ABSTRACT

COLOR IMAGE QUALITY MEASURES AND RETRIEVAL

by
Yan-Yu Fu

The focus of this dissertation is mainly on color image, especially on the images with lossy compression. Issues related to color quantization, color correction, color image retrieval and color image quality evaluation are addressed. A no-reference color image quality index is proposed. A novel color correction method applied to low bit-rate JPEG image is developed. A novel method for content-based image retrieval based upon combined feature vectors of shape, texture, and color similarities has been suggested. In addition, an image specific color reduction method has been introduced, which allows a 24-bit JPEG image to be shown in the 8-bit color monitor with 256-color display. The reduction in download and decode time mainly comes from the smart encoder incorporating with the proposed color reduction method after color space conversion stage. To summarize, the methods that have been developed can be divided into two categories: one is visual representation, and the other is image quality measure.

Three algorithms are designed for visual representation:

(1) An image-based visual representation for color correction on low bit-rate JPEG images. Previous studies on color correction are mainly on color image calibration among devices. Little attention was paid to the compressed image whose color distortion is evident in low bit-rate JPEG images. In this dissertation, a lookup table algorithm is designed based on the loss of PSNR in different compression ratio.

(2) A feature-based representation for content-based image retrieval. It is a concatenated vector of color, shape, and texture features from region of interest (ROI).

(3) An image-specific 256 colors (8 bits) reproduction for color reduction from 16 millions colors (24 bits). By inserting the proposed color reduction method into a JPEG encoder, the image size could be further reduced and the transmission time is also reduced. This smart encoder enables its decoder using less time in decoding.

Three algorithms are designed for image quality measure (IQM):

(1) A referenced IQM based upon image representation in very low-dimension. Previous studies on IQMs are based on high-dimensional domain including spatial and frequency domains. In this dissertation, a low-dimensional domain IQM based on random projection is designed, with preservation of the IQM accuracy in high-dimensional domain.

(2) A no-reference image blurring metric. Based on the edge gradient, the degree of image blur can be measured.

(3) A no-reference color IQM based upon colorfulness, contrast and sharpness.

COLOR IMAGE QUALITY MEASURES AND RETRIEVAL

by
Yan-Yu Fu

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Computer Science

Department of Computer Science

January 2006

APPROVAL PAGE

COLOR IMAGE QUALITY MEASURES AND RETRIEVAL

Yan-Yu Fu

Dr. Frank Y. Shih, Dissertation Advisor — Date
Professor of Computer Science, NJIT

Dr. James McHugh, Committee Member — Date
Professor of Computer Science, NJIT

Dr. Yun-Qing Shi, Committee Member — Date
Professor of Electrical and Computer Engineering, NJIT

Dr. Artur Czumaj, Committee Member — Date
Associate Professor of Computer Science, NJIT

Dr. Qun Ma, Committee Member — Date
Assistant Professor of Computer Science, NJIT

Dr. Cristian Borcea, Committee Member — Date
Assistant Professor of Computer Science, NJIT

BIOGRAPHICAL SKETCH

Author: Yan-Yu Fu

Degree: Doctor of Philosophy

Date: January 2006

Education:

Doctor of Philosophy in Computer Science,
New Jersey Institute of Technology, Newark, NJ, January 2006

Master of Science in Computer Science
New Jersey Institute of Technology, Newark, NJ, January 1996

Bachelor of Arts in English Literature
National Central University, Chung-Li, Taiwan, June 1982

Major: Computer Science

Dissertation Advisor: Professor Frank Y. Shih

Publications:

F. Y. Shih, K. Zhang and Y. Fu,
"A hybrid two-phase algorithm for face recognition,"
International Journal of Pattern Recognition and Artificial Intelligence, vol. 18, no. 8, pp. 1423-1435, Dec. 2004.

F. Y. Shih, Y. Fu and K. Zhang,
"Multi-view face identification and pose estimation using B-spline interpolation,"
Information Sciences, vol. 169, no 3, pp. 189-204, Feb. 2005.

Frank, Y. Shih, and Yan-Yu Fu,
"Color correction for color distorted images,"
Proceedings of the Third IASTED International Conference on CIRCUITS, SIGNALS, and SYSTEMS (CSS 2005), pp. 206-211, Marina del Rey, CA, USA, Oct. 24-26, 2005.

F. Y. Shih and Y. Fu,
"Approximate image quality in low-dimensional domain based on random projection," submitted to *International Journal of Pattern Recognition and Artificial Intelligence*, 2005.

Presentations:

Yan-Yu Fu,
"JPEG image enhancement,"
The 1st Graduate Student Seminar, NJIT, Dec. 2003.

Yan-Yu Fu,
"Dimension reduction and computational image quality,"
The 2nd Graduate Student Seminar, NJIT, Nov. 2004.

Frank, Y. Shih, and Yan-Yu Fu,
"Color correction for color distorted images,"
The Third IASTED International Conference on CIRCUITS, SIGNALS, and SYSTEMS (CSS 2005), Marina del Rey, CA, USA, October 24-26, 2005.

To my beloved family members

ACKNOWLEDGMENT

I would like to express my deepest appreciation to Prof. Frank Y. Shih, who not only academically served as my dissertation advisor, providing valuable and countless resources, insight, and intuition, but spiritually gave me support and encouragement throughout the doctoral years. Especially, I would like to express my gratitude for my advisor's patience that allowed me an enormous degree of freedom to explore all kinds of ideas and possibilities, without which this research would not have been a reality. Special thanks are given to the committee members for their active participation. In particular, Prof. McHugh suggested me to increase the dimension of the feature vector for shape for boosting the query performance. Prof. Shi gave me an idea to conduct a literature survey in order to understand on what types of images the PSNR may fail. Prof. Ma guided me on a theme to help focus many widespread ideas. Prof. Borcea guided me a direction for a possible wireless application for the smart encoder. The random projection method was originally inspired from one of the papers in Prof. Czumaj's class - Advanced Algorithms: Approximation and Randomized Algorithms.

All of my fellow graduate students in the Computer Vision Laboratory are deserving of my recognition for their support. I would like to give special thanks to Kai Zhang for his technical assistance in MATLAB programming. Meanwhile, I would like to express my gratitude to Patrick Shih. With his suggestions on color space related issues, he inspired me some early topics and methods.

Last, but not least, I would like to salute to my mother for her consistent spiritual support during my PhD years. I am a true believer of her saying that "In order to make your dream come true, you must have passion, effort, and perseverance."

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivations and Problem Statement

Color vision involves in three parameters, namely, objects, lighting, and the observer. Human being, as an observer, tries to process and understand color and brightness information from some chosen properties on top of the object. Recently, a paradigm shift on color vision research has taken place. It is the shift from knowledge-driven to application-driven. It is easy to see the knowledge-driven trend through some historical origin of color research. From Newton's study of the physical nature of light on which the perception of color is based, through Young's human perception on which three principal colors, yellow, blue, and red, are based, all the way to Maxwell's calculation of the relative position of color on a color surface, the motivation was driven by knowledge itself. However, nowadays, especially during the past two decades, color vision research has become solely driven by application's needs. Problem domain is usually associated with the invention of a certain digital device, or data transmission and conversion between two digital devices; for example, what we see is what we print, that relies on the color correction between monitor and printer. Observing this paradigm shift, we have to face the application-driven problems and device-driven problems. Two major problems relating to color capture and display device now are color image quality measuring and visual representation problems for a specific application domain. The former problem is important than the latter one. Once the former problem is solved, consequently a great number of application benchmark problems could be solved.

One of the motivations of this dissertation is to develop color image quality measure tailoring to human perception in general, and to develop image quality measure (IQM) in very low dimensional domain, in particular. Past study on IQMs is based on high-dimensional domain including spatial domain and frequency domain; however, for the image of large size and applications like image retrieval, efficiency is the first concern. There is a need to reduce image in high-dimension to low-dimension to obtain image quality and its category. Our goal is to design low-dimensional domain IQM, while still preserving the accuracy of IQM in high-dimensional domain.

In industry, color research is always driven by application's needs. For instance, color printers need the color correction method on color monitor input; Metropolitan Museum of Art needs color reproduction for historical artworks; digital camera company needs the design of a brand new color space for color images on natural scenes; Hewlett Packard needs to develop a color quantization method tailoring to the limited memory of some embedded image capture systems. To these ends, another motivation is to come up with a novel method for color correction, color quantization, and color image retrieval.

1.2 Conceptual Framework

Three conceptual frameworks are presented in this section. Framework 1 is a conceptual model for color image quality measure; Framework 2 is a no-reference model for blurring artifact detection, and Framework 3 is a conceptual model for dimensionality reduction.

1.2.1 A Conceptual Model for Color Image Quality Measure

One of the disadvantages of the referenced color image quality metrics like Peak Signal to Noise Ratio (PSNR) is that it always requires an original image as the reference image. Meanwhile, PSNR fails to correlate to human perception in several situations. To alleviate these problems, the first conceptual model we build is the No-Reference Color Image Quality Measure (NRCIQM) model. With this model, three input parameters are extracted from the input image. They are sharpness, contrast, and colorfulness. Our goal is to come up with a linear regression model for this conceptual model, and to calculate the relative weight with respect to each parameter from the experiments.

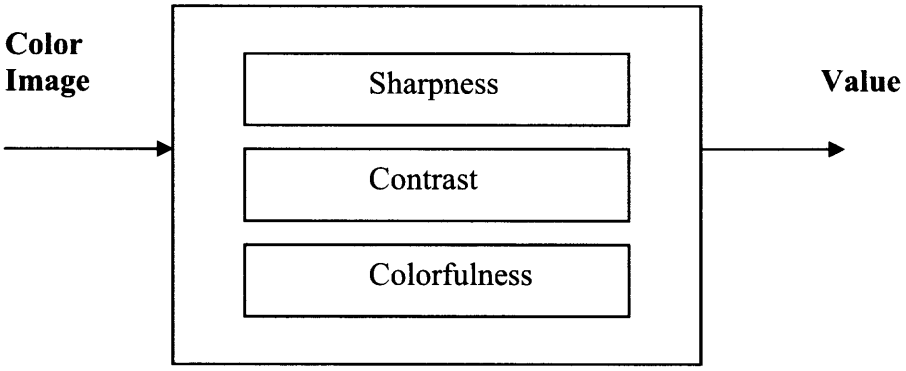


Figure 1.1 A conceptual model for color image quality measure.

1.2.2 A No Reference Edge Model for Blurring Artifact Detection

Our No-Reference Edge Model (NREM) takes only one input, the decoded image, and outputs a decision as well as a value. The decision is to judge the image either having blurring artifact or ringing artifact; whereas the value indicates its severity. The processing black box has two modules, Edge Continuity Measure and False Edge Measure. The output decision will be based upon a linear combination of these two results with proper weights.

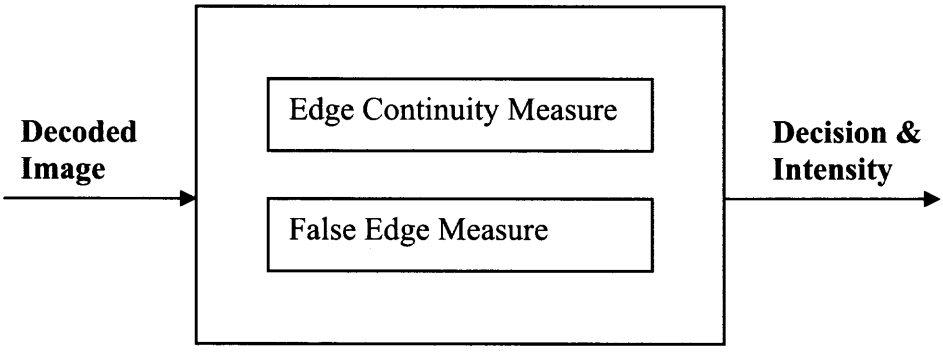


Figure 1.2 A no-reference edge model for blurring artifact detection.

1.2.3 Entity Relationship Model for Dimensionality Reduction

Definition 1: **Entity** is a high dimensional space with its coordinate in a vector or a matrix, such as image. After projection, preserving the entity means preserving the appearance, which includes the shape, the texture, and the edges of the original image.

Definition 2: **Relationship** is defined as the **Euclidean distance** between each pair of Entity; it is commonly used as **L2-Norm.**

Lemma 1: For a given pair of entity, after dimensional reduction or increase, one of the following cases will occur:

(1) Both Entity and Relationship preserve.

(2) Entity does NOT preserve, Relationship preserve.

(3) Both Entity and Relationship do NOT preserve.

(4) Entity preserve, Relationship does NOT preserve.

The case we desire is case 2. That is, the only thing we have in mind is to preserve the

pairwise distance after projection. This property may lead to some breakthrough in applications in aspect of time complexity. Where are we now? The large puzzle is dimension reduction for high-dimensional data; the small puzzle is the area of case 2. What do we plan to do? This dissertation strives for finding the appropriate methods in the case 2.

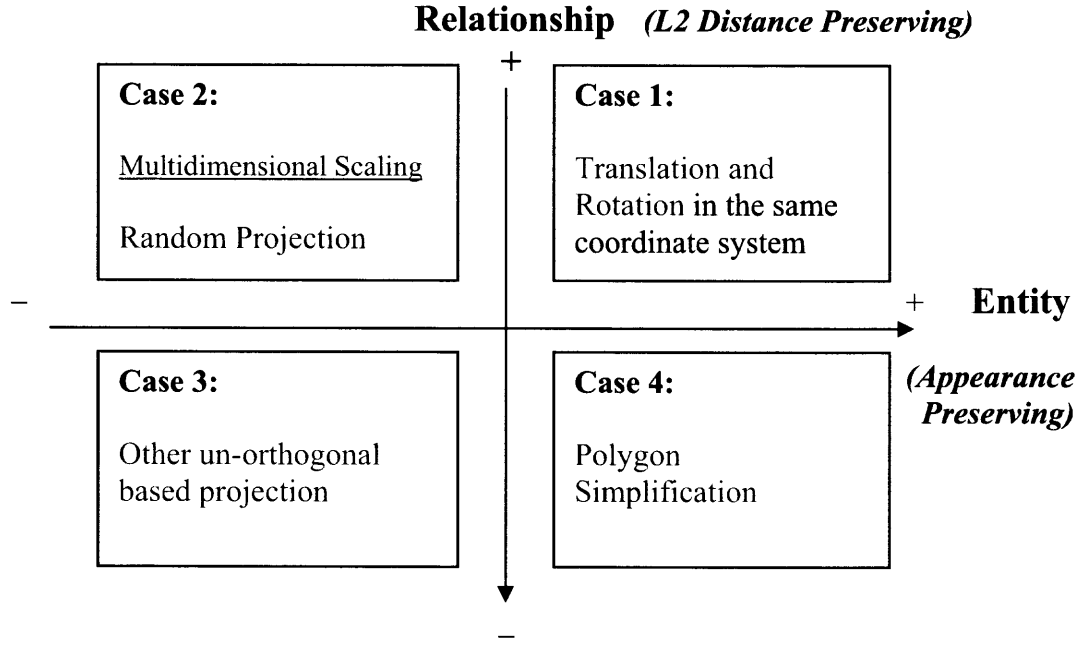


Figure 1.3 ER Model for dimensionality reduction;
Case 2 is the small puzzle inside the large puzzle.

In order to preserve the pairwise distance from high dimension to low dimension, ideally we may want a least distortion in their relationship; practically we have to admit that there are not so many math tools (i.e. MDS and Random Projection) existing by far. Hence, any mathematical method to be invented later on in the second quadrant (case 2) will have a better chance to lead to a breakthrough in data mining and machine learning.

In general, other than depicting the ER model for dimension reduction statically, we have to specify some operations so that it is easy for researchers to change quadrant dynamically. First of all, fixing the Y-axis, a positive appearance can be changed to a negative appearance by any projection that does not preserve the shape. Second of all, fixing the X-axis, a positive relationship can be altered into a negative relationship by any projection that is un-orthogonal-based. In particular, to switch from case 1 to case 2, any projection based on the orthogonal basis can be used; to switch from case 1 to case 3, any projection based on the un-orthogonal basis can be used; to switch from case 1 to case 4, any filter or projection with shape-preserving property can be applied.

Traditional dimension reduction methods, such as PCA, LDA, SVD, have their relative pros and cons. Despite of their relative strength, at least one con they share in common is that the dimensionality reduction process is very time consuming. Our proposed method is random projection, but for certain application like IR and ML, it is inevitable to combine support vector machine for further dimension reduction, especially when in case two concept is interleaving with each other, dimensionality increase by using either linear or nonlinear kernel will be needed.

1.3 Dissertation Organization

This dissertation is organized as follow:

Chapter 2 discusses the referenced Image Quality Measure (IQM) study from efficiency viewpoint. We have created a novel image representation based on random projection to speed up the scoring process. Using this sparse representation, in general speed can be from N reduced to $log\ N$, where N is the size of image set.

Chapter 3 studies the problem of compression artifacts due to JPEG/JPEG2000. Our focus are blurring artifact. In this chapter, we have designed a blurring artifact detection method. Not only the artifact can be detected, but the degree can be measured. The idea of the method is based on the discontinuity of edges. The more discontinuous the edge is, the more blurring the image is.

Chapter 4 studies the known results on no-reference quality measures. We have developed a no-reference quality index for color images. This quality index is determined by a linear regression model of three parameters: sharpness, contrast, and colorfulness.

Chapter 5 analyzes the color correction problem with respect to color distorted images in general, and low-bit rate JPEG images in particular. Apart from the traditional linear and nonlinear methods, we have designed a lookup table method to correct color distortions based upon its compression ratio for the last time.

Chapter 6 performs a concise survey on content-based image retrieval problem, especially pays heed to those methods combining shapes and colors. We have invented a hybrid method based on the mixture features from DC image histogram, chain code histogram and color histogram.

Chapter 7 discusses the intelligent color quantization problem. We developed two methods. One is tailoring to image-specific color reduction, which can be incorporating to the encoder to reduce colors from 24 bits to 8 bits using the careful selected color palette according to the classification of spatial image content in DCT domain. The other method is designed to enhance the quality of the quantization results.

Chapter 8 summarizes all the proposed methods and their corresponding contributions, and outlines the short-term and long-term future research directions.

CHAPTER 2

A REFERENCED QUALITY INDEX IN LOW-DIMENSIONAL DOMAIN BASED ON RANDOM PROJECTION

2.1 Definitions and Notations

Traditionally, Image Quality Measures (IQMs) are mainly conducted in the spatial domain, although with rare exceptions in the frequency domain [Eskicioglu and Fisher, 1995; Avcibas, Sayood, and Sankur, 2002]. Recently, with the advent of compression technology, IQMs in the compressed domain start to gain attentions in which they measure compression artifacts in the first place. This kind of IQM is used to automatically measure the degree of image artifacts such as blocking and ringing effects [Franti, 1998; Marziliano, et al., 2002]. It is often calculated in the spatial domain, which we call the high-dimensional domain. Some IQMs can be conducted in both spatial and compressed domains [Fuhrmann, et al., 1995; Wang and Bovik, 2002].

