics are followed. Many existing approaches are based on mathematical tools such as PDE and variational models or some methods considering texture recovery. Some related applications such as texture synthesis, block recovery and film restoration are also introduced. In Section 5.3, an approach based on Mumford-Shah model, hierarchical level set and texture mapping is proposed. Section 5.4 presents a wavelet-based image inpainting algorithm that takes advantages

of image decomposition and multi-resolution analysis.

5.1   Human Vision Perception and Inpainting

A human brain has an amazing ability to analyze and recognize what we see either
consciously or unconsciously. Kanizsa in his seminal book [69] described how human
visual perception occurs. Perception implicates the entire process from sensory input
to related mental analysis of given information. In other words, *vision* and *thought*
cannot be separated.

Inpainting process reflects *perceptual interpolation*, which implies both primary
and secondary processes. In the first phase of perception, the vision system (e.g.
eyes) receives a set of 2D images from real 3D scene. Then human brain analyzes the
given information obtained from the image sensor (e.g. rod and cones). In this phase,
we need two procedures: 1) detecting damaged regions and 2) filling in the detected
regions. If we ignore the first procedure, the second one is sometimes referred to as
perceptual interpolation. The brain goes beyond the information obtained from the
first phase. In Gestalt psychology[1], this procedure is elucidated as phenomena of
completion or totalization. For example, when we see figure 24(a), a triangle around
the center is amodally perceived although there is no edge. Then we are tempted
to fill in the triangular region by removing the front triangle as if it occludes the
other objects and estimating the occluded contours and surfaces. In addition, most
of us think there are three small dark objects and a large contour object filled white.
It is hard to imagine the occluded small objects are connected to each other. This
procedure is demonstrated in figure 24(c)–(d).

Perceptual interpolation and analysis of human visual system are related to many
topics in image processing, pattern recognition and computer vision: object re-
moval, disocclusion, block recovery, texture synthesis, image interpolation, and super-
resolution to only name a few. They are all related substantially, but each application

---

[1]References on fundamentals of Gestalt theory or its application to computer vision can be found
in [69, 79, 148, 37].
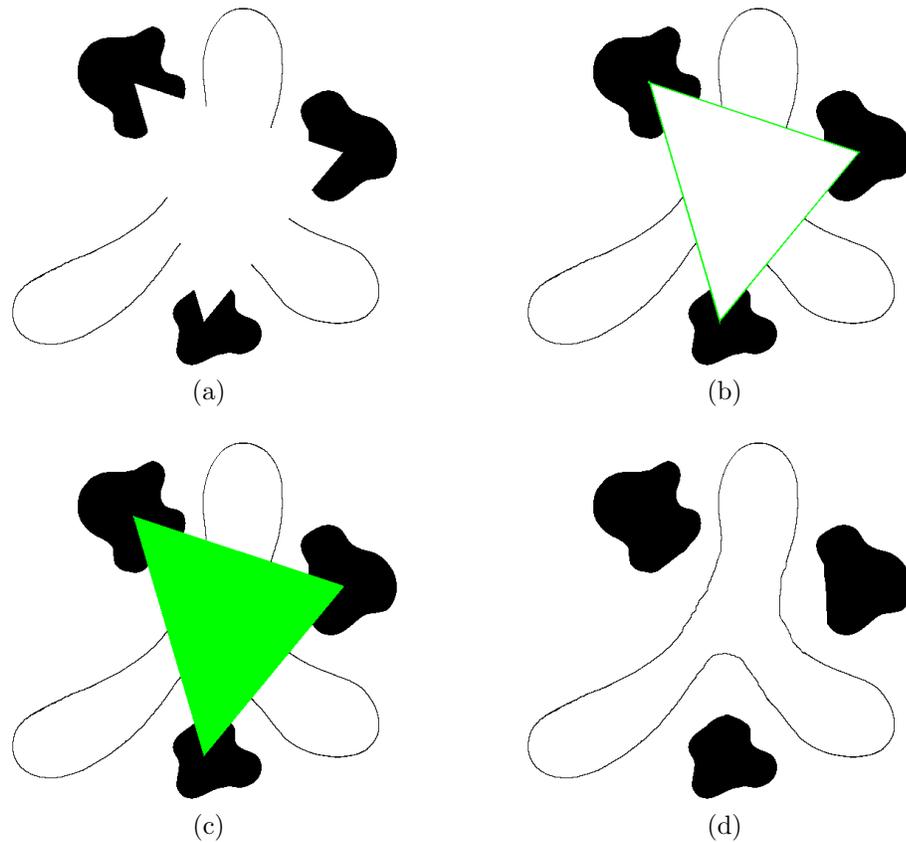
Figure 24: Tendency toward completion of gap for anomalous contours and objects [69]

has slightly different problems. For example, image interpolation usually requires to fill in the regularly spread small unknown regions using regularly spread known information all over the image. On the other hand, disocclusion or object removal requires to estimate one or a few unknown large regions. Also, narrow application of image inpainting usually considers small damaged areas or scratches rather than a large fat region. These different applications result in different approaches.

A fundamental problem in computer vision is concerned with the properties of edges as perceived by human being. Scientists and psychologists have tried to model visual perception mathematically. One of the most challenging problems in image inpainting is to estimate *reasonable* edges which convince human visual system. When we estimate edges in inpainting domain by perceptual interpolation, there are some

(a) Similarity   (b) Proximity   (c) Good continuation

(d) Closure   (e) Symmetry   (f) Periodicity
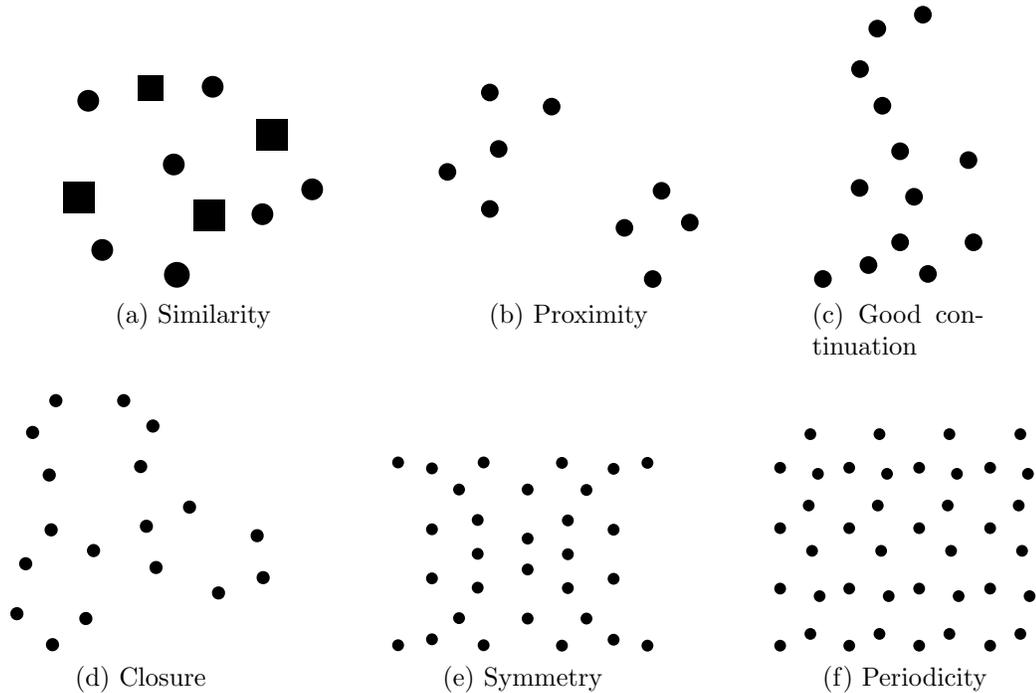
Figure 25: Examples of some of Wertheimer's laws of perceptual grouping [79]

principles in order to satisfy human visual system.

Low-level inpainting can be done by considering local *continuity* and *smoothness*. In this case, the unknown area can be filled by locally looking at the surrounding known image pixels without considering image content. On the other hand, high-level inpainting must consider *reasonable* situation using human visual experiences. Reasonable estimate can be obtained by analogical inference from other examples based on our previous experiences. For instance, we can estimate an occluded object more accurately if we have seen it before. Wertheimer's laws of perceptual grouping including *similarity*, *proximity* or good *continuation*, *closure*, *symmetry*, and *periodicity* are some useful principles for perception of the visual scene [79]. Figure 25 depicts some of the Wertheimer's laws. These grouping laws could be helpful to estimate an occluded region using the properties of other given regions. When we only use global information of a given image to solve inpainting problem, we call it middle-level inpainting and some Gestalt perceptual laws by Wertheimer could be applied.
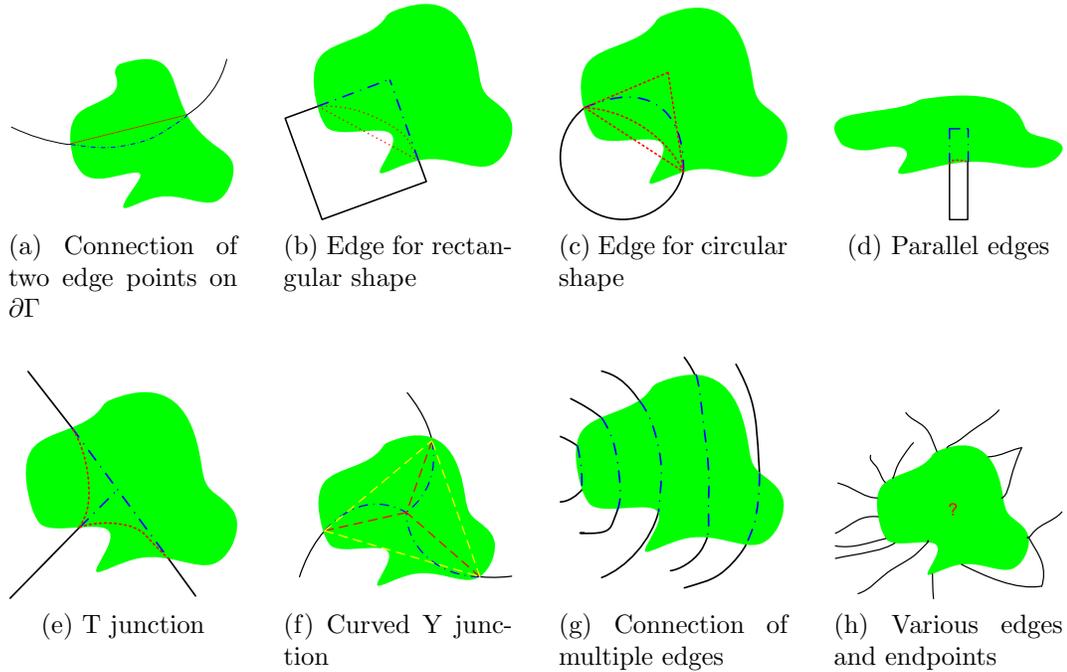
Figure 26: Various cases for perceptual edge estimation: green region is interpolated region and possible estimation for edge is displayed in different colors (blue dash-dot line is most probable estimation)

Now let us consider different examples of simple inpainting problems that only involve strong edge structure without considering any details as shown in figure 26. In figure 26(a), two given endpoints on the boundary of inpainting domain ($\partial \Gamma$) are to be connected to construct one continuous edge (continuity) and the estimated edge must be curvilinear similar to the dashed lines (smoothness). It should be noted that there is no unique estimate in this case. The estimate is acceptable as far as human eyes feel comfortable. In this simple case, there are not many reasonable choices and low-level inpainting is enough for a possible solution. In figures 26(b), (c) and (d), we can also consider more difficult situations with only two endpoints. In these cases, more possible solutions are available compared to figure 26(a) when only low-level inpainting principle is considered. In fact, middle-level or high-level inpainting principle (e.g. the edge estimation using given shapes such as square and circle) should be considered. Based on the given information, we may choose the most probable estimate such as the blue dash-dot lines. This can be explained as a

similarity property in Gestalt principle. However, these estimates could be changed if some other information exists. There are even more complicated examples such as the 'T' junction, or 'Y' junction, and estimations for multiple edges as shown in figures 26(e)–(h). These examples do not have only one solution. There could be many possible solutions if they are meaningful and satisfactory to human visual system.