There are two criteria in evaluating an IQM: accuracy and efficiency. Accuracy indicates how close an IQM is to human vision, and efficiency indicates the time complexity in calculating an IQM. In this chapter, we emphasize on the efficiency. Lately numerous image representation methods, such as curvelet, edgelet, [Donoho, et al., 1999] and ridgelet transform [Do and Vetterli, 2003], have been developed and their efficiency have been shown. Inspired by these novel image representation methods, we reduce image quality problem to image representation problem. Our idea is to project an image in a high-dimensional space into a low-dimensional space to calculate the IQM. We use random projection in the randomized algorithm for projection [Johnson and Lindenstrauss, 1984; Dasgupta and Gupta, 2003]. In L2 norm, when the data are

projected onto a low-dimensional space by random projection, the original pair-wise distance is preserved. [Indyk and Motwani, 1998; Bingham, 2001; Achlioptas, 2001] The detailed proof is in Appendix A.

Definition 1: *Quality Score* (QS) is the score obtained from an IQM. If it is obtained from the original image, we call *High-Dimensional Quality Score* (HDQS). If it is obtained from the low-dimensional image after random projection, we call *Low-Dimensional Quality Score* (LDQS).

Definition 2: *Quality Index* (QI) is the index obtained from a fuzzy function whose input is the QS and output is between 0 and 1. If it is obtained from HDQS, we call *High-Dimensional Quality Index* (HDQI). If it is obtained from LDQS, we call *Low-Dimensional Quality Index* (LDQI).

Definition 3: The L2-norm, also known as the *Euclidean norm*, is the magnitude of a vector.

Definition 4: Let $f(i,j)$ denote an image of size $N \times N$ and $\hat{f}(i,j)$ denote its distorted image. The *Peak Signal-to-Noise Ratio* (PSNR) is defined as:

$$PSNR = 20 * \log_{10}(\frac{255}{\sqrt{MSE}})$$

$$MSE = \frac{\sum[f(i,j) - \hat{f}(i,j)]^2}{N^2}$$

Definition 5: Let $D(x,y)$ denote Euclidean distance between x and y, and let $D'(f(x), f(y))$ denote Euclidean distance between their transformations, $f(x)$ and $f(y)$. The *expansion*, *contraction*, and *distortion* are respectively defined as:

$$expansion = \max_{x,y \in X} \frac{D'(f(x), f(y))}{D(x, y)} \quad (1)$$

$$contraction = \max_{x,y \in X} \frac{D(x, y)}{D'(f(x), f(y))} \quad (2)$$

$$distortion = expansion \times contraction \quad (3)$$

Since we are concerned with the relative comparisons, the relative distortion between the original Euclidean distance D and the random-projected Euclidean distance D' is developed. The relative distortion between LDQS and HDQS is calculated as:

$$distortion = \frac{LDQS - HDQS}{HDQS} \quad (4)$$

2.2 The Proposed Method: Low-Dimensional Quality Index

The procedures of calculating the LDQI are presented as follows:

1. Given N images of size $n \times m$ (i.e., D dimensions, where $D = n \times m$), we calculate the target dimension $D' = log\ N / \varepsilon^2$, where ε denotes the distortion factor.
2. For each image X, we perform the random projection as:
   1) Select D' random vectors from D.
   2) Let $R = (r_1, r_2, ..., r_D)$ whose variance is one.

      If ($r_i \leq 1/6$), then $r_i = \sqrt{3}$;

      else if ($r_i \geq 5/6$), then $r_i = -\sqrt{3}$;

      else $r_i = 0$;

   3) Project X onto D' to obtain $F(X) = (x_1, x_2, ..., x_{D'})$.

3. Calculate the quality score PSNR($F(X)$).

4. Transform $PSNR(F(X))$ by the fuzzy function (FF) to obtain QI (Quality Index) as shown in Figure 1.

5. The *Fuzzy Function* (FF) is defined as:

$$\begin{aligned} FF(\ PSNR(X)\ ) &= 0, && \text{if } PSNR(X) \le 20; \\ &= (PSNR(X) - 20) / 20, && \text{if } 20 < PSNR(X) \le 40; \\ &= 1, && \text{otherwise.} \end{aligned} \quad (5)$$

Therefore, we obtain the low-dimensional quality index as:

$LDQI = FF(\ PSNR(\ F(X)\ )\ )$, where $F(X)$ is the low-dimensional projection of X.

Since the random projection preserves the pair-wise distance of N points from high dimensions to $\log N$ dimensions, we have $d(f(i,j) - \hat{f}(i,j)) \cong d(g(i,j) - \hat{g}(i,j))$, where $f(i,j)$ is the original image; $\hat{f}(i,j)$ is the distorted image; $g(i,j)$ is the projected image from $f(i,j)$; $\hat{g}(i,j)$ is the projected image from $\hat{f}(i,j)$.

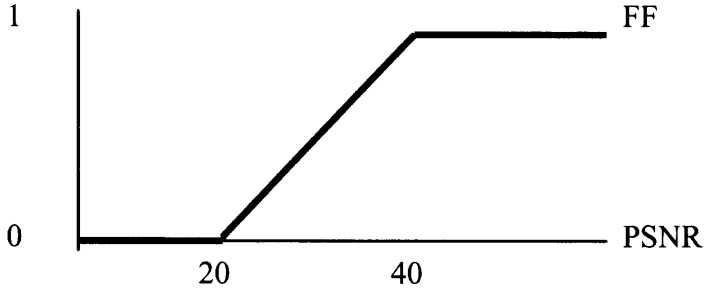
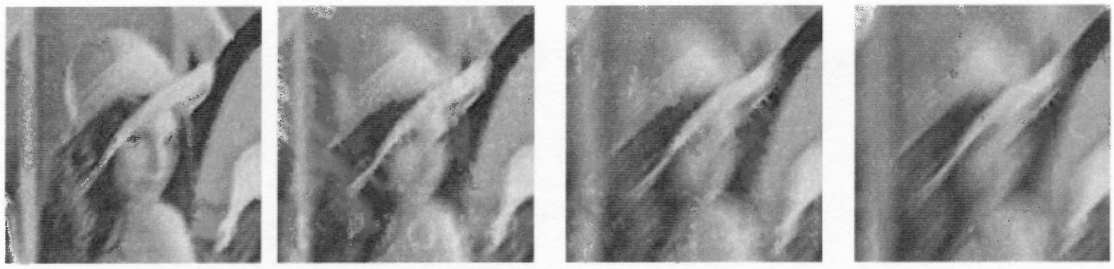


Figure 2.1 The fuzzy function (FF) mapping from PSNR to QI.

2.3 Experimental Design

We test our proposed method on 40 images of size 256 × 256. Each image is conducted by 8 motion blurs, 8 Gaussian blurs, and 8 compression blurs as illustrated in Figure 2. Totally, there are N = 960 images. The image contents are natural scenes or person images. For motion blur, we perform 8 velocity vectors; all have 45° with magnitudes increasing from 5 to 40 at an interval of 5. For Gaussian blur, we perform 8 point spread functions (PSFs) with window sizes increasing from 3×3 to 17×17 at odd numbers. For compression blur, we apply 8 compression ratios in the range of 1:20, 1:40, 1:60, …, 1:160.

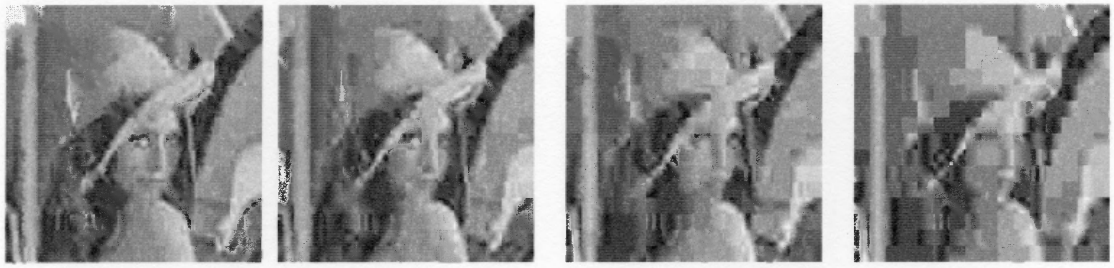
(a)



(b1) (b2) (b3) (b4)

Figure 2.2 (a) Original image;
(b) Motion blur: 45° with magnitudes ranging from 5, 10, 15, , 40, step 5.

Figure 2.2 continued. (c) Gaussian blur with window sizes of 3×3, 5×5, 7×7, …, 17×17; (d) Compression blur with compression ratio of 1:20, 1:40, 1:60, 1:80.



(d5) (d6) (d7) (d8)

Figure 2.2 continued. (d) Compression blur with compression ratio of 1:100, 1:120, 1:140, 1:160.

We perform the sparse representation of the IQM as follows:

1. We call the original image and its blur image as the dense representation. We use the L2 distance and the PSNR to obtain two quality indexes: $HDQS_{L2}$ and $HDQS_{PSNR}$.
2. Perform the random projection on both the original and blur images to create the sparse representation. An example is shown in Figure 3.
3. Calculate the L2 distance and the PSNR on the sparse representation to obtain two quality indexes: $LDQS_{L2}$ and $LDQS_{PSNR}$.
4. Calculate the L2 and PSNR relative errors between dense and sparse representations as:

$$\text{L2 Relative Error} = \frac{LDQS_{L2} - HDQS_{L2}}{HDQS_{L2}} \tag{6}$$

$$\text{PSNR Relative Error} = \frac{LDQS_{PSNR} - HDQS_{PSNR}}{HDQS_{PSNR}} \tag{7}$$

Our experimental results are shown in Figures 2.4-2.6. We use the L2 in Figure 4. The results show that the relative errors of Gaussian blur after random projection are about the same as blurs increase; however, the relative errors of motion and compression blurs decrease slowly. Since the relative errors of all three blurs are very small, we obtain that the image sparse representation measured by the L2 can preserve the pair-wise distance.

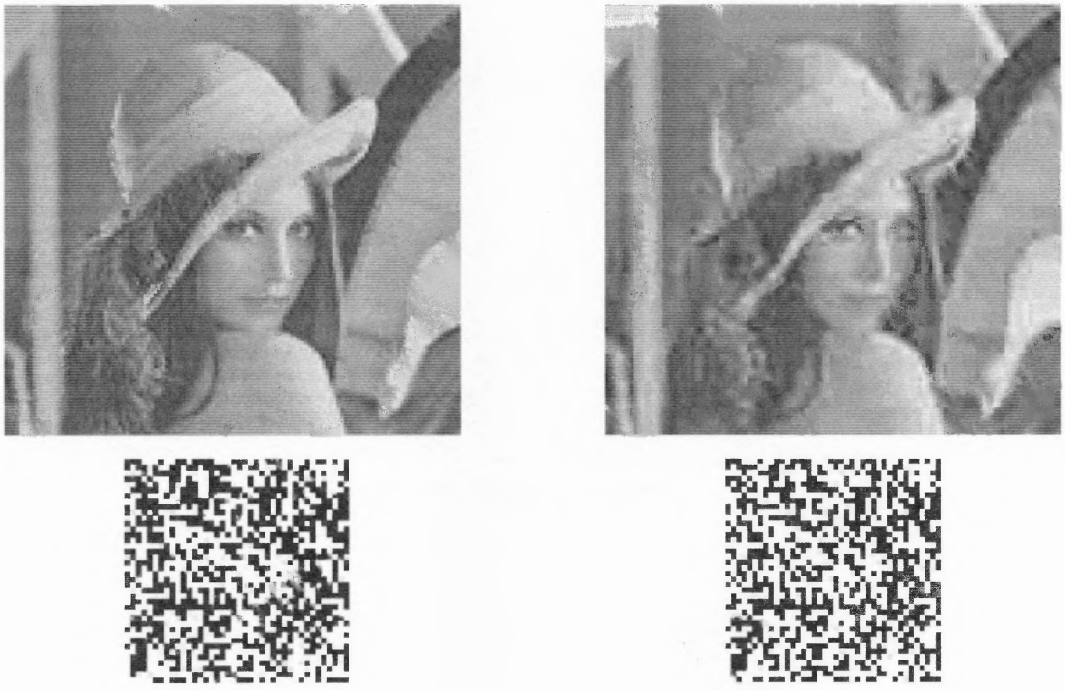


Figure 2.3 An example of dense and sparse representations.
Upper row: Lena image (256×256) and its 1:40 compressed image.
Lower row: the images after random projection into 40×40.

2.4 Experimental Results and Discussions

We use the PSNR in Figure 5. The results show that the relative errors after random projection remain about the same as Gaussian blurs increase; the relative errors of motion blur first decrease and then increase when the magnitude is larger than 10; the relative errors of compression blur decrease until the compression ratio of 1:100 and turn

increasing slightly. We observe that the sparse representation of compression blur shows the desired property in which the errors decreases as blurs increase. The reason is due to the fact that the dense representation in high compression ratio has already lost most of high-frequency information, so that the sparse representation has not much to lose anymore. Therefore, the pair-wise distance between dense and sparse representations becomes closer. Since the sparse representation of the JPEG image after random projection has very small distortion, we can use it for image clustering and indexing. Using the sparse representation of the JPEG image, our QI can gain the speed advantage of O(log N) as opposed to O(N) in the conventional method.

The results in Figure 6 and Tables 1-3 show slightly high QI distortions for three types of blur. The reasons are the QI distortion inherits the PSNR distortion and the QI adds up the distortion due to the fuzzy function. Still the QI is suitable for measuring the compression blur because of its relatively low distortion. We observe that the relative errors are about 0.15 as the compression blurs increase.

Table 2.1 Quality Index Distortion on Motion Blur

(Note: Each magnitude calculates the mean distortion of 40 images.)

magnitude	5	10	15	20	25	30	35	40
Motion	0.1113	0.0071	0.0601	0.1075	0.1482	0.1835	0.2067	0.2334

Table 2.2 Quality Index Distortion on Gaussian Blur

(Note: Each window size calculates the mean distortion of 40 images.)

window size	3x3	5x5	7x7	9x9	11x11	13x13	15x15	17x17
Gaussian	0.2679	0.2706	0.2688	0.2665	0.2667	0.2688	0.2673	0.2665

Table 2.3 Quality Index Distortion on Compression Blur

(Note: Each compression ratio calculates the mean distortion of 40 images.)

compress ratio	1:20	1:40	1:60	1:80	1:100	1:120	1:140	1:160
Compression	0.2415	0.1525	0.0878	0.0383	0.0047	0.0505	0.1054	0.136

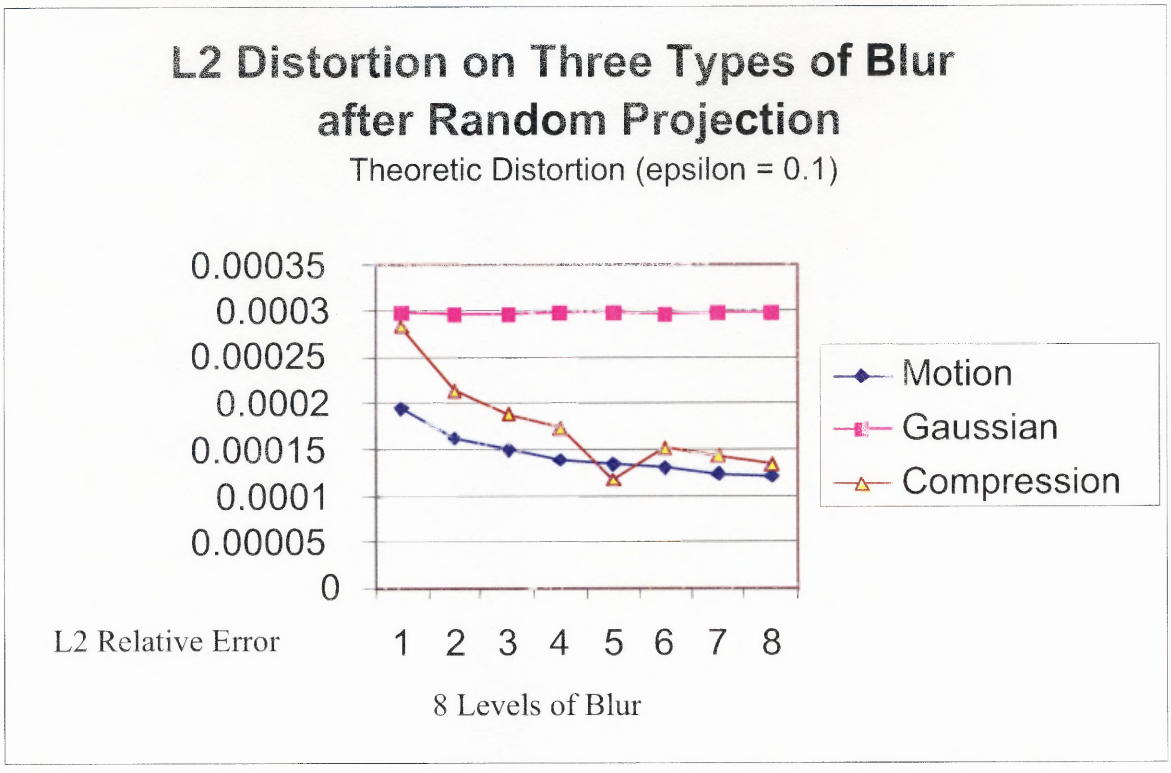


Figure 2.4 The L2 distortion on three types of blur after random projection ($\varepsilon = 0.1$).

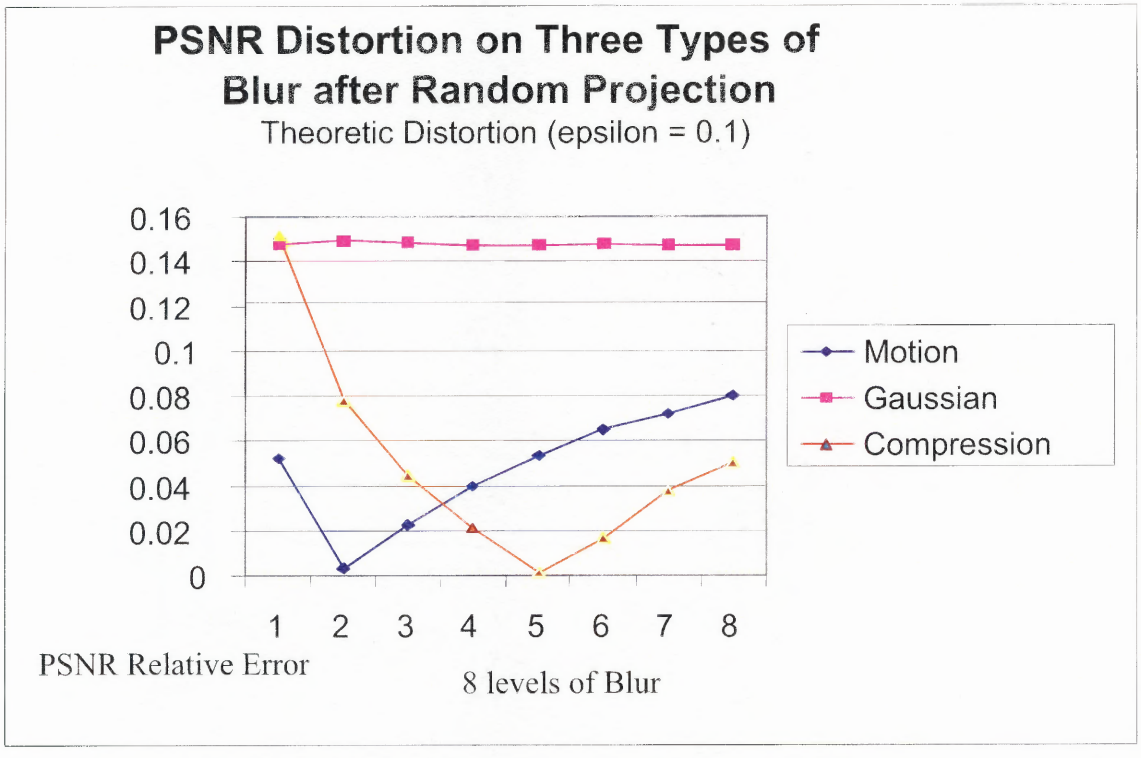


Figure 2.5 PSNR distortion on three types of blur after random projection ($\varepsilon = 0.1$).

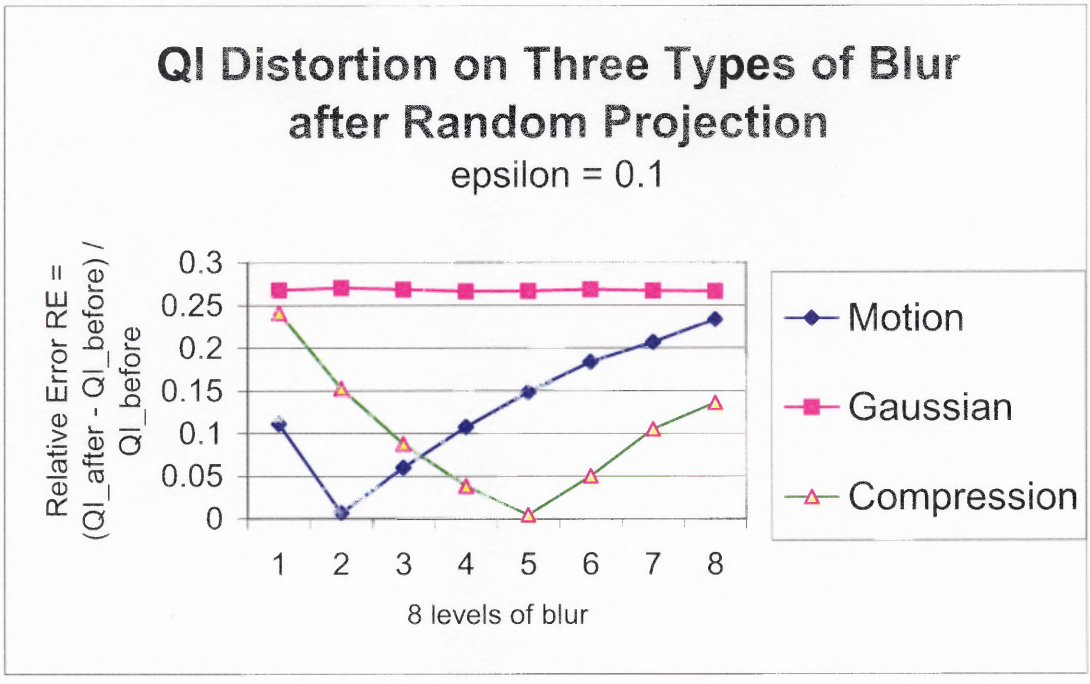


Figure 2.6 QI distortion on three types of blur after random projection ($\varepsilon = 0.1$).

CHAPTER 3

A NO-REFERENCE METRIC FOR IMAGE BLURRING DUE TO COMPRESSION AND NOISE

3.1 Definitions and Notations

There are two kinds of image compression today: **lossy** and **lossless**. Lossless image compression will not allow any information loss when the image is restored, where as lossy image compression takes advantage of limitations of human visual system and throws away some information which people anyways wouldn't easily see; this incurs the compression artifacts. There are numerous artifacts due to lossy compression, such as blocking artifact, blurring artifact, ringing artifact as well as color distortion.

What is compression artifact? It is the distortions incurred from the decompression procedure. Since the distortion is undesired visually, it is also called artifact. For simplicity, we only address three kinds of compression artifacts which were most commonly seen.

Blocking artifact is the artifact due to small square 8x8 blocks all over the images, known both in JPEG and in video decoders; this is mainly due to the hardware burden by design. Blurring artifact means that the decompressed image is smoother than originally. It is due to the effect that two many high frequency values have been taken away. Ringing artifact is the effect that along the edges there are a great deal of oscillating ripples; this is seen typically from the JPEG2000 images.

A lot of artifact detection and removal algorithms have been proposed till now. In general, all the methods can be divided into two categories. One is the spatial domain

methods [Derviaux, et al., 1997]; the other is the frequency domain methods [Bovik and Liu, 2002]. Since too many methods proposed for the blocking artifact removal; we want to pay our attention mainly in the blurring and ringing artifacts.

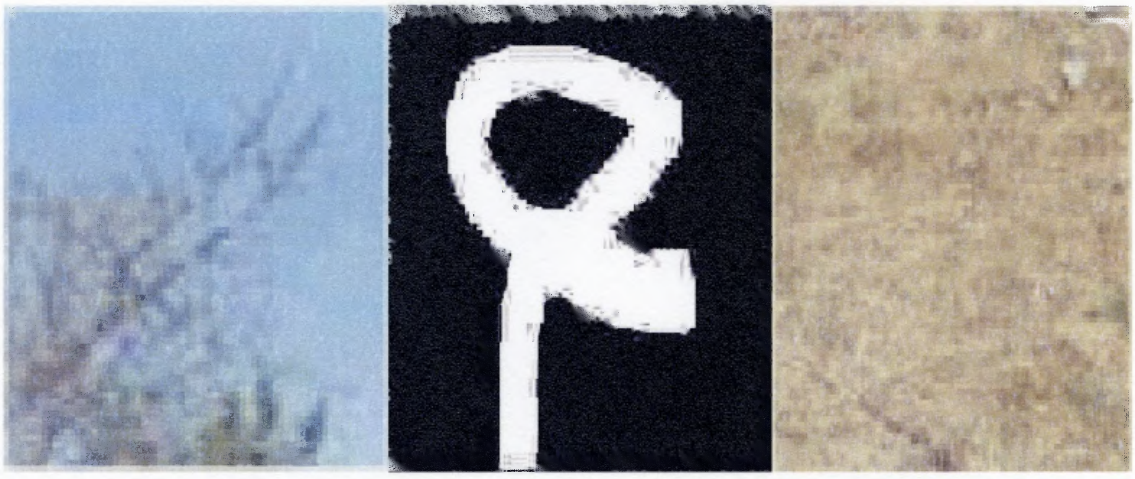


Figure 3.1 From left to right: blocking artifact; ringing artifact; blurring artifact.

3.2 Introduction: Overview of Image Compression Artifacts

In the compressed domain, image blurring is due to the attenuation of the high frequency pixels. The JPEG or JPEG2000 artifacts include distortions due to blurring, ringing and blocking. Up to now, literatures have paid a great deal of attention to the blocking and ringing artifacts removal; however, the blurring artifact has not yet been investigated thoroughly. The unbalanced phenomenon leads to the problem of some unbalanced metric for image quality assessment.

Image-blurring related issues include blur identification [Pavlovic and Tekalp, 1992; Devcic and Loncaric, 2000], blur estimation [Kayargadde and Martens, 1996; Elder and Zucker, 1998], image deblurring [Carasso, 1999] and blind deconvolution [Kundur and Hatzinakos, 1996]. In theory, all blurring and deblurring processes use the

calculation of the point spread function. In practice, most of these methods require iteratively applying algorithms, which are computationally demanding. For a decade, the notion of image blurring has stayed in the areas, such as blurred by motion, blurred by defocus, and blurred by noise. About the same time, the issue of JPEG's blurring artifact was raised and examined, but no blur metric was proposed, even though there exist a lot of perceptual quality metrics based on the blocking effect [Eckert and Bradley, 1998; Chandra and Ellis, 1999]. Not until recently, the blurring metric was proposed in the area of compressed domain images [Marziliano, Dufaux, Winkler, and Ebrahimi, 2004].

Marziliano et al. [Marziliano et al., 2002] proposed a blur metric based on analysis of the spread of edges in an image. This metric works well for the images blurred by smoothing. The disadvantages are that the metric cannot work well for all types of blurring, and the calculation of local maxima and minima could be error-prone. Wang et al. [Wang, Sheikh, and Bovik, 2002] used three parameters, namely blocking, blurring and zero-crossing, to develop a no-reference quality metric. The metric is the absolute difference in between block images, that is based upon the blocking metric in the first place. The assumption of their blur metric is that the blurring and blocking artifacts are closely related in the compressed domain. However, this assumption is only partially true because image blur is primarily due to the loss of high frequency pixels in most of the JPEG / JPEG2000 images.

We thereby present a novel blur metric based on the assumption that the loss of high frequency pixels due to quantization will occur more often. This leads to the effect of edge discontinuity, so that a blur metric in the compressed domain can take advantage. In other words, the more the JPEG image is blurred, the more discontinuous the edge

could be. We also design some experiments to evaluate if this metric not only can work correctly in quantitative measurement, but also is consistent with human perception.

3.3 Our Proposed Method: Blurring Due to Edge Discontinuity

Our algorithm originates from the observation of edge variations corresponding to different levels of compression ratio. All edges have a tendency of becoming more fragmented as the compression ratio increases. We assume that the loss of high frequency pixels due to quantization will lead to the effect of edge discontinuity. Our algorithm for blur metric is described as follows:

1. If the input image is in JPEG format, we decompress it using the Huffman and quantization tables existed in the image header.
2. We only use the luminance portion (i.e., the Y component in the decompressed YC_bC_r image). Scale up its size at least 3 times using the nearest neighbor interpolation.
3. Apply Sobel edge detector and thresholding to obtain a binary edge image.
4. Calculate the column and row gradients in the binary edge image.
5. Sum up the absolute value of the column gradients, and divide it by the total number of edge pixels.
6. Sum up the absolute value of the row gradients, and divide it by the total number of edge pixels.
7. Take the average of the results in steps 5 and 6 to obtain the blur score.

The flowchart of our algorithm is shown in Figure 3.2. The advantage of our blur metric is that we evaluate the compressed images without using the original image as the reference. This metric can also be extended to evaluate the images blurred by noise.

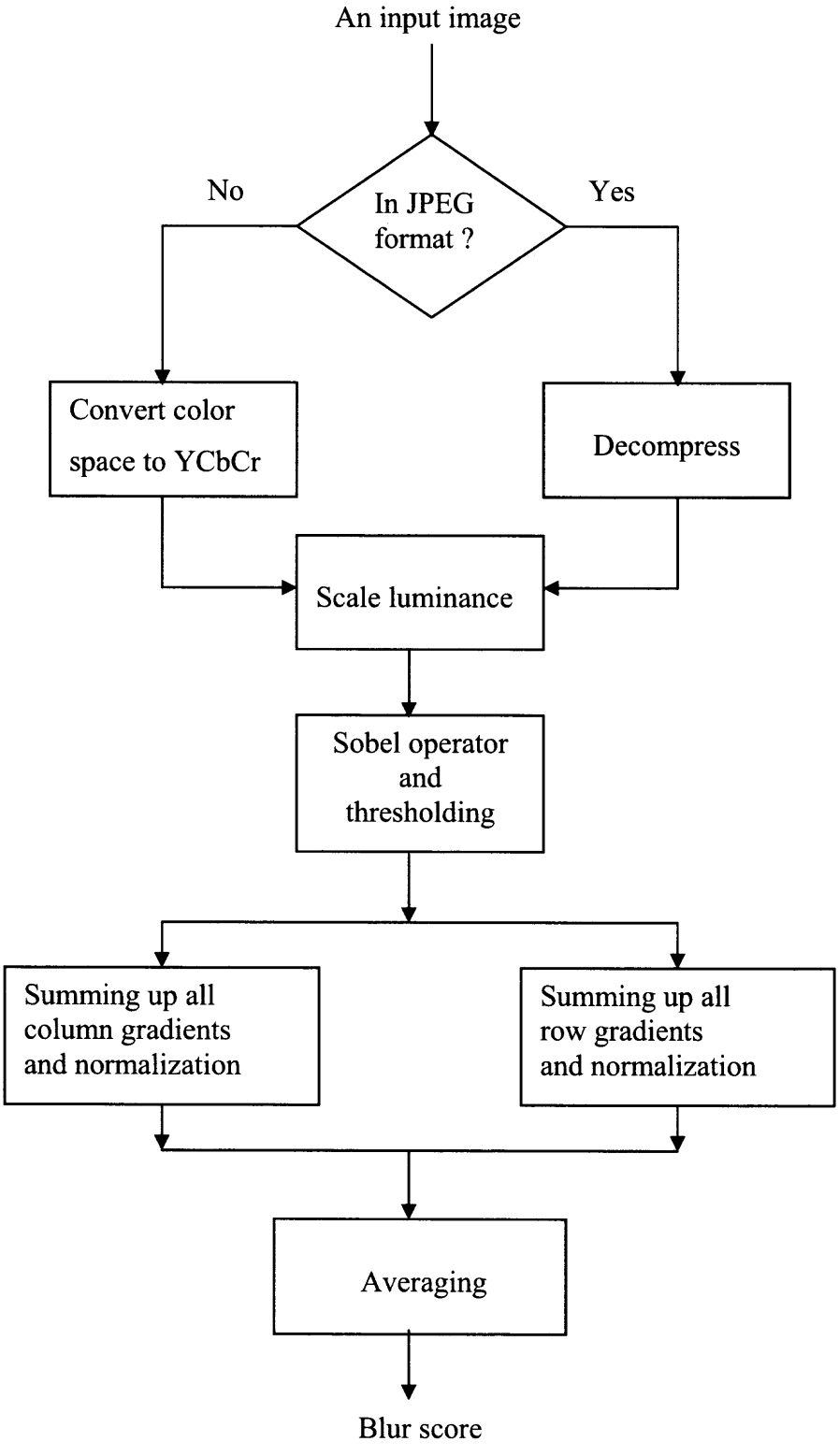


Figure 3.2 The flowchart of the proposed algorithm.

3.4 Experimental Design

In our experiments, we randomly choose 100 images from the image archive of Hubble Space Telescope as of years from 1990 to 2000. The image size is 80 by 80. The first experiment is designed for understanding the blurring effect due to compression as described below. Table 3.1 provides the blur scores for different compression ratios.

1. The original 100 images are compressed in four different compression ratios: CR10, CR20, CR30, and CR40. Three images, A, B and C, and their compressed versions are shown in Figures 3.3, 3.4 and 3.5.
2. Use our blur metric to score each image. Compute the mean of the original and four compressed image sets.

Table 3.1 The Blur Scores (Unit: 1/100) for Different Compression Ratios

	Original	**CR=10**	**CR=20**	**CR=30**	**CR=40**
Image A	8.52	8.8	9.81	10.7	12.1
Image B	8	8.5	9.8	10.6	12.1
Image C	9.1	9.5	10.6	11.1	12.7

The second experiment is designed for understanding the blurring effect due to noise as described below.

1. Perform the blurring process such that all 100 images are added by the following three types of random noises: (a) salt-and-pepper noise, (b) Poisson noise, and (c) Gaussian noise.

2. Perform the deblurring process such that all three sets of noisy images are deblurred by the Wiener filter [Gonzalez and Woods, 1992], which is commonly regarded as an effective noise-removal filter.

Original CR=10 CR=20 CR=30 CR=40

Figure 3.3 An image "A" with four compression ratios.

Original CR=10 CR=20 CR=30 CR=40

Figure 3.4 An image "B" with four compression ratios.

Original CR=10 CR=20 CR=30 CR=40

Figure 3.5 An image "C" with four compression ratios.

Figures 3.6 and 3.7 show the results for the first and second experiments, respectively. Figure 3.8 shows an image blurred by different types of noises and their deblurred images.

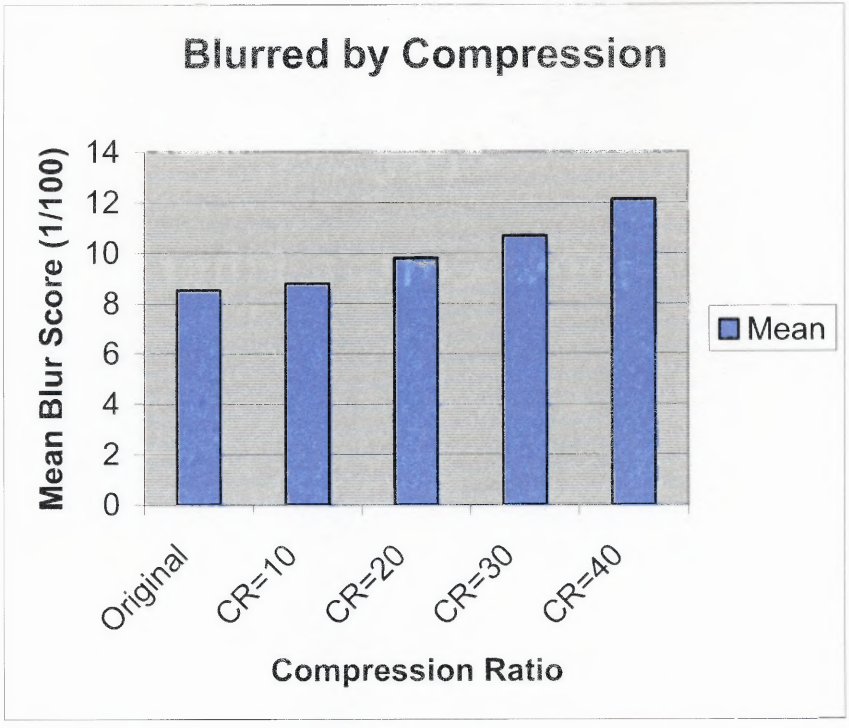


Figure 3.6 Blur metric performance on compression (totally 100 images).