5.2   Review on Image Inpainting

For the last decade, various image inpainting approaches have been proposed. In this section, different approaches for inpainting problem are discussed.

5.2.1   Image Inpainting Using Mathematical Models

In Chapter 2, we have reviewed the mathematical models for image enhancement. Many existing image inpainting approaches including the early works use these models to solve an inpainting problem. This section briefly review those existing inpainting methods.

PDE-based Approaches

In [7], Bertalmio et. al observed the underlying methodology performed by a restoration artist and proposed a PDE model to express the relationship between edge propagation direction and measure of change in the propagated information by time variations. Let u_0 be a given corrupted image with inpainting domain Γ, the region to be inpainted, and its boundary $\partial\Gamma$, and u_n be a resulting image after n iterations. Then an inpainting algorithm can be expressed as an evolution equation in the following:

$$u_{n+1} = u_n + \Delta t \frac{\partial u_n}{\partial t} \text{ in } \Gamma, \tag{62}$$

$$\frac{\partial u_n}{\partial t} = \overrightarrow{\delta L}_n \cdot \overrightarrow{N}_n = \nabla L_n \cdot \overrightarrow{N}_n \cong \nabla(\Delta u_n) \cdot \nabla^\perp u_n, \tag{63}$$

where L_n is the information to be propagated, $\vec{N}_n$ the propagation direction, and $\overrightarrow{\delta L_n}$ measure of the change in L_n. In this case, the iteration stops at steady state, i.e. $u_{n+1} - u_n \to 0$ or $\overrightarrow{\delta L_n} \cdot \vec{N}_n \to 0$. Also a diffusion process (e.g. anisotropic diffusion[111]) is performed after a few iterations in equation (62) in order to avoid an error in calculation of the isophote direction caused by noise. If we consider additional term for anisotropic diffusion, equation (63) can be rewritten as follows:

$$\frac{\partial u}{\partial t} = \nabla(\Delta u) \cdot \nabla^\perp u + \nu \nabla \cdot (c(|\nabla u|)\nabla u), \tag{64}$$

where $c(\cdot)$ is a diffusion coefficient function (usually monotonically decreasing function). Equation (64) is also explained by Navier-Stokes equation (16) in Section 2.2.2.

Another method in [6] derives a third-order optimal PDE that expresses local neighborhoods and ensures continuation of level lines. In [4], global heat transfer principle is decomposed into basic laws and their numerical scheme is developed both for inpainting and denoising instead of directly solving a PDE.

Variational Approaches

Total variation introduced in Section 2.3.1 can be considered to solve inpainting problem as proposed by Chan and Shen [18]. Let Ω and Γ denote the image domain and the inpainting domain respectively. Then TV minimizer for image inpainting can be modeled in the following energy functional

$$F_{TV}(u) \;=\; \int_{\Omega \setminus \Gamma} |u - u_0|^2 \, dxdy + \lambda \int_\Omega |\nabla u| \, dxdy \tag{65}$$

λ must be 0 in the inpainting domain Γ.

Mumford-Shah model can be considered for edge estimation of unknown inpainting domain, which is the critical information for image inpainting problem. Based on equation (20), inpainting formulation of Mumford-Shah functional can be expressed

as follows:

$$F_{MS}(u, C) = \int_{\Omega \setminus \Gamma} (u - u_0)^2 dx dy + \mu \int_{\Omega \setminus C} |\nabla u|^2 dx dy + \nu |C|. \qquad (66)$$

The above equation indicates that only the variance of the image and the length of the segmentation curve are considered inside the inpainting area.

In order to minimize Euler elastica model given by equation (18) for all level lines, Chan et al. proposed the following minimization functional in [17]:

$$F_{EE}(u) = \int_0^1 P(X_\lambda) d\lambda = \int_0^1 \int_{X_\lambda} (\alpha \kappa^2 + \beta) ds d\lambda = \int_\Gamma (\alpha \kappa^2 + \beta) |\nabla u| ds, \qquad (67)$$

where $P(X_\lambda)$ is the Euler elastica model for a given level line X_λ and $\frac{d\lambda}{dt} = |\nabla u|$ or $d\lambda = |\nabla u| dt$.

In [51], an improved model which combines the advantage of Mumford-Shah and Euler (MSE) elastica models has been proposed, called Mumford-Shah-Euler functional,

$$F_{MSE}(u, C | \Gamma) = \int_{\Omega \setminus \Gamma} (u - u_0)^2 dx dy + \lambda \int_{\Omega \setminus C} |\nabla u|^2 dx dy + \int_C (\alpha + \beta \kappa^2) ds. \qquad (68)$$

In this case, the length term in equation (66) is included in the Euler elastica curve model (i.e. $\int_C (\alpha + \beta \kappa^2) ds = \alpha |C| + \beta \int_C \kappa^2 ds$).

Bayesian Framework and Variational Approaches

As shown in Section 2.4, Bayesian framework has very close relationship to variational approach. For inpainting problem, inpainting domain Γ is considered. Then equation (4) in Chapter 2 can be rewritten as follows:

$$\hat{u} = \arg \max_u \ln p(u | u_0, \Gamma) = \arg \max_u [\ln p(u_0 | u, \Gamma) + \ln p(u | \Gamma)], \qquad (69)$$

where u_0 is a given noisy image and Γ is the inpainting domain. More details can be found in [130], which summarizes TV and other variational approaches for image inpainting in terms of Bayesian philosophy of vision.

Image Decomposition, Sparse Representation, and Compressed Sensing

Many recent works on image inpainting consider texture recovery as well as image structure for good continuation. Image decomposition by mathematical transforms such as wavelet and DCT is able to separate an image into image structure and detailed texture. Many existing methods in this category contains the following two steps:

- Inpainting of smooth image space (i.e. piecewise smooth image with visually strong edges)

- Inpainting of textures (i.e. texture synthesis).

In [8], Bertalmio et al. combined their previous work [7] with texture synthesis. This can be achieved by an image decomposition method proposed in [150] introduced in equation (21) of Section 2.3.4. After image decomposition, two different images, i.e. a cartoon and a texture images, are obtained and the above-described two steps are applied. A solution of equations (62) and (63) is used for inpainting a decomposed BV image while a texture synthesis method from [46] is applied to a texture image.

Some inpainting algorithms proposed in [49, 53] use MCA approach introduced in Section 2.6.2. The minimization function in equation (31) by applying a diagonal mask matrix M that indicates inpainting region can be formulated as an image inpainting problem in the following:

$$F(\alpha_v, \alpha_w) = \|\nabla T_v \alpha_v\|_1 + \lambda \|M(u - T_v \alpha_v - T_w \alpha_w)\|_2^2 + \mu(\|\alpha_v\|_1 + \|\alpha_w\|_1). \qquad (70)$$

This can be rewritten in terms of v and w if we define $\alpha_v = T_v^+ v + r_v$ and $\alpha_w = T_w^+ w + r_w$ with simplification assumptions $r_v = r_w = 0$:

$$F(v, w) = \|\nabla v\|_1 + \lambda \|M(u - v - w)\|_2^2 + \mu(\|T_v^+ v\|_1 + \|T_w^+ w\|_1). \qquad (71)$$

Also, equation (70) can be simplified by considering two sparse dictionary matrices as a single unitary transform T (i.e. $u = T\alpha$):

$$F(\alpha) = \lambda \|M(I - T\alpha)\|_2^2 + \mu\|\alpha\|_1 \text{ or } F(Z) = \lambda \|M(I - Z)\|_2^2 + \mu\|T^H Z\|_1. \qquad (72)$$

In this case, equation (72) becomes equivalent to the model used in [61, 62].

Methods using transform-based sparse representations such as discrete cosine transform or wavelet transform were also proposed for the approximation of unknown regions [61, 30, 98]. A work in [112] uses K-SVD for redundant dictionaries and is applied to inpainting, denoising, and demosaicking.

Also, a method based on compressive sensing has recently been proposed in [126]. Equation (10) described in Section 2.6.3 is slightly modified by replacing $\mathbf{M} = \mathbf{S}\mathcal{F}^{\mathbf{H}}\mathbf{G}^{-1}\mathcal{F}$ as follows:

$$\min \|\boldsymbol{\alpha}\|_0 \text{ s.t. } \mathbf{y} = \mathbf{S}\mathcal{F}^{\mathbf{H}}\mathbf{G}^{-1}\mathcal{F}\boldsymbol{\Phi}\boldsymbol{\alpha}, \qquad (73)$$

where $\mathbf{S}$ is a sampling matrix that indicates element locations, $\mathcal{F}$ is Fourier transform matrix, and $\mathbf{S}$ is a diagonal matrix with values of a Gaussian function. The equation (73) is then solved by regularized orthogonal matching pursuit [105]. The method is very efficient if the inpainting pixels are randomly chosen, but it shows a limitation if an inpainting region is too big since randomness of measurement is lost.

Other Models

A method in [38] employs Ginzburg-Landau functional while localized wavelet-based approaches that use the second-order Allen-Cahn equation are considered. In their

previous work [9], Cahn-Hilliard equation was used for inpainting binary images. Nonlocal scheme that uses Mumford-Shah variational model is considered for inpainting color image in [68].

5.2.2 Texture Synthesis and Inpainting

One of the main problems in early research works for image inpainting is that *texture* in an image was not considered seriously. As seen in many approaches based on PDEs and variational models, image smoothness is postulated by considering image as a smooth functional space such as Sobolev space. As pointed out in [96], however, a natural image includes oscillatory texture layer as well as piecewise smooth image (cartoon or structural image). In addition, random noise should be considered. When we refer to the inpainting problem, it should be noted that only a partial part of an image (i.e. inpainting domain) is restored and the restored part is supposed to have the same properties as the other known regions of the image. In other words, the image statistics and the pattern (which can be represented by texture and noise) of the unknown inpainting region should be similar to the known regions of the image. Therefore, it is crucial to understand and utilize the texture properties in an image for inpainting problem. Another important factor for plausible inpainting work is the fact that there must be continuous connection between inpainting region and the other known regions.