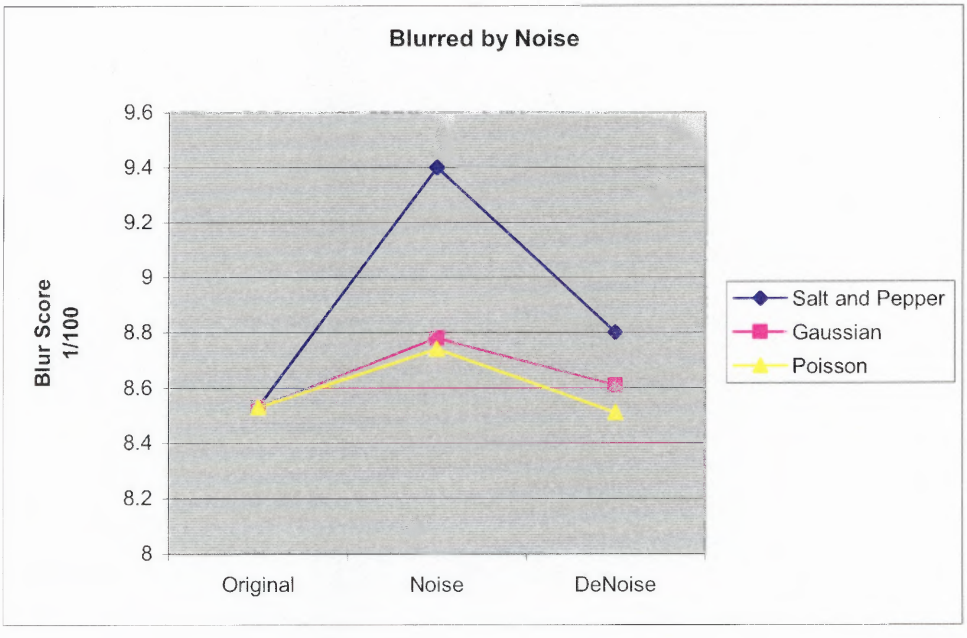
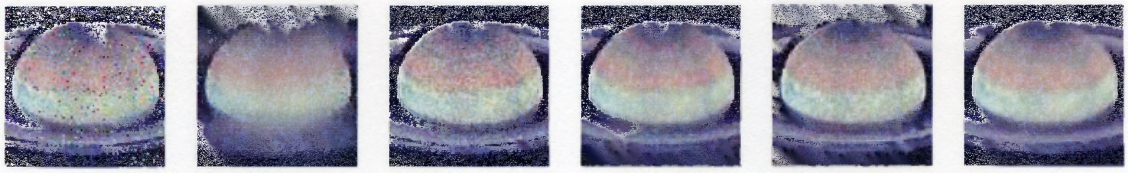


Figure 3.7 Blur metric performance on noises (totally 100 images).



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

Figure 3.8 (a) Salt and pepper noise, (b) its deblurred image, (c) Gaussian noise, (d) its deblurred image, (e) Poisson noise, and (f) its deblurred image.

To understand how the human vision system (HVS) perceives images blurred by both compression and noise, we design the third experiment as follows. We perform a survey by asking 30 persons to evaluate 100 images with the aforementioned 4 compression ratios and 3 types of noises, as well as 3 deblurred images by Wiener filter. Each questionnaire contains 4 images randomly chosen from the set of 100 images; among them, 3 images are blurred by compressions, and 1 image is blurred by noise. The blur score ranks from 1 to 5, with 1 indicating the least and 5 indicating the most blurring.

3.5. Experimental Results and Discussions

The results of the blurred by compression experiment in Figure 3.6 show that as the compression ratio increases, the score of our blurring metric climbs up. At the low compression ratio, the increment of the blur score is small. When the compression ratio is greater than 20, the blur score has a large increment from 8.8 % to 9.81 %. To summarize, human perception can accept blur score below 10.7 %. Above 12 %, it is easy to see the blocking and blurring artifacts. The proposed metric can be used to evaluate the JPEG Quality Transcoder (JQT). By measuring the blur scores of an image

before and after JQT, we can judge its performance. For a moderate gain of image quality, a reduction of blur score by at least 1/100 is required.

The results of the blurred-by-noise experiment in Figure 3.7 show the following:

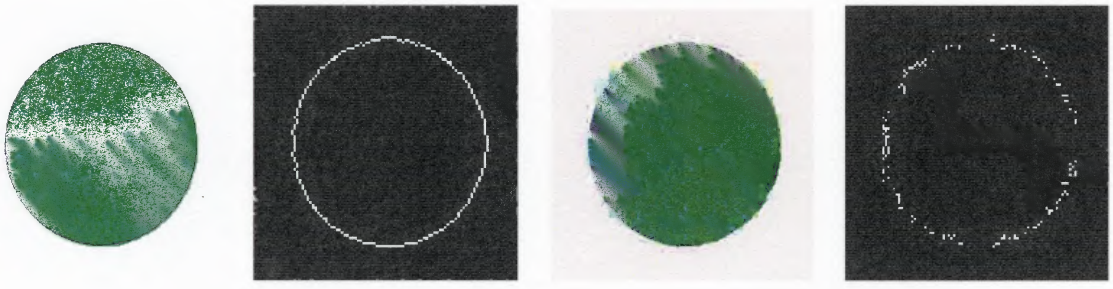
(1) As an image is blurred, the score of our metric climbs up; As the image is deblurred, the score of our metric goes down.

(2) The mean of the 100 images blurred by salt and pepper noise is much higher than the one blurred by Gaussian noise; the mean of the 100 images blurred by Gaussian noise is higher than the one blurred by Poisson noise.

The reason why the blurring score is consistent with the level of blur by compression is that there exists a phenomenon of edge discontinuity after compression. The higher the compression ratio is, the less continuous the edge is. This fact exists in any lossy compression scheme. The only exception is the Region-Of-Interest (ROI) functionality [Santa-Cruz, Grosbois and Ebrahimi, 2002] in certain JPEG2000 images since ROI is normally coded as lossless. However, our blur metric also works on the background of ROI coded images. In experiments, we observe that the phenomenon of edge discontinuity disappears when a certain compression ratio (we call it "the critical CR") is reached. This critical compression ratio is image dependent. The reason why the blurring score is consistent with the level of blur by noise is that the sum of gradient of edges after adding noise will become greater.

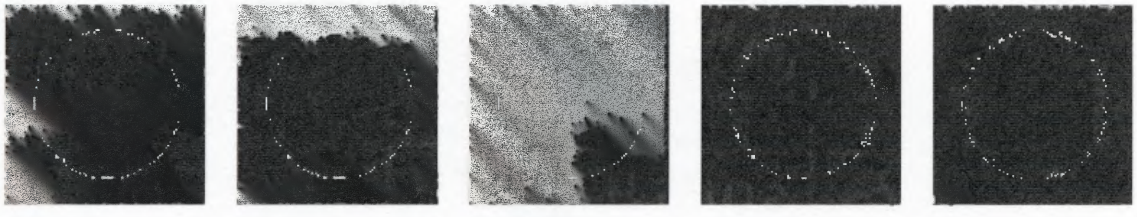
As for the blocking effect, the blocking score will go all the way up as the compression ratio increases. Once the critical CR is reached, it is suggested not to compress the image further because it will be too blurred to view. Hence, the critical CR indicates a point at which human perception quality experiences a major loss.

In order to illustrate the critical CR, we use a synthetic image; for instance, a circle, to make observation by human perception. In Figure 3.9, (a) and (b) show a circle image and its corresponding edge image. Figures 3.9 (c) and (d) show the circle in the compression ratio 200 and its edge image. Figure 3.10 presents the edge images as the compression ratio increases from 20 to 200 in a step of 20.

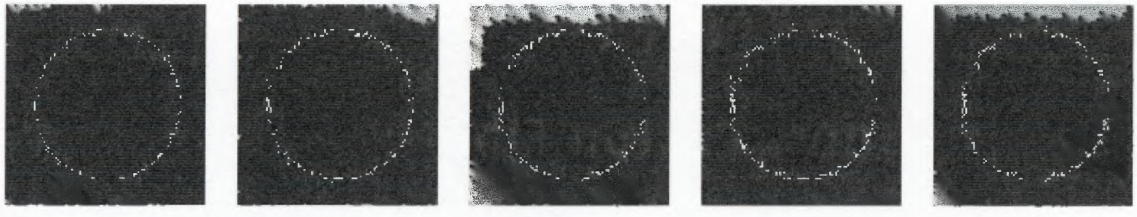


(a) (b) (c) (d)

Figure 3.9 (a) A circle, (b) edge image, (c) CR=200, and (d) its edge image.



CR=20 CR=40 CR=60 CR=80 CR=100



CR=120 CR=140 CR=160 CR=180 CR=200

Figure 3.10 The edge images as the compression ratio increases from 20 to 200.

The results of the third experiment by human vision test are listed in Tables 2 and 3. They indicate the following:

(1) The trend of mean blur score increases corresponding to the compression ratio. It is consistent with Figure 3.6. The results also suggest that for human perception, the compression ratio 10 is almost as good as the original image. Since the means of CR0 and CR10 are too close to separate, CR10 becomes the compression ratio that is visually lossless.

(2) The mean blur score of salt-and-pepper noise is greater than that of Gaussian noise; the mean blur score of Gaussian noise is greater than that of Poisson noise. It is consistent with Figure 3.7.

(3) There is a strong correlation between our blur metric and the human vision system. This can be verified by the correlation coefficients in Table 3.4, which shows that for images blurred by compression, the correlation coefficient is 0.9943, and for images blurred by noise, the correlation coefficient is 0.951. Since they are all close to 1, it is evident that our metric has a strong correlation with HVS.

Because our blur metric is primarily based upon the edge information, what would the metric perform if images contain no clear objects and edges, e.g., the texture images? We further examine the metric score on texture images. We randomly select 30 texture images and perform the image blurring by compression. The setting is the same as in the first experiment. That is, each image is compressed in CR=10, CR=20, CR=30 and CR=40. The images and blur scores for two texture images are shown in Figure 3.11. The results of blur metric performance on 30 texture images are illustrated in Figure 3.12. We observe that as the compression ratio increases, metric score increases in texture images.

Table 3.2 Human Vision Test on Images with Different Compression Ratios

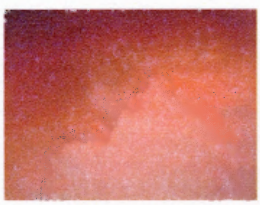
Compression Ratio	CR=0	CR=10	CR=20	CR=30	CR=40
Mean Blur Score	1.44	1.64	2.47	3.56	4.73

Table 3.3 Human Vision Test on Images Blurred by Different Noises

Type of Noises	Salt & Pepper Noise	Gaussian Noise	Poisson Noise
Blurred by Noise (Mean Blur Score)	4.87	3.13	1.8
Deblurred by Wiener Filter (Mean Blur Score)	4.17	2.73	1.2

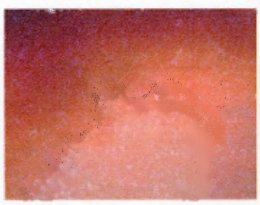
Table 3.4 Correlation Coefficients between Blurred Metric and Human Vision

	Human Vision System	Our Blur Metric	Correlation Coefficient
Blurred by Compression	1.44	8.53	0.9943
	1.64	8.98	
	2.47	9.37	
	3.56	10.44	
	4.73	11.52	
Blurred by Noise	4.87	9.4	0.9510
	3.13	8.8	
	1.8	8.7	



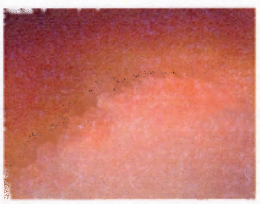
CR=10

Blur Score=0.1314



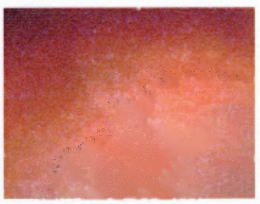
CR=20

Blur Score=0.1357



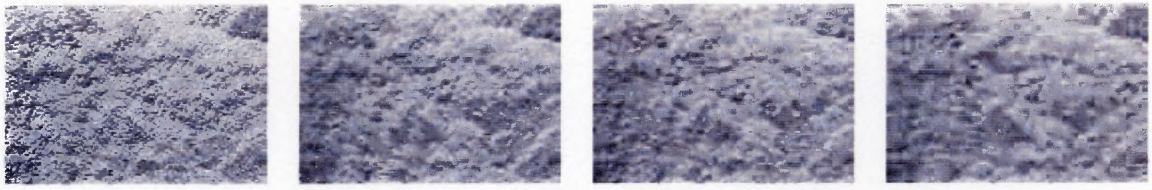
CR=30

Blur Score=0.1365



CR=40

Blur Score=0.1376



CR=10 Blur Score=0.185 | CR=20 Blur Score=0.189 | CR=30 Blur Score=0.1947 | CR=40 Blur Score=0.2115

Figure 3.11 Blur scores for two texture images as the compression ratio increases.

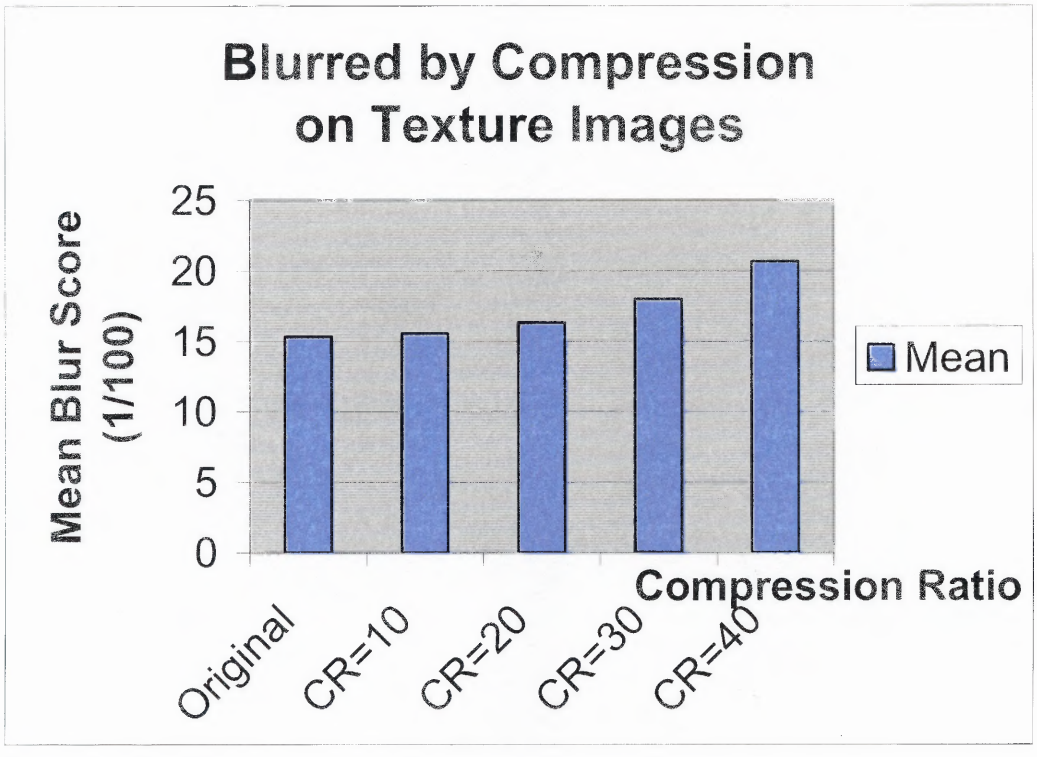


Figure 3.12 Blur metric performance on 30 texture images.

CHAPTER 4

A NO-REFERENCE QUALITY INDEX FOR COLOR IMAGES

4.1 Definitions and Assumptions

Problems associated with whether the power signal-to-noise ratio (PSNR) is a good image quality measure have been addressed throughout the researchers during the last decade [Girod, 1993;Wang, 2002; Callet, 2003]. However, until now there is no paper summarizing and assessing all the cases thoroughly when PSNR may fail to predict the human perception. The importance of understanding PSNR failure cases is easy to see. First of all, knowing what kind of image may lead to PSNR failure and their relative severity can help us design a more accurate perception metric. Second, up to the present time, many computer vision systems rely on PSNR to perform quality measurement, knowing the failure cases can help improve all these machine vision systems. The first goal of this chapter is to design a no-reference image quality index which may overcome the drawbacks of PSNR.

Image quality metrics (IQM) can be divided into two categories. One is the objective IQM, and the other is the subjective IQM. PSNR is a typical example for the former; whereas perceptual metrics are the examples for the latter. Recently, the Just Noticeable Difference (JND) becomes a popular concept to realize the subjective IQM. Many JND metrics were proposed [Zhang, 2005, Chou, 1995, Shen, 1996]. Given a pair of images, an original image and its degraded one (may be caused by noise, blur or compression), it is crucial to measure the score discrepancy between PSNR and JND.

The second goal of this chapter is to distinguish the score discrepancy between PSNR and JND, so that an automatic image quality control system can be made.

The rest of this chapter is organized as follows. In Section 4.2 we conduct a brief literature survey for image quality assessment models. Section 4.3 presents our proposed method. Section 4.4 describes the experimental design. Section 4.5 provides experimental results and discussions. Section 4.6 draws conclusions.

4.2 A Brief Literature Survey

Image quality metrics can be divided into three models:

(1) **Full Reference (FR)** – This model requires an original image as the reference to the distorted image.

(2) **Reduced Reference (RR)** – This model tends to reduce the original image into only the limited parameters, and then compare the parameters with those from the distorted image.

(3) **No Reference (NR)** – This model requires no reference image. Quality assessment is performed only on top of the distorted image.

Ideally, the NR model is the best choice for image quality assessment, since no extra data will be added to the bit-stream throughout the transmitting process. Practically, we face some problems, such as (a) the reference image is too complex to develop through the degradation model and (b) the NR quality index is hard to correlate with human perception. Therefore, a compromising solution becomes the RR model. Recently the RR model attracts a great deal of attention of Video Quality Expert Group (VQEG). Meanwhile, classical IQMs (FR model), such as MSE and PSNR, are still widely used in the digital image communication society. However, since only RR and NR are

appropriate for the quality-of-service (QoS) monitoring, lately the NR metrics test has been initiated by VQEG. As the NR model becomes more and more popular, we may need to take a retrospect to realize all the pre-existing FR and RR metrics as well as their pros and cons.

The FR model, in general, contains two stages. The first stage is called "low-level visual representation" stage, and the second stage is the "error pooling" stage. The former calculates the error between the reference and its distorted images on top of the low-level visual representation, and then to construct the distortion map; whereas the latter pools the individual local error into a global quality assessment. Two categories of FR metrics can be found in literature. Metrics from the first category use a human visual system (HVS) model [Marziliano, 2004; MacLennan, 2001; Osberger, 1998;Winkler, 2001;Winkler, 1999] for low-level perception, such as sub-band decomposition and masking effect, in order to compute distortion maps, but often propose the poor error pooling methods, such as Minkowski summation. Whereas, the metrics in the second category use little information about HVS for error representation, and push the effort on the error pooling stage integrating an *a priori* knowledge on introduced distortions. This makes those metrics in this category specialized, so they are less robust.

The RR model, similar to the FR model, constructs a low-level visual representation first, but wants to extract some high-level visual information from the low-level representation. Feature extraction is the key of RR model. By comparing the extracted features between the distorted image and the original image, a quality index can be reached. These low-level features include edges, corners, or contours; they may also be some statistical features, like mean, variance, and standard deviation. As far as the

domain from which those features are extracted, normally it is either the spatial or the frequency domain. Recently, the wavelet and the compressed domains [Tang, 2003] are gaining more and more attention. Despite of the fact that the reduced space requirement and the high-level visual information are used, RR metrics usually suffer from less robust and less feasible to all kinds of images.

NR metrics are designed to relief the cons of the above FR and RR metrics. At the present time, NR metrics are only at the beginning stage in that there are only limited research results proposed. An NR perceptual blur metric was proposed by Marziliani *et al.* [Marziliano, 2002]; an NR PSNR estimate for the compressed image was presented by Turaga *et al.* [Turaga, 2002]. In addition, Wang et al. [Wang, 2002] developed an NR perceptual quality metric for the compressed images; Tong *et. al.* [Tong, 2004] designed an NR metric for the JPEG2000 image. Because all the afore-mentioned NR metrics do not use the same error evaluation as the VQEG suggested, it is rather difficult to compare their performance at this moment.