Texture synthesis and quilting have been extensively used in image processing, computer vision and computer graphics. The problem requires to generate a texture image which is similar yet not identical to a given example patch of texture. The output texture could be any required size. There are a number of proposed approaches for this problem [35, 46, 45, 3, 153, 154, 145, 146, 166, 116, 117, 55, 59].

Exemplar-based Texture Synthesis

Many of the texture synthesis methods take an example patch (or exemplar) of given texture and estimate a growing pixel or patch by copying the exemplar[2]. A synthesis algorithm proposed by Efros and Leung is one of the well-known methods [46]. The synthesized texture is grown pixel by pixel by considering neighborhood window based on Markov random field (MRF) model. In other words, the probability of an unknown pixel is computed based on the spatially neighboring pixels. Then it is considered for a newly synthesized pixel.

The limitations of this approach can be caused by complicated texture structures and many different types of texture elements (*texels*) in a given texture sample. For example, a growing pixel can choose a wrong estimate from insufficient search space and growing garbage will be generated after many iterations. Also, a particular place can be stuck by verbatim copying. Another problem is that it becomes computationally intensive when a pixel-by-pixel growing method is used since it requires to compare all the possible patches for each generated pixel. Some of the later works based on Efros and Leung are extended for either reduced computational complexity or visual improvement [45, 3, 153].

Texture Synthesis Using Wavelet Transform

Some recent approaches perform texture synthesis based on multi-scale analysis and statistical properties in the wavelet domain [146, 166, 116, 117, 55, 59, 112]. Since texture typically contains oscillatory pattern probably with some random variation, a synthesized texture is supposed to have the similar pattern but different variational factor such as position and orientations. When we analyze a sample texture using wavelet transform, there are various useful statistical properties since wavelet performs multi-resolution analysis and has some useful properties such as locality and directionality. In addition, probability distribution models with a few parameters

---

[2]*Exemplar*, or example patch of texture, is called differently in the literature such as *epitomes* [66, 26] or *textons* [168].

such as generalized Gaussian distribution [90] represent statistics of the wavelet coefficients. Based on the properties of wavelet transform, the approaches usually utilize sample statistics such as mean, variance, and correlation in the wavelet domain.

5.2.3   Exemplar-based Inpainting

The methods based on image decomposition require to estimate the inpainting domain separately. However, it is also possible to consider both piecewise smooth image with strong edges and textures simultaneously. The algorithm proposed in [33] requires to give high priority to a pixel of the inpainting region which is supposed to have strong and continuous edge and then to copy an example patch (or pixels) from a sample image. There are mainly three steps for each iteration to fill a windowed inpainting patch in this algorithm.

First, the priority is decided by considering the product of *confidence* term $C(\mathbf{p})$ and *data* term $D(\mathbf{p})$, i.e. $P(\mathbf{p}) = C(\mathbf{p})D(\mathbf{p})$ and $C(\mathbf{p})$ and $D(\mathbf{p})$ are defined as follows:

$$C(\mathbf{p}) = \frac{\sum_{\mathbf{q} \in \Psi_{\mathbf{p}} \cap \Omega \backslash \Gamma} C(\mathbf{q})}{|\Psi_{\mathbf{p}}|}, D(\mathbf{p}) = |\nabla u_{\mathbf{p}}^{\perp} \cdot \mathbf{n_p}|,$$

where $\Psi_{\mathbf{p}}$ is a given patch centered at a point $\mathbf{p}$ on the boundary $\partial \Gamma$, $|\Psi_{\mathbf{p}}|$ the number of pixels in $\Psi_{\mathbf{p}}$, and $\mathbf{n_p}$ a unit vector orthogonal to $\partial \Gamma$. $P(\mathbf{p})$ is computed for all the pixels on the boundary $\partial \Gamma$. $C(\mathbf{p})$ is initially defined as 0 if $\mathbf{p} \in \Gamma$ and 1 otherwise. Since $\nabla u_{\mathbf{p}}^{\perp}$ can be considered as the isophote direction, the high priority is assigned when $\mathbf{p}$ is on a strong edge and its neighborhood pixels have high confidence. Secondly, once the highest priority pixel $\mathbf{p}$ is decided, the patch $\Psi_{\mathbf{p}}$ is filled by copying the closest exemplar $\Psi_{\mathbf{q}}$:

$$\Psi_{\mathbf{q}} = \arg \min_{\mathbf{q} \in \Omega \backslash \Gamma} d(\Psi_{\mathbf{p}}, \Psi_{\mathbf{q}}),$$

where $d(\cdot, \cdot)$ is a distance function between two patches such as Euclidean distance. Finally, confidence values $C(\mathbf{p})$ are updated in the filled pixels.

Since the approach gives high priority along the strong edges, edge sharpness as well as continuation is well-preserved. Also balanced region filling can be achieved by

changing the location by the priority (not filling by scanning in an orderly manner). However, there are some limitations: 1) if an image is irregularly complicated or there is no similar exemplar, it produces visually unacceptable severe errors, 2) once an error is occurred, newly generated exemplar with some error may be iteratively used, and 3) curved structure is not considered (more efficient for the line structure).

An extension using patch sparsity is proposed in [161] for measuring confidence and sparse linear combination of candidate patches. Another extension to use joint optimization of a single functional that consider the entire inpainting region is made in [156].

5.2.4 Graph-based Approaches

Graph-based structure and texture propagation by either confidence map or interactive guidance could be applied for better results as proposed in [137].

5.2.5 Hybrid Approaches

Some recent works take advantages of the existing schemes. A framework proposed in [12] successfully combines exemplar-based texture synthesis, PDE-based image structure model, and texture coherence into one variational formulation.

5.2.6 Inpainting Approaches in Wavelet Domain

The approaches introduced so far are performed in the spatial domain. However, some coefficients in the transformed domain may be unknown due to transmission or storage problem. For this, total variation model is used in [20, 16, 162, 167, 155].

5.2.7 Video Inpainting Approaches

Video offers spatio-temporal information, which enables to use correlation between temporal frames that may give redundant information for inpainting. In [110], foreground and background of video frames are segmented to build image mosaics and

then inpainted separately, which is followed by texture synthesis for spatio-temporal domain. A method in [132] proposes to use exemplar-based inpainting approach after tracking motion objects while reducing ghost effects in video. Another approach in [80] proposes a patch-based variational Bayesian framework for video inpainting that a nonlocal sparsity-based prior as motion-related information is used instead of explicit motion estimation. An approach in [159] uses prior belief propagation (BP) with regularized structure priors of a spatio-temporal Markov random field. Edge-based structure estimation for inpainting is used in [158]. Some inpainting applications focus on motion tracking information [131]. Recently, there have been new attempts to inpaint 3D videos [34, 104].

5.2.8   Specific Inpainting Approaches

If we know content and properties of image in advance, some assumptions and models can be made before an inpainting algorithm is applied. A method in [76] recovers texture maps of the occluded building facades from spatio-temporal image frames. In [87], *Bandlet* transform is used to remove clouds of remotely sensed images. Red-eye effect often seen in portrait photography with flash light can be corrected by inpainting technique [163]. Human posture model and sequence estimation are used to inpaint video with human in [83, 84].

5.2.9   Other Related Problems

Recovery of Lost Image Blocks and Error Concealment

One of the interesting applications related to image inpainting problem is recovery of missing blocks caused by image communication [152, 119, 108, 99, 160, 165]. The motivation comes from the fact that when an image comprises tiled small blocks by image/video coding algorithms such as JPEG and MPEG, some blocks in the image could be corrupted by a certain unstable communication process such as wireless transmission. The recovery of transmission error is also referred to error concealment

problem.

This problem is a subset of image inpainting problem, which has an additional constraint that inpainting domain is composed of a set of known tiling blocks. Therefore, it is possible to apply any inpainting algorithm. However, it is desirable to take advantages of the additional constraints which may result in faster, more accurate, and more robust algorithm.

In [119], both texture synthesis and non-texture image inpainting algorithms are selectively utilized by block classification. Like [8], a texture synthesis algorithm in [46] is used for a texture image and inpainting approach in [7] is applied to a cartoon image. However, the main difference from [8] is that each missing block is classified as either structure or texture and then synthesis or inpainting algorithm is selectively applied to the classified block. Block classification can be performed by coarseness measure defined as the number of local extrema.

Film Restoration

Automatic restoration of old films is also interesting application [74, 75, 73, 124, 138]. In the sense it is required to restore a sequence of corrupted images, the objective is similar to that of the inpainting problem. In this case, it is also possible to apply any inpainting algorithm if we assume that defects of the given film are known. However, the assumption is practically not useful due to the volume of image frames. Therefore, an automatic defect detection algorithm should be considered. In addition, time complexity should be as low as possible even though it is not always required to achieve real-time system. For example, if an algorithm can process a single frame in 1 minute, 20 minutes of film with 24 frames per second requires $20 \times 60 \times 24 = 28,800$ minutes, or 20 days. Also, useful motion information is available by analyzing neighboring image frames since the difference between adjacent frames is little in most cases.

5.3 Image Inpainting Using Mumford-Shah Model and Texture Mapping[3]

This section presents an image inpainting method that uses a segmentation map using hierarchical level set based on Mumford-Shah inpainting model. In this approach, we combine two schemes to solve image inpainting problem. First, a variational approach has been chosen for estimating image structure such as strong edges in inpainting domain. As we saw in Section 5.2.1, a variational approach that can be also expressed as PDE modeling gives good and smooth continuity for the unknown inpainting domain. Second, a texture synthesis method using segmentation map is employed for preserving texture and the details of image more naturally. Since a proposed Mumford-Shah inpainting model is capable of estimating segmentation of inpainting domain, complete segmentation of both known and unknown regions in an image can be obtained. Based on the segmentation map with good image structure in the unknown inpainting domain, details and texture of inpainting domain can be estimated. The texture mapping approach that copies exemplar patches that are chosen to fill the unknown area with *copy-paste* criteria helps estimating more natural image information. By maximizing the advantages of two different schemes, the proposed approach produces visually pleasing results.

In the following section, we first present an Mumford-Shah inpainting model and its level-set solution using hierarchical approach, which yields smooth segmentation map with fair amount of detailed structure. In Section 5.3.2, the texture filling algorithm using the produced segmentation map is described. Experimental results are also presented in Section 5.3.3. Section 5.3.4 summarizes and concludes briefly the proposed approach.

---

[3]This is a work with Xiaojun Du as well as Prof. Tien D. Bui. The original work was published in *Signal Processing*, 2011 [44].