4.3 The Proposed Method

Our proposed method is based upon the idea of linear combination of three variables, namely, colorfulness, sharpness, and contrast.

$$QI = c_1 * colorfu\ln ess + c_2 * sharpness + c_3 * contrast$$

where QI stands for Quality Index.

Colorfulness (also referred as Chromaticness) is the attribute of a visual sensation according to which the perceived color of an area appears to be more or less chromatic. By modifying the colorfulness definition from [Hesler, 2003], we define colorfulness in

the opponent color space as follows. Let α denote the *rg* channel, and β denote the *yb* channel.

$$Colorfu\ln ess = ((\sqrt{\sigma_\alpha^{\ 2} + \sigma_\beta^{\ 2}} + 0.3 * \sqrt{\mu_\alpha^{\ 2} + \mu_\beta^{\ 2}}) / 85.59)$$

where

$$\alpha = R - G;$$

$$\beta = 0.5 * (R + G) - B;$$

Sharpness is the perception that a picture, or parts of a picture are in focus. It is also the rendition of edges or tonal borders. In literature, sharpness is mostly defined by the local Kurtosis of DCT block of an edge image [Caviedes, 2004]. We define sharpness based upon the total number of edge pixels extracted from Sobel operator. Sharpness is defined as follows:

$$mpep = (tep_{Estimated} - tep_{Sobel}) / tep_{Sobel};$$

$$sharpness = 1 - mpep^{0.2};$$

where

$tep_{Estimated}$ denotes total edge pixels estimated;

tep_{Sobel} denotes total edge pixels counted using Sobel operator;

mpep denotes missing per edge pixel;

Contrast is the difference in tone between the light and dark areas of an image. Normally, it is defined as the ratio of the maximum luminance (intensity) in an image to the minimum luminance. Inspired by Susstrunk and Winkler [Susstrunk, 2004], we define contrast as follows:

Global Contrast = max of Local Contrast;

where $localContrast = \sum_{i=9}^{15} Band_i / \sum_{j=1}^{8} Band_j$. The total 15 bands are the 8×8 blocks of DCT coefficients.

4.4 Experimental Design

Two experiments are conducted. The first experiment aims at testing the sensitivity of three variables; whereas the second aims at finding the corresponding coefficients for the three variables according to human perception.

In experiment I, altogether 100 natural images are chosen from Google image database. To test the sensitivity of colorfulness, we use the percentage of red color from 10 % to 90 %, with an interval of 10 %. This percentage of red will be randomly drawn to the images. To test the sensitivity of sharpness, we use the compressed images from compression ratio 20 to 180, with an interval of 20. To test the sensitivity of contrast, we use the gamma correction to enhance the contrast of the images with gamma from 1.1 down to 0.7 and an interval of 0.05, while gamma is γ in the equation $\mathbf{I}^{\gamma}$.

In experiment II, on one hand, our proposed three variable estimators, namely, colorfulness, sharpness, and contrast measures will be performed on selected 30 images in the survey questionnaire; on the other hand, 30 students will be selected to fill the questionnaire by giving a quality score to each image (The scale is from 1 to 7, with 1 the worst and 7 the best quality. Sample questionnaire is provided in Appendix B.) With our estimated colorfulness, sharpness, and contrast vectors, we can calculate the corresponding coefficients for each variable by using the multiple linear regression method based on the quality score from human perception.

4.5 Experimental Results and Discussions

For experiment I, the results for the sensitivity of colorfulness are listed in Table 4.1, and for the sensitivity of sharpness and contrast are respectively listed in Table 4.2 and 4.3.

Table 4.1: Sensitivity of Colorfulness (Average of 100 images)

Percentage	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %
Colorfulness	0.3405	0.3423	0.4518	0.4919	0.5428	0.5690	0.6061	0.6468	0.6554

Table 4.2: Sensitivity of Sharpness (Average of 100 images)

Compression Ratio	20	40	60	80	100	120	140	160	180
Sharpness	0.7049	0.661	0.6076	0.5687	0.4042	0.3772	0.3529	0.3335	0.3066

Table 4.3: Sensitivity of Contrast (Average of 100 images)

Gamma	1.1	1.05	1.0	0.95	0.9	0.85	0.8	0.75	0.7
Contrast	0.7754	0.6382	0.5172	0.4559	0.3906	0.3311	0.2811	0.2377	0.1938

In Table 4.1, we observe that the colorfulness is in a monotonic increasing order as the percentage of red increases. In Figure 4.1, it is easy to see that when the image turns black, our colorfulness score goes to near 0; whereas, as the image turns on totally red, the colorfulness score will be 1.0. In Table 4.2, we can see the sharpness is in a monotonic decreasing order as compression ratio increases. It is worth noting that before compression ratio 100, the sharpness score drops very fast. However, as compression ratio changes from 100 to 180, the sharpness score goes down smoothly. In Table 4.3, we

can see the contrast is in a monotonic decreasing order as Gamma decreases. Above gamma 1.0, the contrast score experiences a sharp drop. From 1.0 to 0.8, the contrast goes down in a slow manner; while from 0.8 downward, the contrast score goes down even smoothly. Figures 4.2 and 4.3 present the sample results for the sensitivities of sharpness and contrast, respectively.

Figure 4.1: Sample Image Result on the Sensitivity of Colorfulness
Left: Colorfulness=0.03; Middle: Colorfulness=0.5; Right: Colorfulness=1.0

Figure 4.2: Sample Image Result on the Sensitivity of Sharpness
Left: CR=20:1, Sharpness = 0.7049;
Middle: CR=100:1, Sharpness = 0.4042;
Right: CR=180:1, Sharpness = 0.3066.

Figure 4.3: Sample Image Result on the Sensitivity of Contrast
Left: Gamma=1.1, Contrast=0.7754;
Middle: Gamma=1.0, Contrast=0.5172;
Right: Gamma=0.9, Contrast=0.3906;

For experiment II, by using the multiple linear regression, we obtain three corresponding coefficients: c1 = 3.0850, c2=6.8584, and c3=1.5067. By plugging in the coefficients to our proposed color quality index:

QI = c1 * Colorfulness + c2 * Sharpness + c3 * Contrast

we obtain

QI = 3.0850 * Colorfulness + 6.8584 * Sharpness + 1.5067 * Contrast

Hence, the normalized quality index for color images is:

$\mathbf{QI_{color}}$ **= (3.0850 Colorfulness + 6.8584 Sharpness + 1.5067 Contrast) / c1+c2+c3;**

However, as for the gray scale image, the normalized quality index is:

$\mathbf{QI_{intensity}}$ **= (3.0850 Colorfulness + 6.8584 Sharpness + 1.5067 Contrast) / c2+c3;**

In Figure 4.4, we present the experimental results for some images.

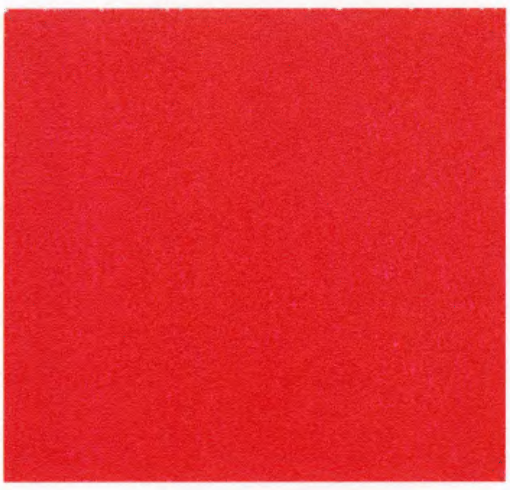
Figure 4.4: Quality Index for Sample Image Results

Sharpness = 0.6563
Contrast = 0.2438
Colorfulness = 0.6521
QI= 0.6008

Sharpness = 0.6416
Contrast = 0.7754
Colorfulness = 0.1550
QI = 0.5300

Sharpness = 0.5610
Contrast = 0.3156
Colorfulness = 0.1282
QI= 0.4121

Sharpness = 0.6197
Contrast = 0.2969
Colorfulness = 0.000
QI = 0.5615



Sharpness = 1.0
Contrast = 0.0
Colorfulness = 1.0
QI= 0.8684

Sharpness = 0.6370
Contrast = 0.5101
Colorfulness = 0.421
QI = 0.5613

Last, we show in Figure 4.5 that two images of the same MSE, which is 225, but they will be perceived drastically different by observers. In this case, the MSE fails. However, they can be correctly discriminated using our proposed QI. The first image is a contrast stretched Lena; the second image is a Gaussian blurred image. Evaluated by our quality index, the former image has QI = 0.6391; whereas the latter image has QI = 0.1646 only.

(a) (b)

Figure 4.5. Evaluation of Lena image distorted by
(a) contrast stretching; MSE = 225, QI = 0.6391
(b) Gaussian blur; MSE = 225, QI = 0.1646.

The root of PSNR failure is the fundamental discrepancy between PSNR and JND; meanwhile, PSNR requires a reference image for quality evaluation. To solve this problem, this chapter attempts a no-reference quality index for color images. The proposed method is based on the linear regression of three factors, namely, colorfulness, contrast, and sharpness. Experimental results indicate that the corresponding weights to three factors in a descending order are sharpness, colorfulness, and contrast. We show that the proposed quality index works successfully in cases when PSNR fails. Meanwhile, the index correlates well with human perception, since the corresponding coefficients derived from the multiple linear regression are based upon the subjective score of human observation.

Future direction will be on incorporating the proposed quality index to the image database for content-based color image retrieval. Quality index will be served as a filter to rule out color images with unqualified quality.

CHAPTER 5

COLOR CORRECTION FOR COLOR DISTORTED IMAGES

5.1 Introduction

Color distortion due to output device has been discussed in the literature [Hung, 1993;Yin, 2004;Balasubramanian, 1998;Finlayson, 1997;Zhang, 2001; Kanomori, 1992; Li, 2004]. Yet rare paper discusses the color distortion due to compression, filter or motion. Observers usually care about artifacts, such as JPEG blocking artifact and ringing artifact, but they seldom care for color distortion. It is because human vision may be not bothered by color distortion and the original uncompressed image cannot be obtained for comparison. To understand the severity of color distortion due to object motion and Gaussian filter, it is necessary to conduct experiments to assess color distortion in different color spaces for blurred images.

Color correction is related to color reproduction [Omer, 2004;Emmel, 2000;Chosson, 2003;Zhang, 1997] in another device. When related to the display device, the gamma correction is performed. The goal of color correction is to make sure the color range (gamut) keeps similar as color goes through input device to output device. Thus, color correction is a color calibration and adjustment process in terms of device, but becomes a process of color enhancement by considering compression or filter. In other words, color correction becomes image enhancement if no color distortion is considered, and is a color reproduction procedure subject to color distortion constraint. The former problem has been tightly related to tone reproduction [Matkovic, 1997]. As a result, combining color and tone reproductions is equivalent to color image enhancement. The

latter is a constraint optimization problem. There is almost no paper discussing color correction in this regard; only some papers addressed the color distortion metric problem [Fuhrmann, 1995;Winkler, 1998;Eckert, 1998;Janssen, 2001].

The motivation of this chapter is that we think the problem of color correction subject to constraint to be very important. One reason is that if JPEG colors can be reproduced and minimized subject to some constraint, then the online JPEG color image downloading time can be reduced and saved, since the bottleneck of JPEG image download and display is on the total number of colors. Another reason is that for artwork color reproduction, color enhancement could not be made as much as possible, it is better to follow certain constraint to make artwork genuine and truthful. To this end, this chapter discusses two issues in general. The first issue is concerning the notion of color distortion due to motion, filter and compression. The second issue is concerning the notion of color correction subject to color distortion constraint. We will investigate these two issues by designing experiments running on top of our proposed methods.

The rest of this chapter is organized as follows. In Section 5.2 we present the proposed methods. Section 5.3 describes the experimental design. Section 5.4 provides experimental results. Section 5.5 draws conclusions.

5.2 Definitions and Assumptions

We present the related definitions of color distortion as follows.

Definition 1: *Color Distortion*: Color distortion denotes an undesired change in color value as it passes from an input device to an output device.

Definition 2: *Color Aberrant*: Color aberrant denotes color deviating from the normal.

Definition 3: *Color Correction*: Color correction denotes the process of correcting or enhancing the image color. In literature, color correction is always related to output device. Here we extends color correction meaning to compression and filter. Hence, color correction can be regarded as a color reproduction process subject to color distortion minimization constraint.

Definition 4: *Gamma Correction*: Gamma correction is the process tailored to the color display device. Before display, the linear RGB data must be processed (gamma corrected) to compensate for the gamma of the display

Definition 5: *Color Calibration*: Color calibration is a process by which the input and output devices are matched to use the same or similar color palette. This insures that the image as seen on the monitor has the same range of colors as the image that is printed, and any adjustments made to the color of the image in the computer are accurately represented when the image is printed.

Definition 6: *Color Reproduction*: Color reproduction denotes the process of reproducing colors on different devices. Two common methods used for reproducing color are additive color mixtures and subtractive color mixtures.

Throughout the chapter, color distortion follows definition 1; whereas the source of distortion is from algorithms or filters. Color correction follows definition 3. Observing that for **referenced color correction (RCC),** correction process will use a referenced image to perform color correction; for **no-reference color correction (NRCC),** normally the correction process means a color reproduction process. The

assumption for RCC is that in the controlled world we have an original image without any distortion for reference purpose; whereas the assumption for NRCC is that it is almost impossible to find a reference for any distorted image in the real world.

The assumption for the both algorithms is that no color gamut mapping between color space is needed since here we mainly deal with motion blur and Gaussian blur filter. In this section, we present two methods for color correction. One method aims at RCC, the other aims at NRCC.

5.3 The Proposed Method (A Lookup Table Method)

The RCC (Referenced Color Correction) Method:

Algorithm:

1. Given an original image O and its blur image B, to compute color distortion (CD_BEGIN) in RMSE for each channel, separately;
2. Perform preprocessing tasks as follows:
   1) to denoise from B using 3×3 Median filter to get a restored image R1;
   2) to deblur from R1 by adaptive restoration (Adjusting the Point Spread Function (PSF)) by minimizing to MSE between R1 and O to get the best restored image R2;
3. Mask CD_BEGIN on top of R2 for each channel to get a restored image R3;
4. Minimize color distortion between R3 and O using the Modified Popularity Algorithm (MPA) to get a restored image R4.
5. Compute color distortion (CD_END) between O and R4 for each channel, separately;

6 Compute the Quality Gain (QG), when

QG = (CD_BEGIN – CD_END) / CD_BEGIN;

where the Modified Popularity Algorithm (MPA) is as follows:

1) Pick the highest K (we use K=10) peaks of the histogram of image O and record the intensities with their corresponding pixel count – Ideal Peak Count (IPC);

2) Stepwise increase the K intensities of the histogram R by decrease the pixels from the consecutive lowest peaks of middle intensity to reach their corresponding IPC. (The reason to do this is that it is better not to bother the color gamut range.)

3) Until the relative color distortion is minimized.

The NRCC (No-Reference Color Correction) Method:

Algorithm:

1 Given a blur image B, to perform preprocessing as follows:

1) to denoise using 3×3 Median filter to get image R1;

2) to deblur using 5×5 Wiener filter to get a restored image R2;

2. From header info of image B, to compute its approximate compression ratio (ACR) [Neelaman, 2003]; Note that ACR is only limited to the most recently compression; it does not trace back to all the compression history if a given JPEG has been compressed several times.

3 From a pre-established color distortion lookup table (see table 1), find the corresponding color distortion (CD_BEGIN) for each channel based on ACR;

4. Approximate the original uncompressed image O, where O = B + CD_BEGIN;

5 Compute the 3-channel color distortion between O and B;

6. Use adaptive unmasking enhancement (AUE) algorithm on top of R2 to obtain the restored image R3 and to obtain the best constant factor k.

7. Mask CD_BEGIN on top of R3 for each channel to get image R4;

8 Compute color distortion (CD_END) between O and R4 for each channel, separately;

9 Compute the Quality Gain (QG),

QG = (CD_BEGIN – CD_END) / CD_BEGIN;

, where the adaptive unmasking enhancement (AUE) algorithm is as follows:

1) Given an image R, perform smoothing to get image S;
2) edge mask = R – S;
3) enhanced R (ER) = R + k * (edge mask);

Note that the limitation of RCC method is that in the real world, it is not feasible to obtain an original image that is the previous version of the color distorted image. The limitation for NRCC is that in case the quantization table is missing, ACR (Approximate Compression Ratio) has to be approximated from the quantization coefficient by counting the zeros as well as near zeros. Although it is possible to obtain the image quality index by IJG provided free software, we have to be aware that a quality score below Q10 is very hard to compute the compression ratio.

5.4 Experimental Design

Altogether three experiments are conducted. Experiment I is a comparative assessment of color distortion in different color space for blurred images. The goal is to understand the color distortion in motion blur and Gaussian blur images. Experiment II is an experiment designed for referenced color correction (RCC) in motion blur and Gaussian blur images. The goal is to minimize the color distortion under the constrained environment, that is, there exists the reference image. Experiment III is an experiment designed for no-reference color correction (NRCC) in low bit-rate JPEG. The goal is to minimize color distortion of low bit-rate JPEG images in situation where no reference image can be found.

For experiment I, 200 natural images are randomly selected from the Goggle image search engine. The image contents are natural scenes images. The size of each image will be resized to resolution 64x64. Each image is conducted by 16 motion blurs and 8 Gaussian blurs. Totally, there are $N = 4800$ images for experiment I. For motion blur, first we fix the orientation and change the magnitude, fixed 45° with magnitudes increasing from 5 to 40 at an interval of 5, altogether we have 8 levels of different magnitudes; second we fix the magnitude and change the orientation, fixed magnitude 10 with orientations increasing from 10° to 80° at an interval of 10°, altogether we have 8 different orientations. For Gaussian blur, we perform 8 point spread functions (PSFs) with window sizes increasing from 3×3, 5×5, all the way to 17×17 at odd numbers. Color distortion will be measured by root mean square error (RMSE) in both 1-channel (luminance) and 3-channel settings. The color spaces tested include:

1. Device Dependent

- RGB
- YcbCr
- HSV
- NTSC (YIQ)

2. Device Independent
   - XYZ
   - LAB
3. Human Vision System (HVS) Related
   - Opponent color space

For experiment II, the testing sample in figure 1 is chosen from the high color distortion motion blur images, that is 200 images in magnitude 40 with fixed orientation 45°. We run our proposed RCC method on the testing sample. For each restored image, we record (1) RMSE between the blur image and its reference image; (2) RMSE between the restored image after our method and its reference image. Then we compute the average RMSE of the sample set for both (1) and (2). Finally we compute the color distortion in terms of quality gain in percentage.

For experiment III, first we remain use the 200 natural images in JPEG format from Experiment I. Then we manually compress them in 200:1 compression ratio with size 256x256. Now the testing sample in figure 2 is the latter set of 200 with high compression ratio. We run our NRCC method on the testing sample. For each restored image, we follow the same procedure as done in experiment II to come up with the quality gain in percentage. The only difference here is that since we do not possess the reference image,

we may need to compute an approximate version of the reference image based upon the entry from lookup table searching by its approximate compression ratio. It is worth noting that to approximate the reference image in our NRCC method, we pre-computed a lookup table (in figure 3) which records the average RMSE in R, G, B each channel for compression ratios ranging from 10, 20, 30 all the way to 200.

Figure 5.1 Testing samples (motion blur image with its reference image) for experiment II; (RCC-Referenced Color Correction).

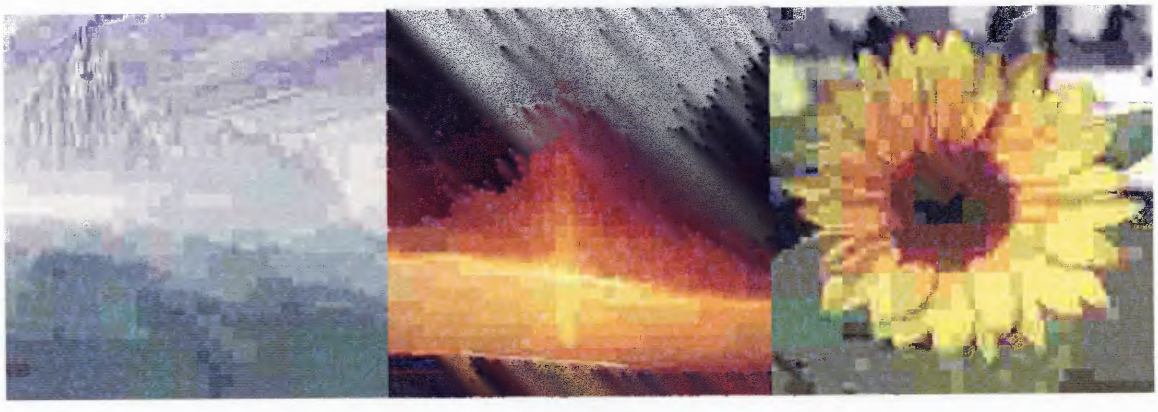


Figure 5.2 Testing samples (JPEG images with CR 200:1) for experiment III; (NRCC - No-Reference Color Correction).

Table 5.1 Lookup Table Containing the Pre-computed Color Distortion (RMSE) in Each Channel for Different Compression Ratio

	RMSE_R	RMSE_G	RMSE_B	RMSE_3Channel
CR10	2.1454	1.8053	2.5095	3.762903068
CR20	3.6904	3.363	4.0135	6.406013067
CR30	4.2574	3.9299	4.5713	7.380132279
CR40	4.6462	4.299	4.9525	8.037153208
CR50	4.945	4.5832	5.2646	8.554224711
CR60	5.1955	4.8369	5.5291	8.997764649
CR70	5.4352	5.0627	5.7323	9.382515314
CR80	5.621	5.2517	5.9373	9.71738263
CR90	5.8138	5.3988	6.0929	10.00353649
CR100	5.9633	5.5665	6.2192	10.25793926
CR110	6.1431	5.7529	6.3791	10.56060854
CR120	6.2789	5.8831	6.4704	10.76575715
CR130	6.4503	6.0396	6.6074	11.03362465
CR140	6.5795	6.2073	6.8467	11.34450062
CR150	6.7315	6.3368	6.9286	11.55307857
CR160	6.8365	6.4985	7.1322	11.82524889
CR170	6.9944	6.6349	7.1959	12.03014988
CR180	7.226	6.8705	7.3391	12.38067991
CR190	7.4181	6.8953	7.5904	12.65652171
CR200	7.7906	7.015	7.5238	12.90392343

5.5 Experimental Results and Discussions

For experiment I, the experimental results are shown in Figures 5.4-5.6. We summarize the experimental finding as follows:

(1) HSV has the least color distortion on both motion blur and Gaussian blur images; NTSC has the second smallest color distortion;

(2) XYZ has the most color distortion on both motion blur and Gaussian blur images; In terms of RMSE, XYZ always exceeds 1.0 for any kind of settings; the color distortion of XYZ is far above the color space with second largest distortion - the Opponent color space whose distortion does not exceed 0.085 in motion blur and does not exceed 0.055 in Gaussian blur.

(3) For all the color space, color distortion on 3-channel is only a little higher than that of 1-channel, the difference will not exceed 0.01, except the 0.5 on the XYZ color space.

(4) The distortion of motion blur increases as the magnitude of velocity vector increases while fixing the orientation. It is monotonically increasing.

(5) The distortion of motion blur does not reveal any pattern as the orientation of velocity vector increases while fixing the magnitude.

(6) The distortion of Gaussian blur does not increase (remain flat) as the window size increases. In fact, the distortion is always a constant, regardless of the window size.

For experiment II, the experimental results are shown in Figures 5.7-5.8. We summarize the experimental finding (see Table 5.2) as follows:

(1) The average quality gain for the whole testing sample is 88.3 % by RCC.

(2) The average running time for one image is less than 1 second, even though the most time consuming part is the blind restoration of the preprocessing step.

(3) The resulting image after preprocessing has less distortion than the blurred image by 30.24 % in average.

(4) The resulting image after color correction has less distortion than the resulting image after preprocessing by 55.86 % in average.

(5) The resulting image after MPA has less distortion than the resulting image after

For experiment III, the experimental results are shown in Figures 5.9-5.10. We summarize the experimental finding (see Table 5.3) as follows:

(1) The average quality gain for the whole testing sample is 60.71 % by NRC

(2) In case where we can obtain the compression ratio of a given JPEG image from an advanced software, the average running time for one image is less than 1 second. In case where we estimate the compression ratio by deriving the quality score from IJG software, i.e. Q50 is equal to 20:1 and Q25 is equal to 30:1, then the estimation will be error prone when quality score is lower than Q10. In this case the average running time is still less than 1 second. However, if we want to estimate the compression ratio by counting the zeros and near-zeros from the quantization coefficients for all the 8x8 block, then the average running time will be increased to 2-3 second.

(3) The resulting image after pre-processing has more distortion than that of the low bit-rate image by 17.67 % in average.

(4) The resulting image after AUE has less distortion than the resulting image after pre-processing by 3 % in average.

(5) The resulting image after color correction has less distortion than the resulting image after AUE by 64.2 % in average.

Table 5.2 Quality Gain (QG) in Each Step of RCC Algorithm

	QG from Preprocessing	QG from Color Correction	QG from MPA	Total QG
Net Gain	30.24 %	55.86 %	60.7 %	88.3 %
Gross Gain	30.24 %	70.51 %	88.3 %	88.3 %

Table 5.3 Quality Gain (QG) in Each Step of NRCC Algorithm

	QG from Preprocessing	QG from AUE	QG from Color Correction	Total QG
Net Gain	-17.67 %	3 %	64.2 %	60.71 %
Gross Gain	-17.67 %	-9.75 %	60.71 %	60.71 %

In Experiment I, the most important findings are (4) and (6). That is, the distortion of motion blur increases as the magnitude of velocity vector increases while fixing the orientation. Meanwhile, the distortion of Gaussian blur does not increase (remain flat) as the window size increases. It is always a constant. The reason for (4) is close to intuition in that as we change the magnitude of velocity vector by an amount of 5 pixels, the increasing speed will lead to the distortion increase. However, The reason for (5) is intuitively violated. Originally we know that Gaussian may cause smoothing effect that leads to a distortion, which is correct according to the finding of experiment I. What this experiment tells us more is that the distortion is always a constant no matter what size of the window is. This needs a further proof.

10	10	10	10	10	10	10
20	20	20	20	20	20	20
40	40	40	40	40	40	40
220	220	220	220	220	220	220
40	40	40	40	40	40	40
20	20	20	20	20	20	20
10	10	10	10	10	10	10

Figure 5.3 The original Image A with size 7×7.

0.0113	0.0838	0.0113
0.0838	0.6193	0.0838
0.0113	0.0838	0.0113

Figure 5.4 A 3×3 Gaussian low-pass filter.

0	0	0.0002	0	0
0	0.0113	0.0837	0.0113	0
0.0002	0.0837	0.6187	0.0837	0.0002
0	0.0113	0.0837	0.0113	0
0	0	0.0002	0	0

Figure 5.5 A 5×5 Gaussian low-pass filter.

8.9349	18.8215	50.9658	162.3097	50.9658	18.8215	8.9349
10	21.0651	57.0411	181.6575	57.0411	21.0651	10
10	21.0651	57.0411	181.6575	57.0411	21.0651	10
10	21.0651	57.0411	181.6575	57.0411	21.0651	10
10	21.0651	57.0411	181.6575	57.0411	21.0651	10
10	21.0651	57.0411	181.6575	57.0411	21.0651	10
8.9349	18.8215	50.9658	162.3097	50.9658	18.8215	8.9349

Figure 5.6 The result image G3: Gaussian blur by 3×3 filter.

8.9376	18.859	50.9389	162.1958	50.9389	18.859	8.9376
10.0026	21.1064	57.0092	181.5243	57.0092	21.1064	10.0026
10.0053	21.112	57.0242	181.5722	57.0242	21.112	10.0053
10.0053	21.112	57.0242	181.5722	57.0242	21.112	10.0053
10.0053	21.112	57.0242	181.5722	57.0242	21.112	10.0053
10.0026	21.1064	57.0092	181.5243	57.0092	21.1064	10.0026
8.9376	18.859	50.9389	162.1958	50.9389	18.859	8.9376

Figure 5.7 The result image G5: Gaussian blur by 5×5 filter.

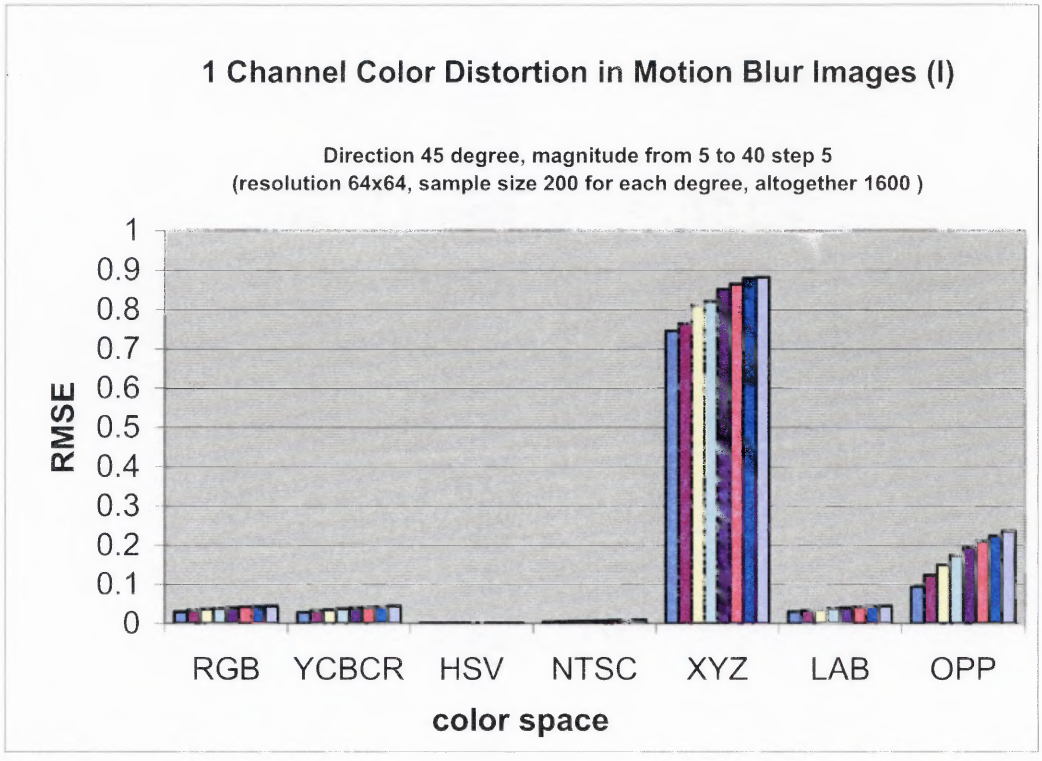
From Figure 5.3 on, we know that the original image A convolves with Figure 5.4, the result image is G3 in Figure 5.6. Similarly image A convolves with Figure 5.5, the result image is G5 in Figure 5.6. Since the Euclidean distance D3 from G3 to A is

57.1151 and the Euclidean distance D5 from G5 to A is 57.2483, we know that D5-D3 is 0.13, which is due to arithmetic of double precision. Similar situation happens between window size 5×5 and 7×7, 7×7 and 9×9. This shows that the distortion will be always a constant despite of the variation of window size, as long as the standard deviation of the filter remains the same. In experiment I, we keep all the window sizes with the standard deviation 0.5.

In experiment II, the most important finding is that despite the fact that color correction has improved the distorted image by 70.51 % compared to the reference image, MPA still can make a further correction to the distorted image by 88.3 % compared to the reference image. It is not surprising that color correction can gain most of the quality in each channel, but it is only a global amount that has no idea on which color is missing and which color is more than enough. As a consequence, we apply MPA to adjust the histogram of the image after color correction. The changes in histogram in Figure 5.10 may best explain why MPA is necessary. What we expect further is to minimize the MSE between the corrected image and the reference image. The adjusting process needs to be adaptive because it is very easy to adjust the peak more than enough such that the MSE become increasing. It is worth noting that if we apply MPA before color correction, then it may cost more time and the resulting MSE will not be as small.

In Experiment III, the most surprising result is that the pre-processing tasks make a quality gain –17.67 %. This means that, when we apply denoise and deblur methods for low bit-rate image, although the resulting image looks smooth and better visually, the MSE is even increasing. Even though the performance of AUE – the net quality gain 3 %, is not good, AUE still can reach its goal which enhances the edge part, according to

the experiment results. The most important role in experiment III is the prepared lookup table on which the color correction process is based. By masking the adequate amount to each channel, the MSE would drop which makes the total quality gain from negative to 60.71 %. Meanwhile, the property of sharp edge is also preserved. The only limitation of our NRCC method is that there is no way to understand which color is missing according to the corrected histogram, since we do not possess the reference image and its histogram. Although we estimate the original image by adding the channel distortions to the low bit-rate image, it is merely an estimated reference image whose MSE distance to the distorted image is exactly the same as the distance between the original image and the distorted image. This is to say, we cannot use the histogram of the estimated original image as the referenced histogram. This explanation may also helps to understand why the resulting image after NRCC will be different from that of RCC. In summary, the similarity of the resulting image after RCC and NRCC is that they both will have near-zero MSE from the original image. The difference of the resulting image after RCC and NRCC is that the histogram of the resulting image from RCC will be similar to the histogram of the reference image, whereas the histogram distance between the resulting image from NRCC and the original image (once if we can find it) can not very far.



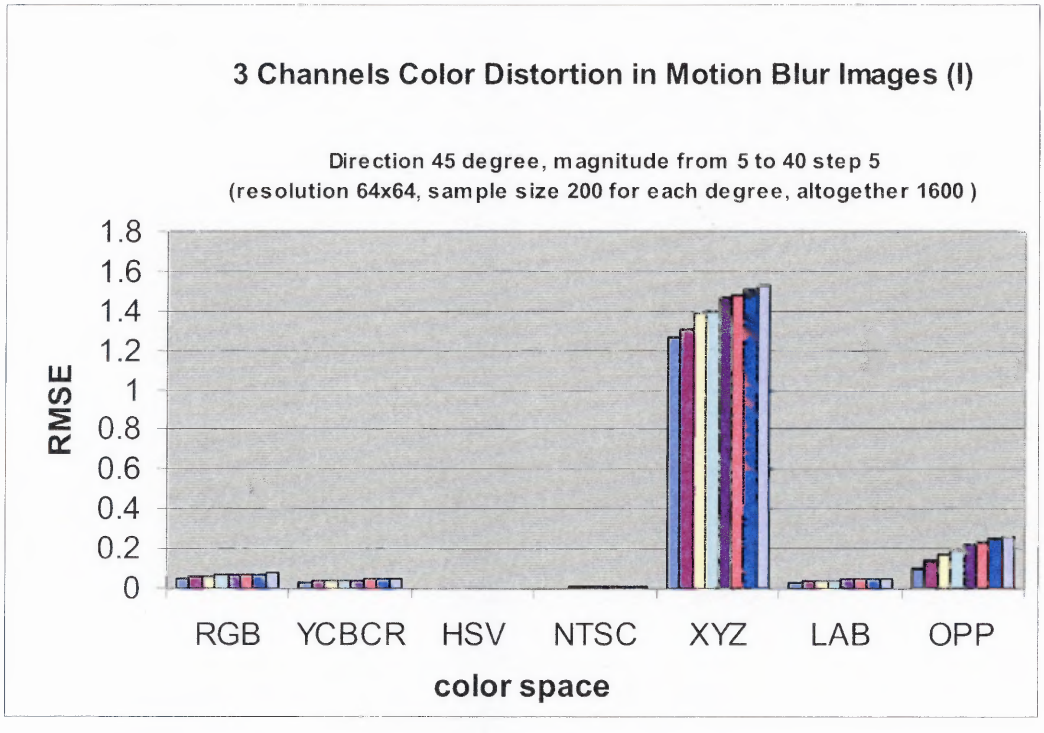
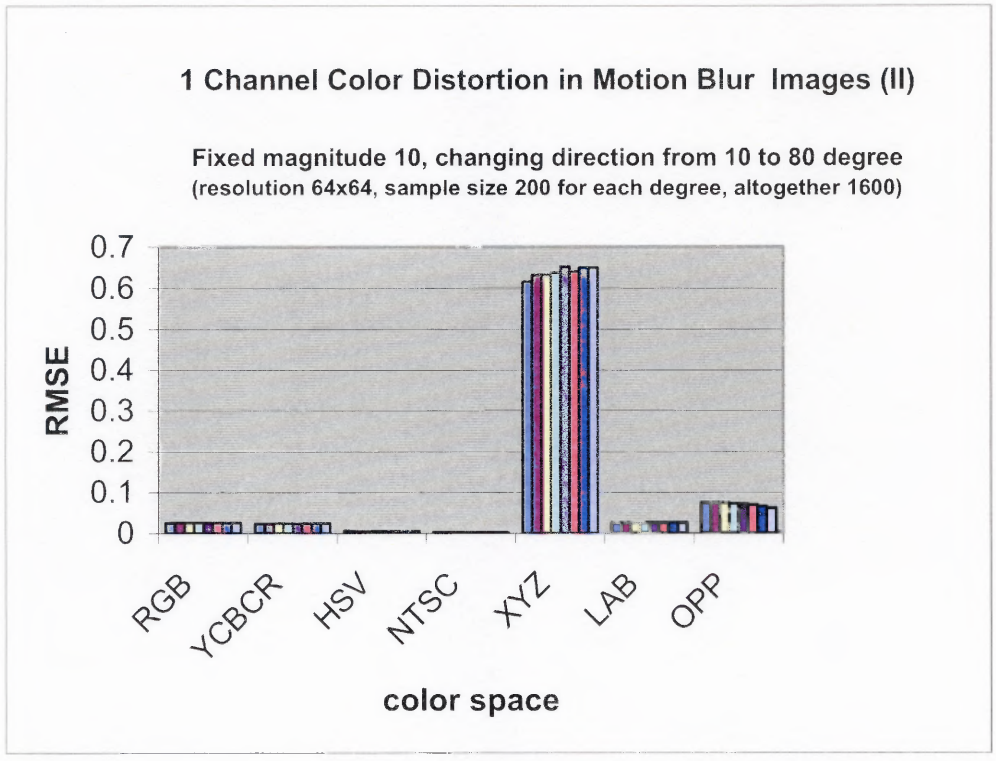


Figure 5.8 Color distortion due to variable magnitudes in motion blur images.