5.3.1 Mumford-Shah Inpainting Model and Hierarchical Approach for Segmentation

As we saw in Section 2.3.3, Mumford-Shah (MS) variational model has been successfully used in image segmentation. In equation (66), Mumford-Shah model for image inpainting was introduced. This can be rewritten in the following form if we assume that the curve C is closed and an image u is partitioned into an inside region u_1 and an outside region u_2:

$$
\begin{aligned}
F_{MS}(u_1, u_2, C) &= \int_{\Omega_1} \lambda(x, y) |u_1 - u_0|^2 \, dxdy + \mu \int_{\Omega_1} |\nabla u_1|^2 \, dxdy \\
&+ \int_{\Omega_2} \lambda(x, y) |u_2 - u_0|^2 \, dxdy + \mu \int_{\Omega_2} |\nabla u_2|^2 \, dxdy + \nu|C|, \quad (74)
\end{aligned}
$$

where Ω_1 and Ω_2 are the regions inside and outside C respectively, and $\lambda(x, y)$ is a given mask function, i.e. $\lambda(x, y) = 0$ if $(x, y) \in \Gamma$ and $\lambda(x, y) = 1$ otherwise. Then two region segmentation using one level set function can be defined as follows:

$$
\phi(x, y, t) = \begin{cases} > 0 & \text{if } (x, y) \in \Omega_1 \\ = 0 & \text{if } (x, y) \in C \\ < 0 & \text{if } (x, y) \in \Omega_2 \end{cases} \quad (75)
$$

Equations (74) and (75) lead us the following segmentation curve evolution equation [149]:

$$
\begin{aligned}
\frac{\partial \phi}{\partial t} &= \delta(\phi) \Big[\nu \nabla \cdot \Big(\frac{\nabla \phi}{|\nabla \phi|} \Big) - \lambda(x, y)(u_1 - u_0)^2 - \mu|\nabla u_1|^2 \\
&+ \lambda(x, y)(u_2 - u_0)^2 + \mu|\nabla u_2|^2 \Big].
\end{aligned} \quad (76)
$$

In equation (74), u_1 and u_2 can be replaced by constants c_1 and c_2 in the piecewise constant approximation. By considering two phase segmentation using the simplified Mumford-Shah model and level set method, we can define the following variational

model:

$$F(c_1, c_2, \phi) = \int \lambda(x,y)(c_1 - u_0)^2 H(\phi) dxdy \tag{77}$$
$$+ \int \lambda(x,y)(c_2 - u_0)^2 (1 - H(\phi)) dxdy + \nu \int \delta(\phi)|\nabla \phi| dxdy$$

where $H(x)$ is the Heaviside function, i.e. $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$. For more details see [44].

In our inpainting model, we employ the hierarchical scheme of Gao and Bui proposed in [60]. In this approach, the PDEs of two level-set functions are decoupled in hierarchical order instead of solving the coupled PDEs. In general, n level set functions are required to segment 2^n regions. The coupled equation in this case results in complicated and time-consuming computations. Since PDEs for curve evolutions in Gao and Bui's approach are decoupled into segmentation and diffusion, the implementation is fast and robust with regard to the initial condition. For inpainting, the segmentation of a given image is performed using piecewise constant approximation of Mumford-Shah model as shown in the previous section.

After the segmentation, diffusion can be used for inpainting purpose. Diffusion-based filling can achieve smooth region filling with good edge preservation and denoising. Therefore, it is efficient for small or long and narrow inpainting regions. However, for large inpainting regions, good estimation of overall structure and preservation of texture patterns is difficult.

5.3.2   Segmentation Map and Texture Filling

Variational or PDE approaches for image inpainting produce good image structure and edge continuity. However, they do not fill the inpainting domain by the statistical properties of other parts of the images since the defined model is usually spatially continuous and smooth. Therefore texture and image details are not easily recovered. On the other hand, natural image has some random and stochastic properties. For example, if the image is contaminated by randomly distributed noise, the inpainting

domain is also supposed to have similar kind of noise although it may be desirable to enhance the whole image. In fact, details with little information, texture with oscillatory or irregular patterns, and even random noise can be considered for more natural and visually satisfactory inpainting. In addition, the heterogeneous smooth filling with Mumford-Shah approach is more noticeable for the relatively large inpainting region while the thin and long regions such as scratches can be restored relatively better. In order to achieve filling both image structure and texture details, a region filling approach using segmentation map and similar patches is proposed.

Once we get a segmentation map by solving Mumford-Shah inpainting model described in Section 5.3.1, the overall image structure of the unknown region can be estimated. On top of the overall structure, it is desirable to *overpaint* using the *most plausible texture* patches. Recently some inpainting approaches utilize texture synthesis algorithm to obtain close similarity between the unknown inpainting region and the surrounding image regions as we saw in Sections 5.2.2 and 5.2.3. In our approach, we take the scheme that could fill in plausible example texture patches into the inpainting regions.

Inpainting of texture patterns and details in image is performed by a few steps described in the following. First, the complete segmentation map of an image is obtained as explained in the last section. It contains the estimated edge structure of the inpainting domain and consists of all the segmented partitions in the whole image. The segmentation map is capable of representing the overall image structure without details as we can see in figure 27.

In the second step, patch priorities that decide the filling orders in the inpainting domain are computed. Texture mapping procedure is to *copy* a good example patch and then *paste* it into a corresponding inpainting area iteratively. In every copy and paste iteration, the inpainting region that has already been filled is also considered as example patches in order to achieve good continuity. Therefore, it is important to have right the filling orders and in fact the inpainting performance is affected by the filling orders. The priorities are computed in a similar way as described in [33], i.e.

85

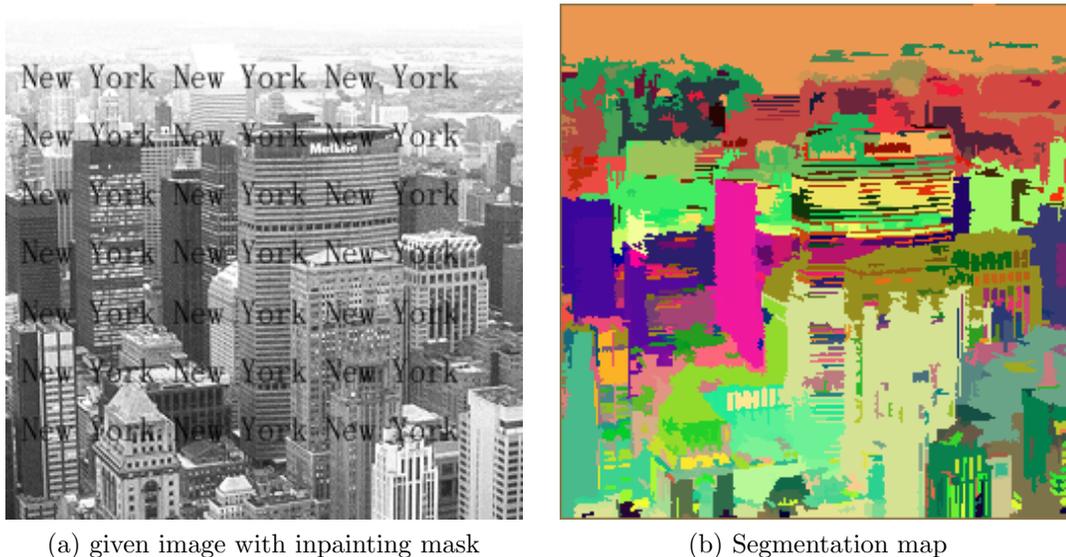(a) given image with inpainting mask   (b) Segmentation map

Figure 27: Segmentation map of *City 2* image.

the inpainting boundary with the segmentation edge gives higher priority while plain region without any edge structure and far from the inpainting boundary has lower priority.

Once we decide the priorities, the patch around the highest priority pixel on the segmentation edge is filled first. The patch is filled by finding the most similar patch (exemplar) to the patch to be filled that excludes the unknown pixels. Euclidean distance measure between two patches is used for similarity measure due to its simplicity and popularity. However, a good similarity measure between two patches is important to get good continuity and please HVS. The exemplar patch is found by searching and comparing all the image regions that do not belong to the inpainting domain. Instead of searching all the image regions, the search area can be confined to speed up the algorithm.

In the last step of the filling procedure, we fill in the remaining inpainting areas which have no segmented edges from boundary to inside of the inpainting domain. The search of similar patch can be done in the same manner as the previous step that fills in the high priority regions.

86

5.3.3 Experimental Results

In the experiments, we have applied our inpainting algorithm to various different images including gray-scale and color images, natural images, and paintings. Figure 28 shows a real artistic inpainting example to the wall painting with some damages (e.g. cracks). The painting is a small part of the 'Last Judgment', which is Michaelangelo's fresco in Sistine Chapel. The inpainting regions around the damaged cracks are masked manually as can be seen in figure 28b. Segmentation map is constructed for edge estimation and the exemplar-based filling method has been applied to the damaged image. Although the continuous edges are discontinued by the cracks (or mask), the image structure including the discontinuous edges are estimated smoothly well in the inpainting domain. Our approach is performed well to estimate edge structure especially with narrow mask.

Figures 29 and 30 are the examples of object removal or disocclusion that have larger and fatter masks than the previous example in figure 28. We compare the inpainting results of our approach to some other methods. The method proposed in [106] uses a simple implementation that convolve a small 3×3 filter iteratively. In our experiment, we use 1000 iterations. The method usually produces smooth inpainting results that does not consider either image structure or texture. Also, implementations of OpenCV library [10] for Navier-Stokes approach and the algorithm in [139] were used for the evaluation. These implementations are fast and efficient, but we can notice some partitions with a certain degree of discontinuity. These methods do not consider texture or details. Another PDE-based approach proposed in [7] is also implemented for the comparison. In these examples, PDE and variational approaches have difficulty in inpainting of natural images where it is important to consider texture, patterns, and statistical properties of the image. The estimated boundary between different textured regions is defined in the segmentation map. For example, the ridge of a mountain occluded by the rightmost pillar in figure 30h is recovered based on edge information of figure 30g.

(a) Original image
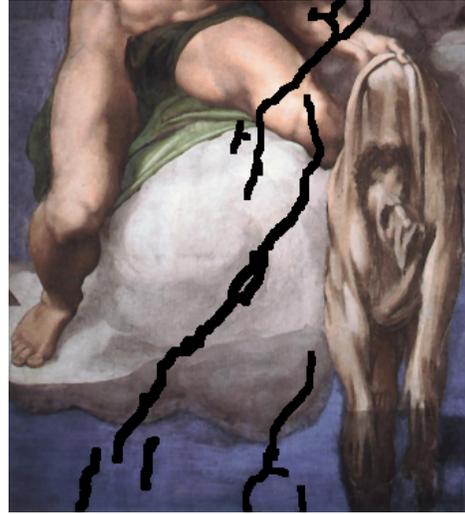
(b) Masked image
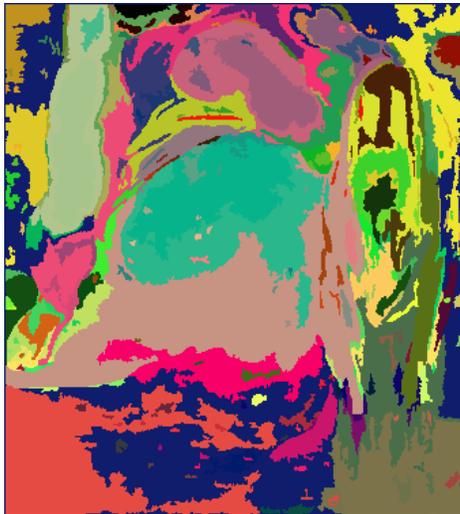
(c) Segmentation map
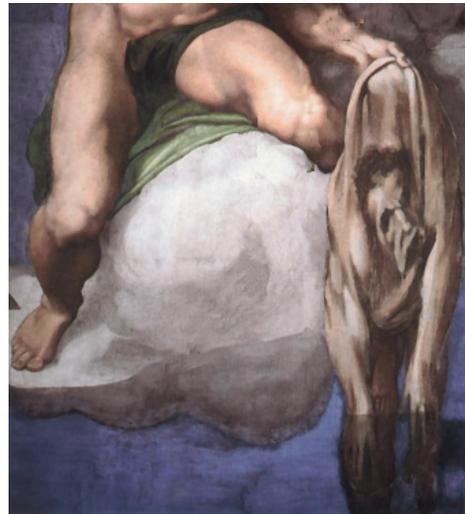
(d) Inpainted image