(1 channel versus 3 channels)



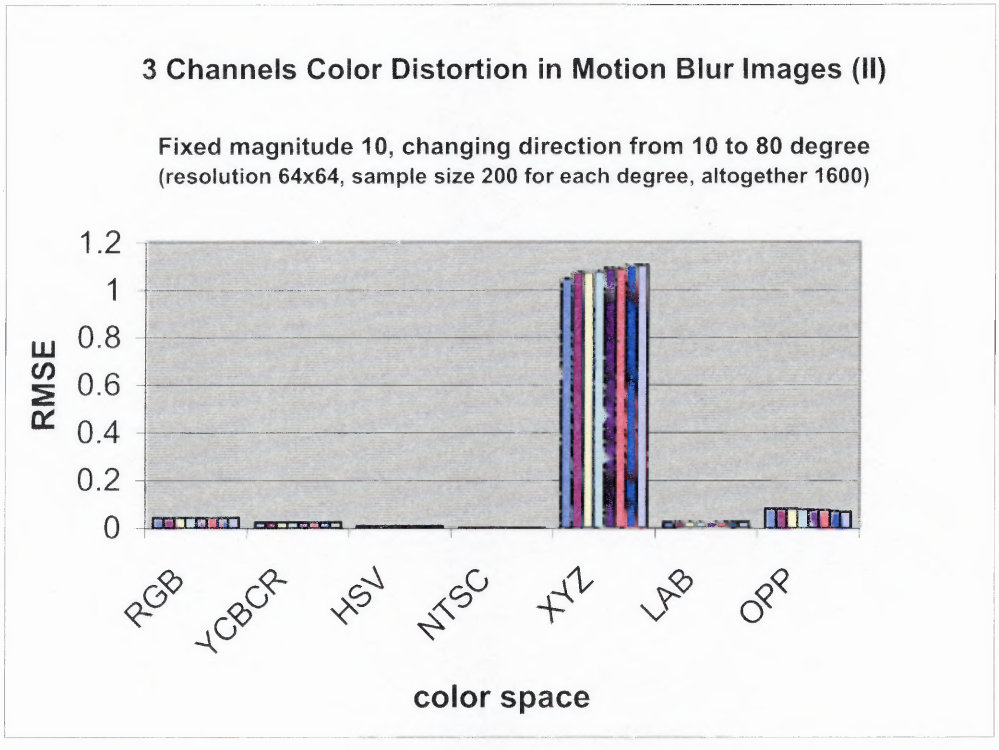
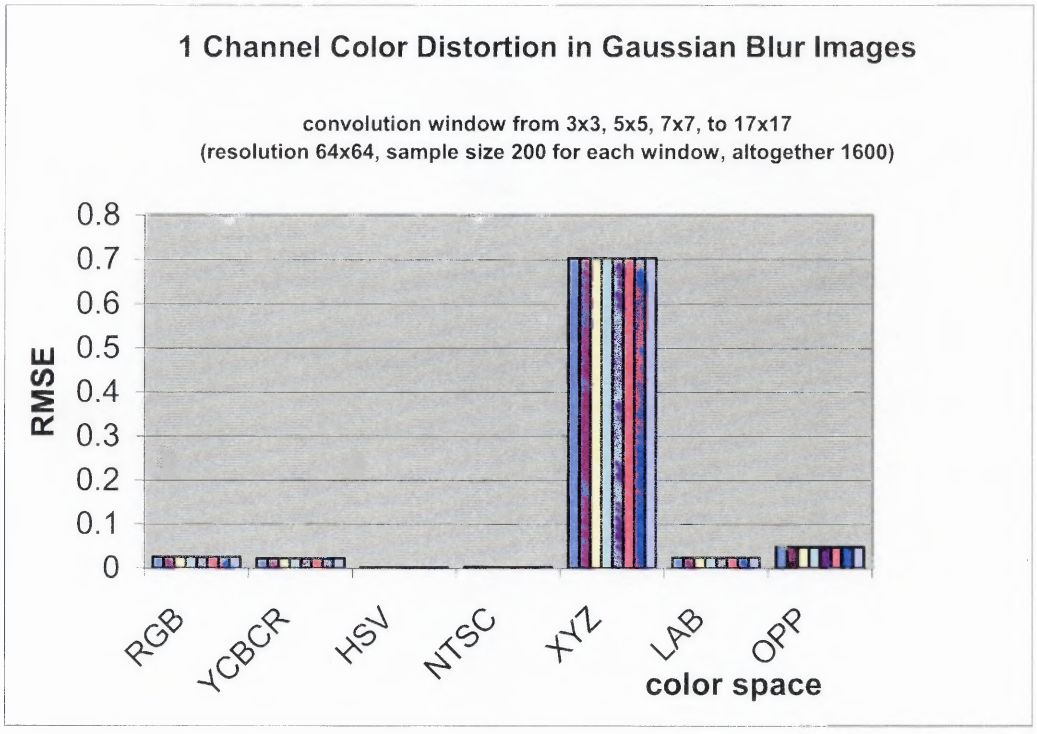


Figure 5.9 Color distortion due to variable orientations in motion blur images. (1 channel versus 3 channels)



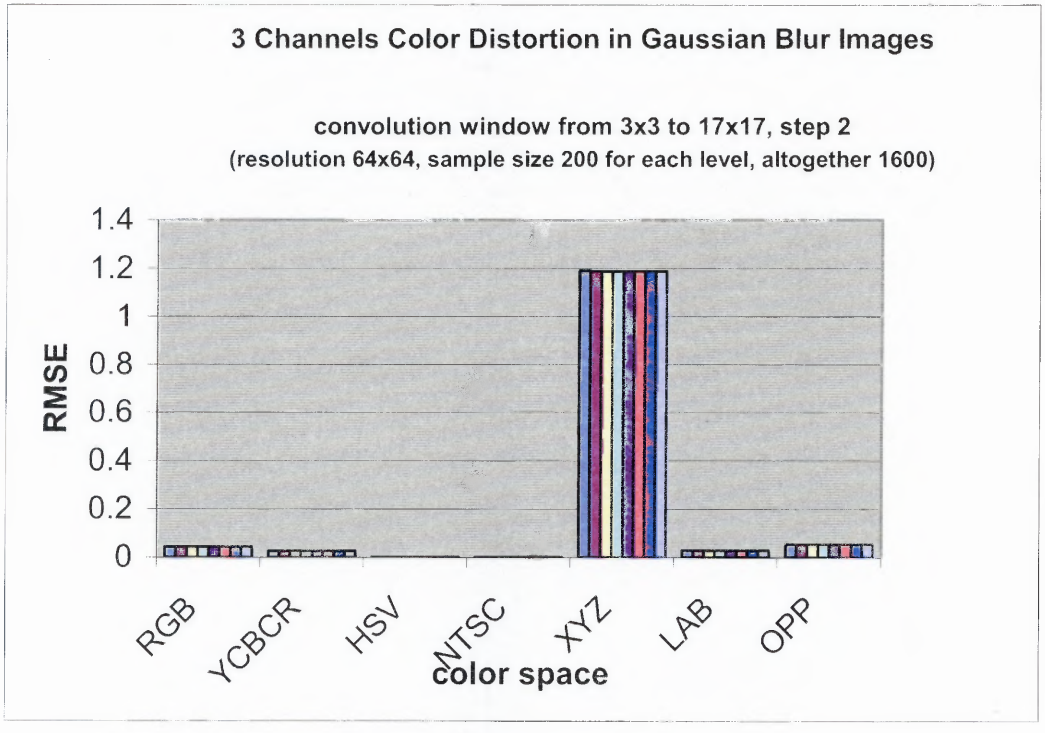


Figure 5.10 Color distortion due to different window size in Gaussian blur Images. (1 channel versus 3 channels)

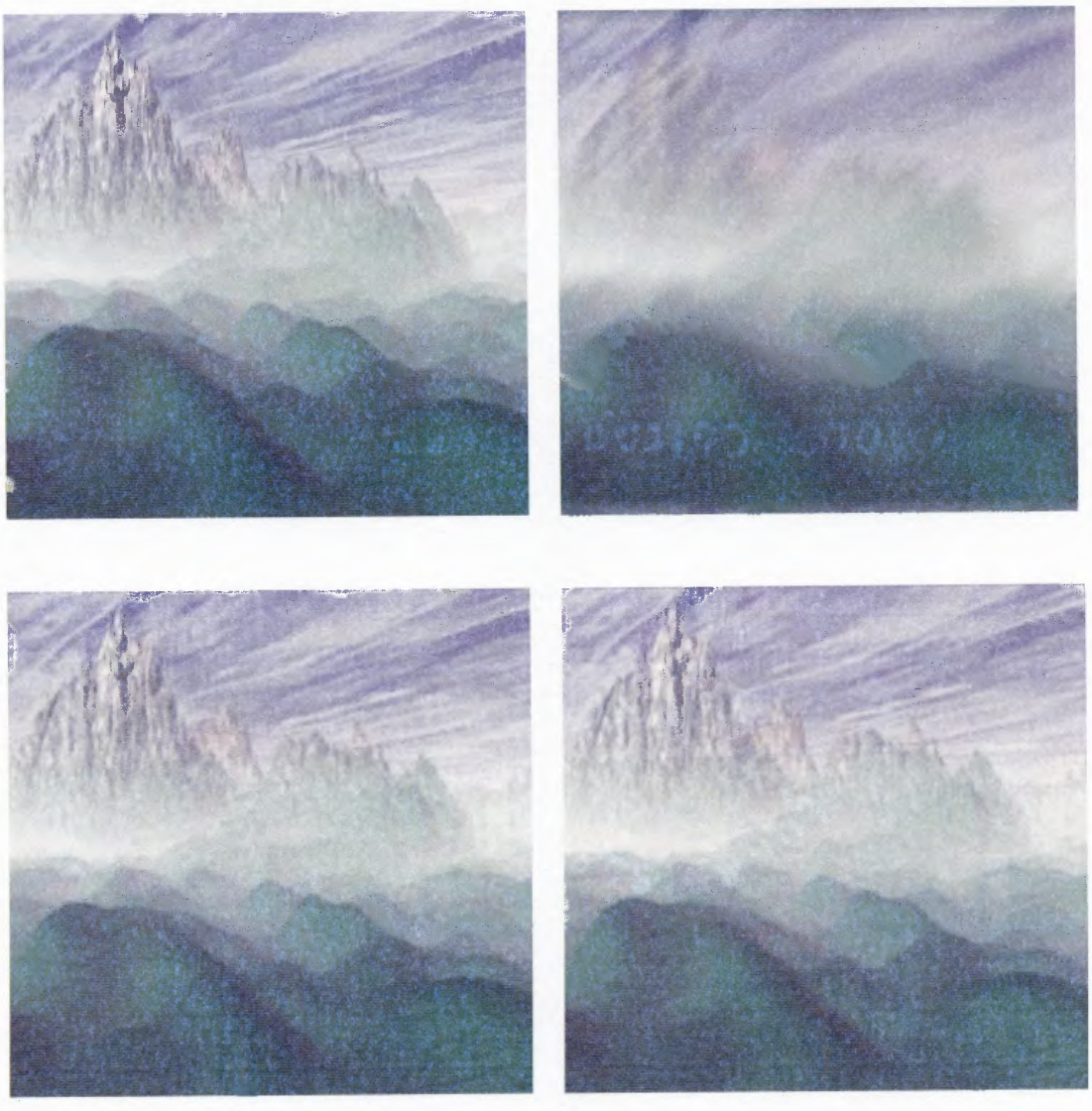


Figure 5.11 Referenced color correction (RCC) – sample result 1.
Top left: the original image Top right: after motion blur
Bottom left: after preprocessing Bottom right: after CC+MPA

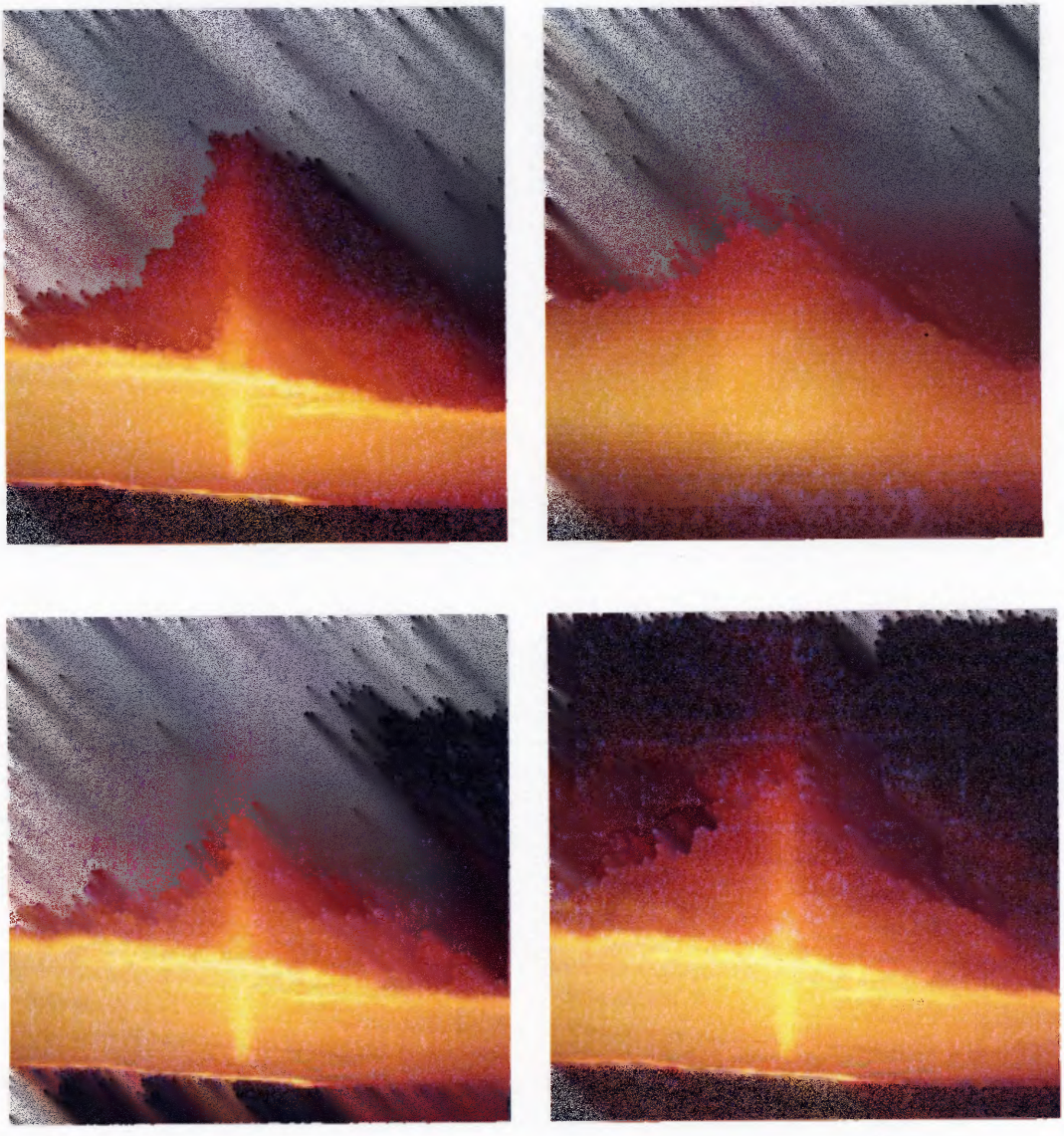


Figure 5.12 Referenced color correction (RCC) – sample result 2.
Top left: the original image Top right: after motion blur
Bottom left: after preprocessing Bottom right: after CC+MPA

Figure 5.13 No-reference color correction (NRCC) – sample result 1.

Top: JPEG image with compression ratio 200:1

Middle left: after Median filter Middle right: after Wiener filter

Bottom left: after AUE algorithm Bottom right: after Color Correction

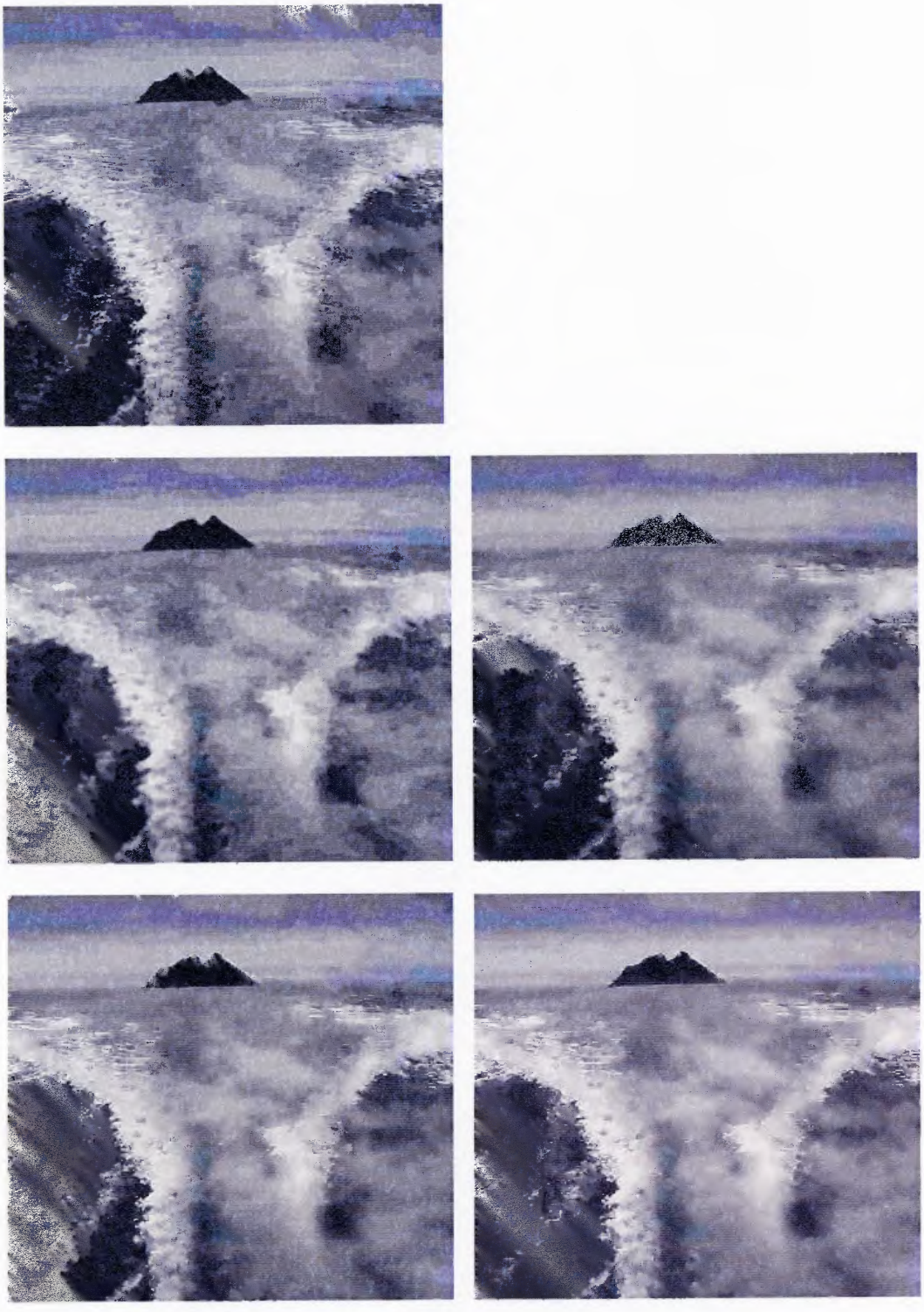


Figure 5.14 No-reference color correction (NRCC) - sample result 2.