Figure 28: Inpainting results using Michaelangelo's Last Judgment image (partial)

(a) Original image

(b) Masked image

(c) Linear filter [106]

(d) Telea [139]

(e) Navier-Stokes [10]

(f) Bertalmio [7]

(g) Proposed segmentation map

(h) Proposed

Figure 29: Inpainting results using 'fish' image

(a) Original image

(b) Masked image

(c) Linear filter [106]

(d) Telea [139]

(e) Navier-Stokes [10]
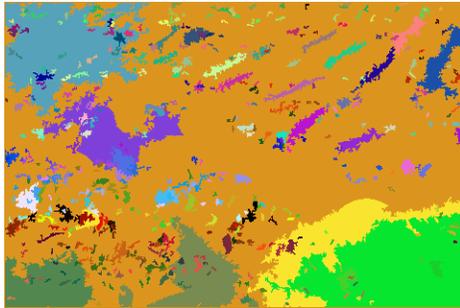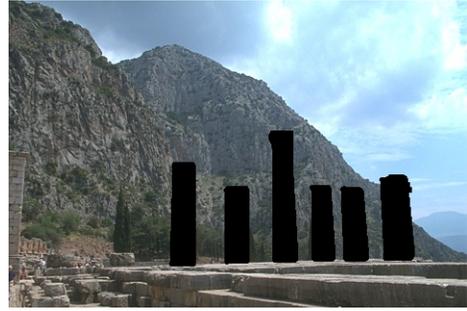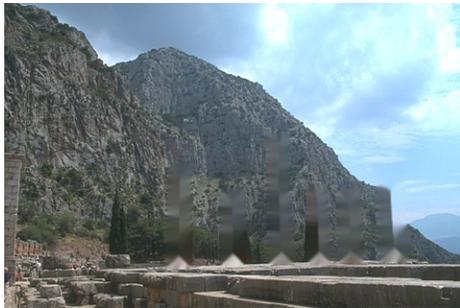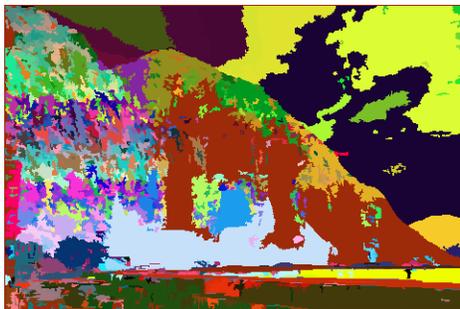

(f) Bertalmio [7]

(g) Proposed segmentation map

(h) Inpainted image

Figure 30: Inpainting results using 'pillar' image

5.3.4 Conclusion

We have presented an image inpainting approach that considers both edge structure and details in image for visually pleasant and more natural inpainting performance. The proposed Mumford-Shah inpainting model with hierarchical level-set approach is able to detect both main image structure and detailed edges in the inpainting domain and produce a partition-based segmentation map. However, inpainting using solely the Mumford-Shah model [51, 19] does not estimate structure of complicated and detailed texture. To overcome this problem, texture filling approach efficiently uses the segmentation map to inpaint details and texture patterns in an image. Examples from the experiments show that the method performs visually well.

5.4 Image Inpainting Using Wavelet-based Inter- and Intra-scale Dependency[4]

As discussed in Section 2.5, wavelet transform has been used for a good image representation and analysis tool due to its multi-resolution analysis, data separability, compaction and sparsity features in addition to statistical properties. In this section, we propose a wavelet-based approach for image inpainting. There are a few major advantages to use wavelet analysis for image inpainting problem. First, multi-resolution analysis of wavelet transform can be helpful to predict coarse-to-fine image structure in the inpainting domain. When we recall an example in figure 24, filling process by perceptual interpolation requires to analyze a *global* image structure as a whole (e.g. perception of triangle) and at the same time consider *local* objects (small objects or pixel level analysis). Both global and local analysis is necessary for reasonable estimation of image structure. Analysis of different image scale will help to analyze and estimate both global and local image content. Second, separable data into low frequency scaling coefficients and high frequency wavelet coefficients make it possible

---

[4]This work was originally published in *ICPR* 2008 [30].

to analyze both structure and texture independently. In addition to the image structure, it is also important to analyze texture and detailed patterns for natural images. Wavelet transform can decompose an image into low-pass filtered image structure and high-pass filtered details. As we reviewed image decomposition approaches for image inpainting in Section 5.2.1, both image structure and texture contribute inpainting quality and performance. Wavelet is capable of treating these elements altogether.

5.4.1 Proposed Approach

We first recall inpainting problem : given an image u with unknown regions Γ, find an ideal image ν:

$$\nu(x,y) = \begin{cases} \hat{u}(x,y) & \text{if } (x,y) \in \Gamma \\ u(x,y) & \text{otherwise} \end{cases} \tag{78}$$

Now the goal is to find a good estimation $\hat{u}$ in Γ.

Many inpainting solutions are achieved by analyzing image models and utilizing them to obtain probable estimation. In our proposed algorithm, wavelet and scaling coefficients are estimated after wavelet decomposition of a given incomplete image. This approach does not require iterative decomposition and reconstruction, which gives computational advantage. Discrete wavelet decomposition for each scale can be formulated as follows:

$$s_{j+1} = h_j * h_j * s_j,$$

$$w_{j+1}^{LH} = g_j * h_j * s_j,$$

$$w_{j+1}^{HL} = h_j * g_j * s_j,$$

$$w_{j+1}^{HH} = g_j * g_j * s_j,$$

where h_j and g_j represent low-pass scaling and high-pass wavelet filters, s_j and w_j are scaling and wavelet coefficients for scale j respectively. Wavelet transform is applied to the given original image u and image of the mask m. The mask is a user defined area that covers the inpainting region.
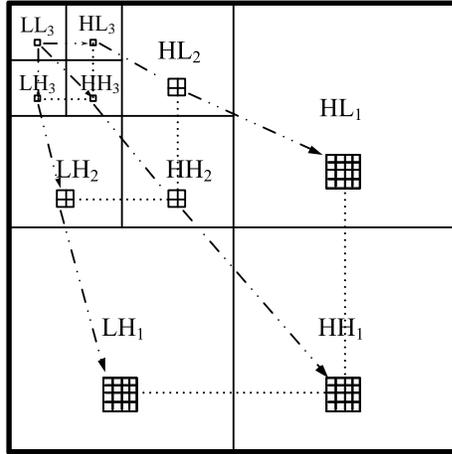
Figure 31: Inter-scale dependency of wavelet coefficients

Inpainting process is performed in the wavelet domain by predicting both scaling and wavelet coefficients from coarse to fine scales in the unknown regions. As shown in figure 31, each coefficient in the wavelet domain is dependent on neighboring coefficients in the same subband and corresponding coefficients located in the other subbands. Therefore, it is important to consider both inter and intra-scale dependency. On the other hand, if the dependency of these coefficients is not considered, it would be difficult to estimate visually meaningful coefficients.

Once forward transform has been carried out for both the original image and the image of the mask up to level L (the coarsest level); scaling coefficients of the coarsest scale need to be estimated for global image structure. The coarsest scaling subband is required to be filled first as depicted in figure 31. Since the coarsest scaling coefficients have the same properties as the low-pass filtered smooth image, any inpainting existing algorithm described in the previous Section 5.2 could be applied here. We have applied the total variation (TV)-based algorithm used in [17] since the approach works well in smooth image domain. After filling the scaling subband, three wavelet subbands (HL, LH, and HH) in the coarsest level are filled simultaneously to avoid visually annoying artifacts after the inverse transform. In order to complete the wavelet subbands, exemplar-based scheme proposed in [33] has been used. In other words, unknown areas in the wavelet subbands are filled by example patches of the

same subband based on a priority map (see [33]) which is obtained from the scaling subbands of u and m. In this case, example patches in three wavelet subbands should be geometrically corresponding to each other to reduce visual artifacts. An example patch should resemble a patch area in the unknown region, i.e. the difference or distance between a patch including the unknown area and an example patch should be minimized. For this, we propose the following distance measure:

$$d = \alpha|w_{LH}^d - w_{LH}^e| + \beta|w_{HL}^d - w_{HL}^e| + \gamma|w_{HH}^d - w_{HH}^e| + |s_d - s^e| \qquad (79)$$

where w_{LH}^e and w_{LH}^d are example patch and destination patch in the LH subband respectively. The destination patch w_{LH}^d includes the unknown area in the LH subband, similarly for the other subbands. s^e and s^d are example patch and destination patch from the scaling subband. Finally α, β, and γ are parameters to be determined. By default, they are set to 1.

Once all the coefficients are estimated, inverse transform for one level is performed to obtain the approximation of the next finer scale. In the subsequent scales, as in the coarsest scale, reconstructed scaling subband is used for the estimation of wavelet subbands. The scaling subband can also be corrected by blending with the example patch corresponding to the patches decided for the wavelet coefficients. Blending can be done linearly, i.e.

$$s^d \leftarrow \alpha s^d + (1 - \alpha)s^e \qquad (80)$$

where $\alpha \in [0, 1]$ is a parameter.

Summary of proposed algorithm is described as follows:

1. Apply forward wavelet transform to a given image u.

2. Apply forward wavelet transform to the image of the masks m.

3. Set scale parameter $j = L$ (coarsest level).
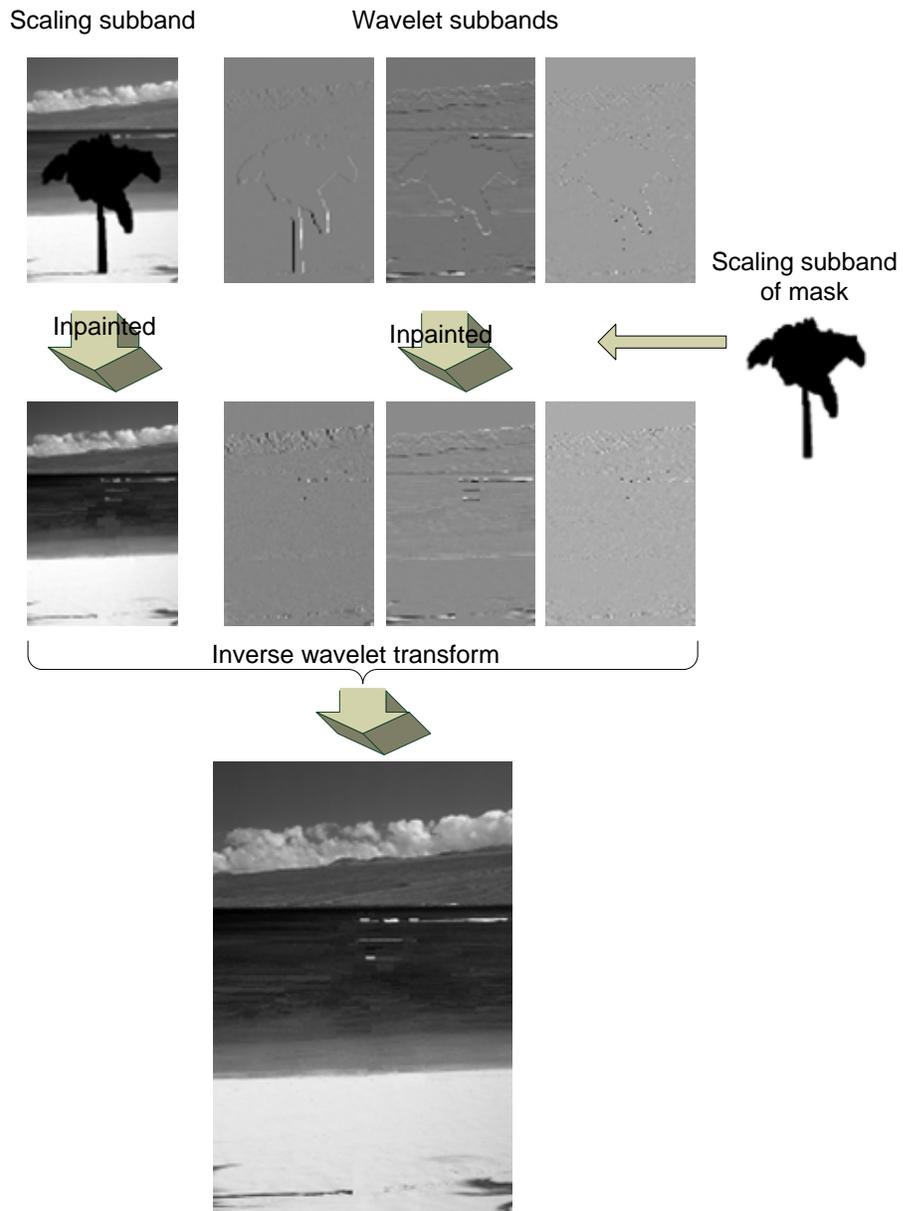
4. While $j > 0$ (from coarse to fine),

Figure 32: Inpainting in wavelet domain (one level): scaling and wavelet subbands are estimated and reconstructed.

(a) If j is L, fill the unknown area of scaling subband by using PDE or variational inpainting approach (e.g. [17]). Otherwise, scaling coefficients have been already reconstructed in 4.(d) of the previous scale.

(b) Estimate wavelet coefficients in the wavelet subbands by finding closest texture patches that can minimize distance d given by equation (79). Geometrically corresponding scaling coefficients to the estimated wavelet coefficients are also stored.

(c) The reconstructed scaling coefficients are blended with the estimated scaling coefficients in the step 4(b). Blending function is in equation (80).

(d) Perform inverse wavelet transform to obtain scaling subband of the next finer scale $j - 1$.

(e) $j \leftarrow j - 1$.

5. (Optional) Post-processing (e.g. blending)

5.4.2 Experimental Results

For the experiments, we assume that masks for the unknown regions are defined manually. The mask image, m, is transformed by wavelet and the scaling subband for each level is used for the inpainting of the subbands of the image at the same level. Different wavelets could be considered. In our experiments, we simply use the Haar basis with $L = 2$ or 3. The proposed algorithm has been applied to various gray and color images. In figure 34, inpainting results of some color images are shown. Both texture patterns and image structure are well-preserved.

We also compared the results of existing algorithms such as PDE-based approach [7], total variation [17], exemplar-based approach [33], and decomposition and sparse reconstruction [61]. Figure 34 shows inpainting results from these different methods. A narrow region like a golf driver can be smoothly filled by most of the methods.

However, for large regions, PDE-based or variational algorithms result in smooth surface even for textured background in addition to diminution of edge sharpness.

Figure 33: Inpainting results of proposed method: original (top row), masked images (middle row) and completed (bottom row) images

(a) given image

(b) PDE-based [7]

(c) total variation [17]

(d) exemplar-based [33]

(e) sparse reconst. [61]

(f) proposed

Figure 34: Inpainting results of different approaches.

Exemplar-based approach mostly produces plausible textures in the unknown area, but some abrupt changes between patched areas are observed. Our proposed method also has slight seam from time to time, but overall quality in terms of image structure and texture looks comparatively better than the other algorithms.

5.4.3 Conclusion

We have presented wavelet-based approach for image inpainting. Wavelet is a decent mathematical tool for estimating global structure of image and texture analysis thanks to its multi-scale analysis and separability features. The proposed method takes advantages of wavelet by utilizing inter- and intra-scale dependency for maintaining image structure and texture quality.

Chapter 6

Conclusion and Future Work

In this thesis, image enhancement techniques using useful mathematical tools have been discussed. Image enhancement algorithms that include image denoising, contrast enhancement, and inpainting are indispensable to the field of image processing and computer vision. As shown in this thesis, there exist various mathematical approaches that play critical roles in solving image enhancement problems. In the following paragraphs, the summary of results, contributions, and future works for each image enhancement problem presented in this thesis are presented briefly.

In Chapter 2, various mathematical approaches and their close relationships to image enhancement problems have been discussed. Some examples show that a mathematical model can be expressed by PDE, calculus of variations, wavelet transform, Bayesian framework, sparse representation, or some other mathematical approaches and they are fundamentally equivalent or share common properties. Establishment of more complete theory on unifying these mathematical approaches as a whole for image processing could be interesting future work.

In Chapter 3, an efficient color image denoising framework using wavelet transform was presented. The framework decorrelates the color components and plugs in a shrinkage method flexibly. The wavelet-based approach produces superior results compared to some spatial-based algorithms. Most of experimental results show the

proposed denoising framework that considers correlations between color channels usually gives about 1 *dB* higher than the RGB color space-based approach in terms of PSNR measure when other conditions are same. The proposed approach only uses wavelet transforms, but denoising performance can probably be improved by combining some schemes such as PDE, variational behaviors, or patch-based analysis.

In Chapter 4, a novel algorithm for brightness, contrast, and color enhancement in the compressed wavelet domain was proposed. To our knowledge, this is a unique work for a wavelet encoded image that considers fast complexity and advantages of wavelet properties while scaling coefficients are modified for global brightness enhancement. In addition, the results are comparable or better than the other existing spatial domain-based approaches visually and numerically.

In Chapter 5, a review on image inpainting problem was presented and two inpainting approaches were proposed. The problem was described as mathematical and psychological perception with extensive literature survey. A hybrid image inpainting approach using Mumford-Shah model and texture mapping was proposed. This method can estimate the image structure by a proposed variational model while image details such as texture are considered. The experimental results show that the approach works well for many cases. However, there is a limitation when the inpainting region is too complex or too big. Another inpainting approach using wavelet transform is also presented. Since discrete wavelet analysis produces different scales of image information, it is possible to estimate global image structure in a different scale. In addition, wavelet coefficients contain texture information. This multiscale and data separation properties of wavelet transform produce relatively good image completion results. Proposed inpainting approaches can be improved by using a *good* local model, e.g. an image patch. More intelligent similarity measure between image patches will leverage the inpainting algorithms. Also, invariant learning-based criteria from other images could be helpful for ambiguous cases.

In addition to some detailed issues to be solved in each enhancement problem, some applications to image enhancement can be considered as future works. For

example, detection of image degradation types could be considered by analyzing a degraded image statistically and geometrically. Once we understand the nature of degradation in an image, one or more image enhancement approaches can be applied. Another possible future work is to employ these methods in a specific problem. For instance, we may consider to restore partial occlusion of character scene, remove unwanted objects in an image, or enhance a specific cell image. Also, applications to video or 3D scene and surface are being recently studied and it could be interesting to consider them as a future research. An image could have multiple degradation sources that may require to solve different enhancement problems.

Bibliography

[1] S. S. Agaian, B. Silver, and K. A. Panetta. Transform coefficient histogram-based image enhancement algorithms using contrast entropy. *IEEE Transactions on Image Processing*, 16(3):741–758, 2007.

[2] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311 –4322, nov. 2006.

[3] M. Ashikhmin. Synthesizing natural textures. In *ACM Symp. Interactive 3D Graphics*, pages 217–226, 2001.

[4] M.-F. Auclair-Fortier and D. Ziou. A global approach for solving evolutive heat transfer for image denoising and inpainting. *IEEE Transactions on Image Processing*, 15(9):2558–2574, 2006.

[5] Jean Babaud, Andrew P. Witkin, Michel Baudin, and Richard O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):26 –33, jan. 1986.

[6] M. Bertalmio. Strong-continuation, contrast-invariant inpainting with a third-order optimal PDE. *IEEE Transactions on Image Processing*, 15(7):1934–1938, 2006.

[7] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 417–424, 2000.

[8] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE Transactions on Image Processing*, 12(8):882–889, 2003.

[9] A.L. Bertozzi, S. Esedoglu, and A. Gillette. Inpainting of binary images using the Cahn-Hilliard equation. *IEEE Transactions on Image Processing*, 16(1):285–291, 2007.

[10] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[11] M. Braukus and K. Henry. Retinex image processing. NASA News Release 01-086, August 2001. Accessed Sep., 2010.

[12] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro. A comprehensive framework for image inpainting. *IEEE Transactions on Image Processing*, 19(10):2634–2645, 2010.

[13] T. T. Cai and B. W. Silverman. Incorporating information on neighbouring coefficients into wavelet estimation. *Sankhya: The Indian Journal of Statistics, Serie A*, 63:127–148, 2001.

[14] E.J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489 – 509, feb. 2006.

[15] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21 –30, march 2008.

[16] R.H. Chan, Y.-W. Wen, and A.M. Yip. A fast optimization transfer algorithm for image inpainting in wavelet domains. *IEEE Transactions on Image Processing*, 18(7):1467–1476, 2009.

[17] T. F. Chan, S. H. Kang, and J. Shen. Euler's elastica and curvature based inpainting. *SIAM J. Appl. Math.*, 63(2):564–592, 2002.

[18] T. F. Chan and J. Shen. Mathematical models for local non-texture inpainting. *SIAM J. Appl. Math.*, 62(3):1019–1043, 2002.

[19] T. F. Chan and J. Shen. *Image Processing And Analysis: Variational, PDE, Wavelet, And Stochastic Methods.* SIAM, Philadelphia, 2005.

[20] T. F. Chan, J. Shen, and H. Zhou. Total variation wavelet inpainting. Technical Report 04-47, University of California at Los Angeles, CA, 2004.

[21] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[22] G. Y. Chen and T. D. Bui. Multiwavelet denoising using neighbouring coefficients. *IEEE Signal Processing Letters*, 10(7):211–214, 2003.

[23] S.-D. Chen and A. R. Ramli. Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE Transactions on Consumer Electronics*, 49(4):1310–1319, 2003.

[24] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[25] H. D. Cheng, R. Min, and M. Zhang. Automatic wavelet base selection and its application to contrast enhancement. *Signal Processing*, 90:1279–1289, 2010.

[26] V. Cheung, B.J. Frey, and N. Jojic. Video epitomes. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 42–49, 2005.

[27] D. Cho. Image denoising using wavelet transforms. Master's thesis, Concordia University, Montreal, 2004.

[28] D. Cho and T. D. Bui. Monochrome and color image denoising using neighboring dependency and data correlation. In *Proceedings of SPIE Symposium on*

Optics & Photonics (Wavelets XI), volume 5914, page 59141E, San Diego, CA, 2005.

[29] D. Cho and T. D. Bui. Multivariate statistical modeling for image denoising using wavelet transforms. *Signal Processing: Image Communication*, 20(1):77–89, July 2005. * This paper won the Best Paper Award by the European Society for Signal Processing (EURASIP) in 2010.

[30] D. Cho and T. D. Bui. Image inpainting using wavelet-based inter- and intra-scale dependency. In *International Conference on Pattern Recognition (ICPR)*, pages 1–4, Tampa, FL, December 2008.

[31] D. Cho, T. D. Bui, and G. Y. Chen. Image denoising based on wavelet shrinkage using neighbor and level dependency. *Int. Journal of Wavelets, Multiresolution and Information Processing*, 7(3):299–311, 2009.

[32] M. J. Cree. Observations on adaptive vector filters for noise reduction in color images. *IEEE Signal Processing Letters*, 11(2):140 – 143, 2004.

[33] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13(9):1200–1212, 2004.

[34] I. Daribo and H. Saito. A novel inpainting-based layered depth video for 3dtv. *IEEE Transactions on Broadcasting*, 57(2):533–541, 2011.

[35] Jeremy S. De Bonet. Multiresolution sampling procedure for analysis and synthesis of texture images. In Turner Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, pages 361–368. ACM SIGGRAPH, Addison Wesley, August 1997. ISBN 0-89791-896-7.

[36] A. de Masi, R. Esposito, and J. L. Lebowitz. Incompressible navier-stokes and euler limits of the boltzmann equation. *Comm. Pure Appl. Math.*, 42(8):1189?1214, 1989.

[37] A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis: A Probabilistic Approach.* Springer Verlag, 2008.

[38] J.A. Dobrosotskaya and A.L. Bertozzi. A wavelet-laplace variational technique for image deconvolution and inpainting. *IEEE Transactions on Image Processing*, 17(5):657–663, 2008.

[39] D. Donoho and I. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[40] D. L. Donoho. Denoising by soft-thresholding. *IEEE Transactions on Information Theory*, 41:613–627, 1995.

[41] David L. Donoho. Sparse components of images and optimal atomic decomposition. *Constr. Approx*, 17(3):353–382, 2001.

[42] D.L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845 –2862, nov 2001.

[43] D.L. Donoho and J. Tanner. Precise undersampling theorems. *Proceedings of the IEEE*, 98(6):913 –924, june 2010.

[44] X. J. Du, D. Cho, and T. D. Bui. Image segmentation and inpainting using hierarchical level set and texture mapping. *Signal Processing*, 91:852–863, 2011.

[45] A. A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. *Proceedings of SIGGRAPH 2001*, pages 341–346, August 2001.

[46] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision*, pages 1033–1038, Corfu, Greece, September 1999.

[47] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736 –3745, dec. 2006.

[48] M. Elad, M.A.T. Figueiredo, and Yi Ma. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, 98(6):972 –982, june 2010.

[49] M. Elad, J-L. Starck, P. Querre, and D.L. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *J. Applied and Computational Harmonic Analysis*, 19:340–358, Aug 2005.

[50] Michael Elad. *Sparse and Redundant Representations - From Theory to Applications in Signal and Image Processing.* Springer, 2010.

[51] S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *European J. Appl. Math.*, 13:353–370, 2002.

[52] J.M. Fadili, J.-L. Starck, M. Elad, and D.L. Donoho. MCALab: Reproducible research in signal and image decomposition and inpainting. *Computing in Science & Engineering*, 12(1):44–63, 2010.

[53] M.J. Fadili and J.-L. Starck. EM algorithm for sparse representation-based image inpainting. In *IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 61–63, 2005.

[54] M.J. Fadili, J.-L. Starck, J. Bobin, and Y. Moudden. Image decomposition and separation using sparse representations: An overview. *Proceedings of the IEEE*, 98(6):983 –994, june 2010.

[55] G. Fan and X.-G. Xia. Wavelet-based texture analysis and synthesis using hidden Markov models. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 50(1):106–120, 2003.

[56] C. Fernández, J.Osiewalski, and M. Steel. Modelling and inference with v-spherical distributions. *Journal of the American Statistical Association*, 90(432):1331–1340, 1995.

[57] M.A.T. Figueiredo, R.D. Nowak, and S.J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):586 – 597, dec. 2007.

[58] B. Funt, F. Ciurea, and J. McCann. Retinex in MATLAB. *Journal of Electronic Imaging*, 13(1):48–57, 2004.

[59] C. Gallagher and A. Kokaram. Non-parametric wavelet based texture synthesis. In *IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 462–465, 2005.

[60] S. Gao and T.D. Bui. Image segmentation and selective smoothing by using Mumford-Shah model. *IEEE Transactions on Image Processing*, 14(10):1537–1549, Oct 2005.

[61] O. G. Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: Part I - theory. *IEEE Transactions on Image Processing*, 15(3):539–554, 2006.

[62] O. G. Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: Part II - adaptive algorithms. *IEEE Transactions on Image Processing*, 15(3):555–571, 2006.

[63] K. Q. Huang, Z. Y. Wu, G.S.K. Fung, and F. H. Y. Chan. Color image denoising with wavelet thresholding based on human visual system model. *Elsevier Signal Processing: Image Communications*, 20(2):115 – 127, 2005.

[64] D. J. Jobson, Z. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6(7):965–976, 1997.

[65] D. J. Jobson, Z. Rahman, and G. A. Woodell. Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*, 6(3):451–462, 1997.

[66] N. Jojic, B. J. Frey, and A. Kannan. Epitomic analysis of appearance and shape. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 34–41, 2003.

[67] C.R. Jung and J. Scharcanski. Sharpening dermatological color images in the wavelet domain. *IEEE Journal of Selected Topics in Signal Processing*, 3(1):4–13, 2009.

[68] M. Jung, X. Bresson, T.F. Chan, and L.A. Vese. Nonlocal Mumford-Shah regularizers for color image restoration. *IEEE Transactions on Image Processing*, 20(6):1583–1598, 2011.

[69] G. Kanizsa. *Organization in Vision: Essays on Gestalt Perception*. Praeger, New York, 1979.

[70] J. Kim, M. Cetin, and A. S. Willsky. Nonparametric shape priors for active contour-based image segmentation. *Signal Processing*, 87(12):3021–3044, 2007.

[71] S. Kim. PDE-based image restoration: a hybrid model and color image denoising. *IEEE Transactions on Image Processing*, 15(5):1163–1170, 2006.

[72] N. G. Kingsbury. Image processing with complex wavelets. *Phil. Trans. Royal Society London A*, 357:2543–2560, September 1999.

[73] A.C. Kokaram. On missing data treatment for degraded video and film archives: a survey and a new Bayesian approach. *IEEE Transactions on Image Processing*, 13(3):397–415, 2004.

[74] A.C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P.J.W. Rayner. Detection of missing data in image sequences. *IEEE Transactions on Image Processing*, 4(11):1496–1508, 1995.

[75] A.C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P.J.W. Rayner. Interpolation of missing data in image sequences. *IEEE Transactions on Image Processing*, 4(11):1509–1519, 1995.

[76] T. Korah and C. Rasmussen. Spatiotemporal inpainting for recovering texture maps of occluded building facades. *IEEE Transactions on Image Processing*, 16(9):2262–2271, 2007.

[77] J. Kuang, G. Johnson, and M. Fairchild. RIT MCSL high dynamic range image database. Web Site of Munsell Color Science Laboratory. Accessed Dec., 2011.

[78] S. Lee. An efficient content-based image enhancement in the compressed domain using retinex theory. *IEEE Transactions on Circuits Syst. Video Techn*, 17(2):199–213, 2007.

[79] Steven Lehar. *The World In Your Head: A Gestalt View of the Mechanism of Conscious Experience*. Lawrence Erlbaum Associates, Mahwah, NJ, 2003.

[80] X. Li and Y. Zheng. Patch-based video processing: A variational Bayesian approach. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(1):27–40, 2009.

[81] N.-X. Lian, V. Zagorodnov, and Y.-P. Tan. Color image denoising using wavelets and minimum cut analysis. *IEEE Signal Processing Letters*, 12(11):741–744, 2005.

[82] N.-X. Lian, V. Zagorodnov, and Y.-P. Tan. Edge-preserving image denoising via optimal color space projection. *IEEE Transactions on Image Processing*, 15(9):2575–2587, 2006.

[83] C.-H. Ling, Y.-M. Liang, C.-W. Lin, Y.-S. Chen, and H.-Y.M. Liao. Human object inpainting using manifold learning-based posture sequence estimation. *IEEE Transactions on Image Processing*, 20(11):3124–3135, 2011.

[84] C.-H. Ling, C.-W. Lin, C.-W. Su, Y.-S. Chen, and H.-Y.M. Liao. Virtual contour guided video object inpainting using posture mapping and retrieval. *IEEE Transactions on Multimedia*, 13(2):292–302, 2011.

[85] L. Lucchese and S. K. Mitra. A new class of chromatic filters for color image processing. theory and applications. *IEEE Transactions on Image Processing*, 13(4):534–548, 2004.

[86] F. Luisier and T. Blu. SURE-LET multichannel image denoising: Interscale orthonormal wavelet thresholding. *IEEE Transactions on Image Processing*, 17(4):482–492, 2008.

[87] A. Maalouf, P. Carre, B. Augereau, and C. Fernandez-Maloigne. A bandelet-based inpainting technique for clouds removal from remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2363–2371, 2009.

[88] M. Malfait and D. Roose. Wavelet-based image denoising using a Markov random field a priori model. *IEEE Transactions on Image Processing*, 6:549–565, April 1997.

[89] S. G. Mallat. A theory for multiresolution signal decomposition:the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[90] S. G. Mallat. A theory for multiresolution signal decomposition:the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[91] S.G. Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397 –3415, dec 1993.

[92] S. Masnou and J.-M. Morel. Level lines based disocclusion. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 259–263, 1998.

[93] S. C. Matz and R. J. P. de Figueiredo. A nonlinear image contrast sharpening approach based on Munsell's scale. *IEEE Transactions on Image Processing*, 15(4):900–909, 2006.

[94] J. J. McCann. Lessons learned from Mondrians applied to real images and color gamuts. In *Proc. IS&T/SID 7th Color Imag. Conf.*, pages 1–8, 1999.

[95] T. Melange, M. Nachtegael, and E.E. Kerre. Fuzzy random impulse noise removal from color image sequences. *IEEE Transactions on Image Processing*, 20(4):959–970, 2011.

[96] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Philadelphia, 2005.

[97] M. K. Mihcak, K. Ramchandran I. Kozintsev, and P. Moulin. Low-complexity image denoising based on statistical modeling of wavelet coefficients. *IEEE Signal Processing Letters*, 6(12):300–303, 1999.

[98] J. Miral, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Transactions on Image Processing*, 17(1):53–69, 2008.

[99] M.Ma, O.C. Au, S.-H.G. Chan, and M.-T. Sun. Edge-directed error concealment. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(3):382–395, 2010.

[100] J. Mukherjee and S. K. Mitra. Enhancement of color images by scaling the DCT coefficients. *IEEE Transactions on Image Processing*, 17(10):1783–1794, 2008.

[101] D. Mumford. Elastica and computer vision. In C. L. Bajaj, editor, *Algebraic Geometry and Its Applications*, pages 491–506. Springer-Verlag, New York, 1994.

[102] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–684, 1989.

[103] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24(2):227–234, 1995.

[104] P. Ndjiki-Nya, M. Koppel, D. Doshkov, H. Lakshman, P. Merkle, K. Muller, and T. Wiegand. Depth image-based rendering with advanced texture synthesis for 3-d video. *IEEE Transactions on Multimedia*, 13(3):453–465, 2011.

[105] Deanna Needell and Roman Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found. Comput. Math.*, 9:317–334, April 2009.

[106] M. M. Oliveira, B. Bowen, R. McKenna, and Y.-S. Chang. Fast digital image inpainting. In *International Conference on Visualization, Imaging and Image Processing*, pages 261–266, 2001.

[107] Wumo Pan, T.D. Bui, and C.Y. Suen. Text detection from scene images using sparse representation. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1 –5, dec. 2008.

[108] J. Park, D.C. Park, R. J. Marks, and M. A. El-Sharkawi. Recovery of image blocks using the method of alternating projections. *IEEE Transactions on Image Processing*, 14(4):461–474, 2005.

[109] Y. C. Pati, R. Rezaiifar, Y. C. Pati R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proceedings of the 27 th Annual Asilomar Conference on Signals, Systems, and Computers*, pages 40–44, 1993.

[110] K.A. Patwardhan, G. Sapiro, and M. Bertalmio. Video inpainting under constrained camera motion. *IEEE Transactions on Image Processing*, 16(2):545–553, 2007.

[111] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[112] G. Peyre. Texture synthesis with grouplets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):733–746, 2010.

[113] G. Peyre. A review of adaptive image representations. *Selected Topics in Signal Processing, IEEE Journal of*, 5(5):896 –911, sept. 2011.

[114] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising. *IEEE Transactions on Image Processing*, 11(5):545–557, May 2002.

[115] K. Plataniotis and A. N. Venetsanopoulos. *Color Image Processing and Applications*. Springer-Verlag, 2000.

[116] J. Portilla and E. P. Simoncelli. Texture modeling and synthesis using joint statistics of complex wavelet coefficients. In *IEEE Workshop on Statistical and Computational Theories of Vision*, 1999.

[117] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int'l Journal of Computer Vision*, 40(1):49–71, 2000.

[118] J. Portilla, M. Wainwright V. Strela, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, November 2003.

[119] S.D. Rane, G. Sapiro, and M. Bertalmio. Structure and texture filling-in of missing image blocks in wireless transmission and compression applications. *IEEE Transactions on Image Processing*, 12(3):296–303, 2003.

[120] A. M. Rao and D. L. Jones. A denoising approach to multisensor signal estimation. *IEEE Transactions on Signal Processing*, 48(5):1225 – 1234, 2000.

[121] J. Romberg. Imaging via compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):14 –20, march 2008.

[122] R. Rubinstein, A.M. Bruckstein, and M. Elad. Dictionaries for sparse representation modeling. *Proceedings of the IEEE*, 98(6):1045 –1057, june 2010.

[123] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithm. *Physica D*, 60:259–268, 1992.

[124] P. Schallauer, A. Pinz, and W. Haas. Automatic restoration algorithms for 35mm film. *Videre*, 1(3):60–85, 1999.

[125] S. Schulte, V. De Witte, and E.E. Kerre. A fuzzy noise reduction method for color images. *IEEE Transactions on Image Processing*, 16(5):1425–1436, 2007.

[126] P. Sen and S. Darabi. Compressive rendering: A rendering application of compressed sensing. *IEEE Transactions on Visualization and Computer Graphics*, 17(4):487–499, 2011.

[127] L. Sendur and I. W. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on Signal Processing*, 50(11):2744–2756, 2002.

[128] L. Sendur and I. W. Selesnick. Bivariate shrinkage with local variance estimation. *IEEE Signal Processing Letters*, 9(12):438–441, 2002.

[129] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, December 1993.

[130] J. Shen. Inpainting and the fundamental problem of image processing. *SIAM News*, 36(5), Jun 2003.

[131] T. K. Shih, N. C. Tang, J. C. Tsai, and J. Hwang. Video motion interpolation for special effect applications. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 41(5):720–732, 2011.

[132] T.K. Shih, N.C. Tang, and J.-N. Hwang. Exemplar-based video inpainting without ghost shadow artifacts by maintaining temporal continuity. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(3):347–360, 2009.

[133] A. Skodras, C. Christopoulos, and T. Ebrahimi. The JPEG 2000 still image compression standard. *IEEE Signal Processing Magazine*, 18(5):36–58, 2001.

[134] J.-L. Starck, M. Elad, and D.L. Donoho. Redundant multiscale transforms and their application for morphological component analysis. *Advances in Imaging and Electron Physics*, 132:287–348, 2004.

[135] J.-L. Starck, M. Elad, and D.L. Donoho. Image decomposition via the combination of sparse representation and a variational approach. *IEEE Transactions on Image Processing*, 14(10):1570–1582, 2005.

[136] J. L. Starck, F. Murtagh, E. J. Candes, and D. L. Donoho. Gray and color image contrast enhancement by the curvelet transform. *IEEE Transactions on Image Processing*, 12(6):706–717, 2003.

[137] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum. Image completion with structure propagation. In *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 417–424, 2000.

[138] P. Tang, Chiou-Ting Hsu, Chih-Wen Su, T.K. Shih, and H.-Y.M. Liao. Video inpainting on digitized vintage films via maintaining spatiotemporal continuity. *IEEE Transactions on Multimedia*, 13(4):602–614, 2011.

[139] A. Telea. An image inpainting technique based on the fast marching method. *Journal of Graphics, GPU, and Game Tools*, 9(1):23–34, 2004.

[140] B. A. Thomas and J. J. Rodriguez. Wavelet-based color image denoising. In *IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 804–807, September 2000.

[141] R. Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B (Methodological)*, 58(1):267–288, 1996.

[142] J.A. Tropp and S.J. Wright. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6):948 –958, june 2010.

[143] A. Tsai, A. Yezzi, and A. S. Willsky. Curve evolution implementation of the Mumford-Shah functionalal for image segmentation, denoising, dnterpolation, and magnification. *IEEE Transactions on Image Processing*, 10(8):1169–1186, 2001.

[144] D. Tschumperle and R. Deriche. Vector-valued image regularization with PDE's: A common framework for different applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):506–517, 2005.

[145] G. Turk. Texture synthesis on surfaces. In *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 347–354, 2001.

[146] A. van Nevel. Texture synthesis via matching first and second order statistics of a wavelet frame decomposition. In *IEEE International Conference on Image Processing (ICIP)*, volume 1, pages 72–76, 1998.

[147] K. V. Velde. Multi-scale color image enhancement. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 584–587, 1999.

[148] I. Verstegen. *Arnheim, Gestalt, and art : a psychological theory.* Springer Verlag, New York, 2005.

[149] L. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293, 2002.

[150] L. A. Vese and S. J. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19(1-3):553–572, 2003.

[151] L. Wang, L. He, A. Mishra, and C. Li. Active contours driven by local Gaussian distribution fitting energy. *Signal Processing*, 89(12):2435–2447, 2009.

[152] Y. Wang and Q.-F. Zhu. Error control and concealment for video communication: a review. *Proceedings of the IEEE*, 86(5):974 – 997, 1998.

[153] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 479–488, 2000.

[154] L.-Y. Wei and M. Levoy. Texture synthesis over arbitrary manifold surfaces. In *ACM International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 355–360, 2001.

[155] Y.-W. Wen, R. H. Chan, and A. M. Yip. A primal-dual method for total-variation-based wavelet domain inpainting. *IEEE Transactions on Image Processing*, 21(1):106–114, 2012.

[156] B. Wohlberg. Inpainting by joint optimization of linear combinations of exemplars. *IEEE Signal Processing Letters*, 18(1):75–78, 2011.

[157] J. Wright, Yi Ma, J. Mairal, G. Sapiro, T.S. Huang, and Shuicheng Yan. Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98(6):1031 –1044, june 2010.

[158] G.-L. Wu, C.-Y. Chen, and S.-Y. Chien. Algorithm and architecture design of image inpainting engine for video error concealment applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(6):792–803, 2011.

[159] H. Xiong, Y. Xu, Y.F. Zheng, and C. W. Chen. Priority belief propagation-based inpainting prediction with tensor voting projected structure in video compression. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(8):1115–1129, 2011.

[160] Z. Xiong, X. Sun, and F. Wu. Block-based image compression with parameter-assistant inpainting. *IEEE Transactions on Image Processing*, 19(6):1651–1657, 2010.

[161] Z. Xu and J. Sun. Image inpainting by patch propagation using patch sparsity. *IEEE Transactions on Image Processing*, 19(5):1153–1165, 2010.

[162] A. C. Yau, X.-C. Tai, and M. K. Ng. L0-norm and total variation for wavelet inpainting. In *2nd Int. Conf. Scale Space & Variational Methods in Computer Vision (SSVM'09)*, pages 539–551, 2009.

[163] S. Yoo and R.-H. Park. Red-eye detection and correction using inpainting in digital photographs. *IEEE Transactions on Consumer Electronics*, 55(3):1006–1014, 2009.

[164] H. Yu, L. Zhao, and H. Wang. Image denoising using trivariate shrinkage filter in the wavelet domain and joint bilateral filter in the spatial domain. *IEEE Transactions on Image Processing*, 18(10):2364–2369, 2009.

[165] G. Zhai, X. Yang, W. Lin, and W. Zhang. Bayesian error concealment with DCT pyramid for images. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(9):1224–1232, 2010.

[166] J. Zhang, D. Wang, and Q.N. Tran. A wavelet-based multiresolution statistical model for texture. *IEEE Transactions on Image Processing*, 7(11):1621–1627, 1998.

[167] X. Zhang and T. F. Chan. Wavelet inpainting by nonlocal total variation. *Inverse Problem and Imaging*, 4(1):191–210, Feb. 2010.

[168] S.-C. Zhu, C. Guo, Y. Wang, and Z. Xu. What are textons? *International Journal of Computer Vision*, 62(1-2):121–143, 2005.