Top: JPEG image with compression ratio 200:1

Middle left: after Median filter Middle right: after Wiener filter

Bottom left: after AUE algorithm Bottom right: after Color Correction

In this chapter, we present a comparative assessment of color distortion in different color spaces for blurred images. Experimental results indicate that color distortion increases as the magnitude of motion blur increases; color distortion remains in the same level as the window size of Gaussian blur increases. Once the color distortion is above the JND, we suggest the distorted image be converted to either HSV or NTSC color space for less distortion in monitor. Meanwhile, we present two color correction methods. One of which is used in a referenced case; whereas the other is used in a no-referenced case. We show that it is possible to take care two goals at the same time: (1) minimize the existing color distortion; (2) image enhancement. Experimental results show that the proposed color correction methods can achieve a relatively lower distortion. RCC can correct distortion by 88.3 % and NRCC can correct distortion by 60.71 %. Future research will be focusing on no-reference color reproduction in JPEG images to minimize total colors for saving download and display time.

CHAPTER 6

COLOR IMAGE RETRIEVAL VIA ROI CHAIN CODE HISTOGRAM, DC IMAGE HISTOHGRAM, AND COLOR HISTOGRAM

To commercialize current content-based image retrieval (CBIR) systems, researchers pinpoint some underlying problems and issues. Of the problems CBIR faces now, scalability lies in the heart. However, scalability relies heavily on the high precision of the retrieval result. The first objective of this chapter is attempting to develop a simple yet accurate feature representation method that will be used for indexing when a huge volume of new images are added up to the image database.

To meet the high precision requirement, it is essential to figure out the best combination of features. From the previous research results in CBIR area, we know that color and shape are a sound combination to boost the precision. However, most of the pre-established results are based on the color histogram and shape information in the global level; in other words, they are not based on the color and shape in object itself. The second objective of this chapter is to narrow down the feature domain to the local level, and then retrieve the color, texture and shape features based on the object itself.

The rest of this chapter is organized as follows. In Section 6.2, a brief literature survey with respect to the feature extraction is conducted. Section 6.3 presents our proposed methods. Section 6.4 provides experimental design and results to validate our methods. Section 6.5 concludes with a summary and future work.

6.1 Definitions and Assumptions

Image Retrieval: A technology to retrieve similar images by some search key that is a feature of the image data, such as form, color and pattern.

Content-Based Image Retrieval: Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. "Content-based" means that the search makes use of the contents of the images themselves, rather than relying on human-inputted metadata such as captions or keywords. .

Assumption of CBIR: Most of the CBIR methods rely on the assumption that human will see the center of image. Hence Region of Interest (ROI) area can be chosen in the center as a rectangle or square box.

6.2 A Brief Literature Survey on Feature Representation

Feature extraction is the key step in CBIR. According to literature, traditional feature representation is mainly based on three feature descriptors: color, shape and texture. Among them, color descriptor has relatively more discussions. It is worth noting that although color descriptor has a relatively large impact on precision; for instance, shape descriptor has a reasonably good precision in the trademark application. However, till now any single feature descriptor does not outperform the mixture of any two or three descriptors. In this brief literature survey, we do not pay much heed to a detailed exploration of all the publications so far; instead, we attempt to discuss the relative

advantage and drawback with respect to a certain feature representation, so that it is possible to find out some complementary feature descriptors to boost the performance.

Color histogram is the most widely used color descriptor for color representation. It can be divided into two categories: one is the global color histogram, and the other is the local color histogram. The former suffers greatly from the mixture between background and objects, so the query results may often have the problem of different objects in the similar color backgrounds. The latter has already released most of the above problem and has gained a better precision, but the local color histogram still does not contain shape and texture information of the object. Both global and local histogram methods also have the drawback that they are influenced a great deal from the illumination. The reason why the local histogram is still receiving researchers' attention is due to the fact of low computational complexity. Meanwhile, if the local histogram can cope with some other feature descriptors, for example, shape, then the query precision will be always boosted.

Shape representation, in general, is the most difficult feature descriptor because image segmentation tasks have a high computational complexity, and the shape feature extracted from the image segmentation is error-prone. Usually the problem is rooted from the fact that automatic segmentation is hard to clearly separate foreground and background for each different image. As a result, the result after segmentation sometimes may introduce more than enough details for shape representation. Basically, shape representation falls into two categories: shape through feature vectors and shape through transforms. Further more, shape through feature vectors can be divided into internal representation and external representation. The most famous internal representation is the

use of digital moment by IBM QBIC system, and the use of coefficients of 2D discrete Wavelet transform by Daubechies; the most prominent external representation method is the chain code contour method and 1D Fourier descriptor. Recent advance is the use of chain code histogram to model the objects with irregular shape.

Texture representation can be further divided into three categories: spatial domain methods, frequency domain methods, and texture signatures like contrast, coarseness and directions. Most of the spatial domain methods use the statistical property of brightness patterns. Among them, MRSAR (Multiple Resolution Simultaneous Auto Regressive) method has been demonstrated to offer the best intra-class recognition rate. The most prominent method for frequency domain analysis is the power spectrum method, which creates power spectrum density function in frequency domain. Coefficients of a 2D transform will indicate the correlation of a brightness pattern in the image. Since the degree of texture coarseness is in proportional to a spatial period, coarse texture will have spectral energy focusing at low spatial frequencies; fine texture will be focusing at high spatial frequencies. In addition, Gabor wavelet decomposition model has been shown to be almost the same recognition power as that of MRSAR method.

6.3 The Proposed Methods

Before we jump into the description of our algorithms, it is essential to state the research assumptions. Assumption 1 is that we assume there is only one object in each image in the database; once if more than 1 object locates in the image, then only the object with the largest area will be counted. Assumption 2 is that the size of the object occupies at least one fourth of the image size. These two assumptions allow us to narrow down the

query domain into the Region of Interest (ROI), so that it is possible for the local color descriptor and local shape descriptor to be extracted.

Based upon the above assumptions, our method is based on the similarity measure on the combination features of color, shape and texture within the ROI part. Our proposed content-based image retrieval system diagram is in Figure 6.1.

ROI extraction – The most important task before feature extraction is ROI extraction. In our definition, since the assumption of our system is only one object within the image and no user interaction is allowed, hence ROI is the entire object itself. The extraction of ROI task mainly becomes the segmentation of foreground and background. Note that this definition can reduce a great number of online ROI feature extractions into the offline ROI feature extractions. The detailed ROI extraction procedure is as follows:

The ROI Extraction method: (for luminance)

1. Convert a color image from RGB to YC_bC_r color space.; divide the luminance channel (i.e., Y component) into blocks of 8×8.
2. Compute DCT on each block. Let $C_{m,n}$ denote the DCT coefficient at location (m, n), where $0 \le m \le 7, 0 \le n \le 7$. Note that the origin is placed at the upper-left corner. Let V be the sum of two vertical columns along the left border of the block, H be the sum of two horizontal rows along the upper border, and D be the sum of the diagonal direction index. The detailed definitions are as follows [21]:

$$V = \sum_{m=2}^{7} \sum_{n=0}^{1} C_{m,n};$$

$$H = \sum_{m=0}^{1} \sum_{n=2}^{7} C_{m,n};$$

$$D = \sum_{k=2}^{5} C_{k,k} + \sum_{k=2}^{5} C_{k,k-1} + \sum_{k=2}^{5} C_{k-1,k}$$

Compare the horizontal, vertical, and diagonal energies and select their maximum as $M = \max\{V, H, D\}$.

Let α and β be the separating constants such that each block can be divided into the following categories:

Flat block,	if	$M \leq \alpha$;
Texture block,	if	$\alpha < M < \beta$;
Edge block	if	$M \geq \beta$;

3. Find the upmost and downmost, leftmost and rightmost edge block, and create a rectangle as ROI. Later on, all the color histogram, chain code histogram and DC image histogram will be based on this ROI area, instead of the original image size.

ROI Color Histogram (ROI_CH) – Color histogram of a given grayscale image has been defined in the traditional color histogram based CBIR. Here, the ROI color histogram is a 1-D vector histogram counted from the concatenation of Y-CB-CR three channels of ROI area. It is important to normalize the ROI Color Histogram by dividing each color in the histogram over the total ROI pixel count.

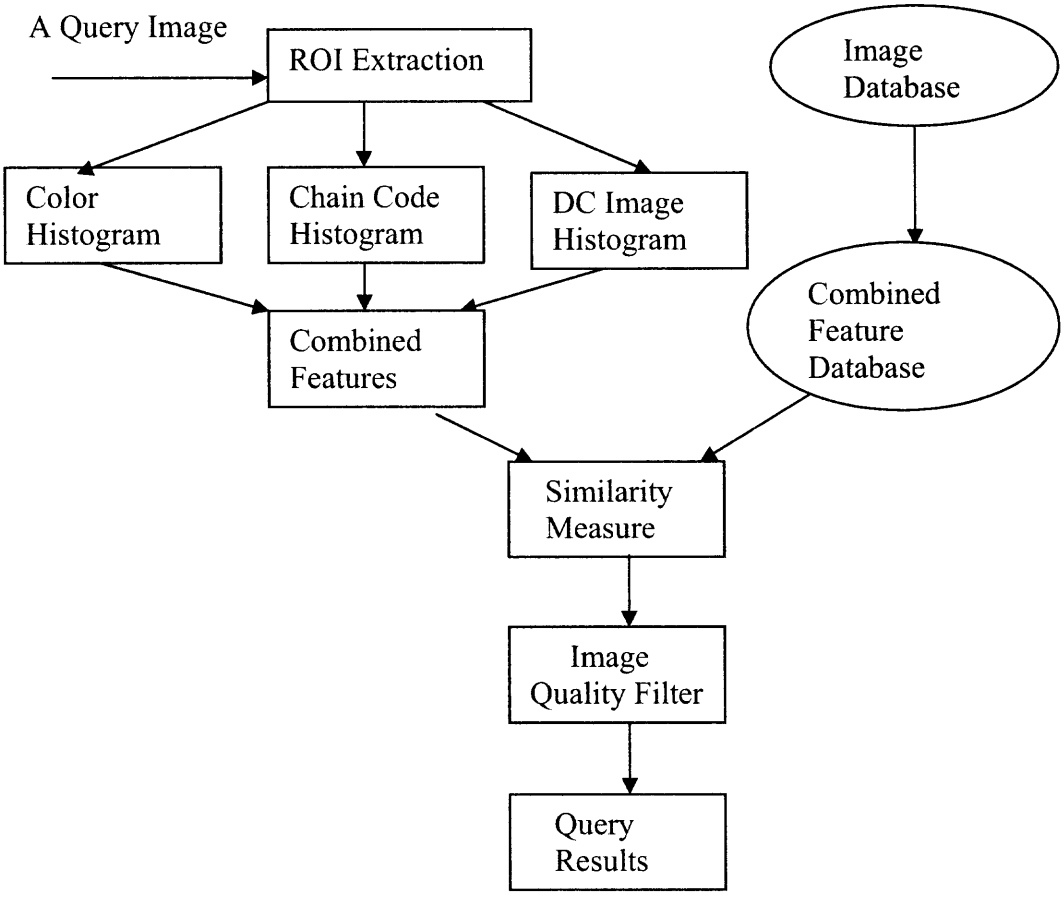


Figure 6.1 Our proposed CBIR system diagram.

ROI Chain Code Histogram (ROI_CCH) – Chain code histogram has the capability of translation and scaling invariant. Meanwhile, the curvature from the object is also rotation-invariant. The entire tasks for ROI Chain Code Histogram (CCH) and Curvature Histogram (CUH) are described as follows:

1) Obtain chain code from the binary contour image, which was extracted from the ROI image segmentation part;
2) Construct the 8-bin chain code histogram according to each direction count;
3) Obtain the curvature based on the difference of chain code;
4) Construct the curvature histogram according to each of the direction count in the curvature;
5) The final normalization step. The idea is to get a ratio of its shape direction, so that the shape can be scale-invariant. Now both CCH and CUH will be divided by its corresponding total count.

ROI DC Image Histogram (ROI_DCH) – DC image histogram is a powerful tool for image texture analysis, especially stable when the area has been reduced to a limited region within the edges. We apply DC image histogram on top of ROI. The entire tasks of ROI DC image histogram formulation are described as follows:

1) Divide an image into blocks of 8 by 8 pixels;
2) Perform Discrete Cosine Transform (DCT) on each block;
3) Generate the DC image from the DC coefficients of all 8×8 blocks;
4) Construct DC image histogram by quantizing the DC values into N bins and counting the number of coefficients falling into each bin. In our CBIR system, altogether 15 bins are used;
5) Normalization of DC histogram: It is the step that the bin count in each bin is divided by the total number of pixels in DC image.

The Combined Feature Vector -- A 1-D vector obtained from the concatenation of color histogram, chain code histogram, curvature histogram, and DC image histogram. Hence, it is easy to see that this combined feature vector is a mixture vector of color, shape, and texture. It is worth noting that since either ROI color histogram and DC image histogram has 256 dimensions, therefore in our system we have to multiply a scaling factor 11 for chain code histogram (8 dimensions) and curvature histogram (15 dimensions) so that altogether we have 23*11 = 253 dimensions, which are roughly equal to 256 dimensions for color histogram and DC image histogram.

The Image Quality Filter -- The quality filter serves as the final screening process of the query results. The color image quality measure method we apply here is the method we have discussed in Chapter 4 – a no-reference quality index for color image, which is determined by a linear regression model of three parameters, namely, sharpness, contrast, and colorfulness. To serve the specific purpose of different application, of course the demand for sharpness, colorfulness, and contrast could be subjective to be tuned. In our method, the tuning process is nothing but a simple giving of the corresponding weight to the three variables. But here, as a default rule, we set the accepted image quality as 0.5.

6.4 Experimental Design

Experimental design aims at understanding the performance of our proposed methods. 70 images (10 objects) in JPEG format are randomly selected from Google image search engine. Each object is gone through scaling and rotation; that is, there are 70 scaling images and 70 rotation images created manually, which make altogether 210 color

images in our object image database. The language we use is MATLAB and the OS platform is SUN Solaris 2.0. All the color images are resized to 256×256 resolution.

Two instrumentation tasks have to be set up here. One is the definition of precision of the query, and the other is the similarity measure.

Precision of Query – (number of retrieved images – number of false positives) / number of retrieved images;

Similarity Measure – The Euclidian distance measure is used; that is, L2 distance will be measured between the feature vector of the query image and the feature vector of each image in the image database.

6.5 Experimental Results and Discussions

Precision is the most practical issue in content-based image retrieval (CBIR). This chapter attempts to tackle the precision problem via feature representation by the combination of color, shape and texture information from Region Of Interest (ROI). Experimental results in Table 6.1 demonstrate that the corresponding precision for chain code histogram, DC image histogram, and color histogram are respectively 51%, 56%, and 57%, if top 10 matches are considered. The precision of the combined feature vector is 72 %. This result shows that no single feature can exceed 60 % precision; however, if they are combined together, the precision can be lifted up to 76 %. The gains are 25 %, 20 % and 19 %, respectively.

Experimental results in Table 6.2 demonstrate that the precisions for chain code histogram for top 1, top 2, top 3 matches are 100%, 94%, and 70%, respectively; the precisions for DC image histograms for top 1, top 2, top 3 matches are 100 %, 100%,

88%, respectively; the precisions for color histograms for top 1, top 2, top 3 matches are 100 %, 100%, 93%, respectively. It is worth noting that since our image database is relatively small, only 210 images, the high precisions are not a surprise. We also observe that in chain code histogram, from the top 3 on, the precision experiences a sharp fall; this shows that chain code together with its curvature is not stable if only shape information is considered. One interesting phenomenon is that although the top 1 match for these three single feature methods is totally correct, they differ in that the matching image is most of the times not the same image.

Table 6.1 Precision Comparisons on Four CBIR Methods

Feature-based Method	Color Histogram	DC Image Histogram	Chain Code Histogram	Our Method	Image-based Method
Precision	57 %	56 %	51 %	76 %	51.5 %

Table 6.2 Precisions for Top 3 Matches on Four CBIR Methods

	Color Histogram	DC Image Histogram	Chain Code Histogram	Our Method
Top 1 Match	100 %	100 %	100 %	100 %
Top 2 Match	100 %	100 %	94 %	100 %
Top 3 Match	93 %	88 %	70 %	98 %

How much progress we have to make from the precision of our proposed method? First of all, even though the combined feature vectors have been extracted offline; when

the image query is performed, it took in average 57 seconds for one query. This indicates that the running time issue is the biggest problem for improvement. Second, the dimension of the combined feature vector is still too large, not only too time-consuming in image comparison, but also a lot of redundancies existed. Hence, the second issue that needs to be improved for later performance tuning is the elimination of redundant dimensions. Third, throughout the research, we did not consider the matching criteria-the matching function; all we used is the standard L2 distance measure. It is not enough for a detailed comparison of the precision. Hence, we believe that the third issue to be improved is to compare more than 5 matching functions, and the best matching function can boost the precision to a small percentage.

Meanwhile, if industry need is the most important issue, probably we need to incorporate the automatic indexing method to tune the running time for large scalable image database. Since indexing methods are solely the focus of database research, it is almost inevitable to have a joint effort of computer vision and database in the near future.

CHAPTER 7

COLOR REDUCTION VIA INTELLIGENT ENCODER -- AN IMAGE SPECIFIC COLOR QUANTIZATION METHOD

7.1 Definitions and Assumptions

Currently, the bottleneck of slow download and display time of JPEG and JPEG2000 images is due to 16.7 million of colors, which were originally designed to enrich the display colors for human perception. To alleviate this problem, it is necessary to incorporate certain mechanism in an encoder, so that the overall load of the decoder can be reduced. Besides, for some display device, only 256 colors can be displayed by 8-bit color gamut. As a consequence, JPEG 16.7 million colors have to be reduced to 256 colors to meet the hardware requirement. Therefore, it is crucial to develop color reduction method to be inserted into the encoder, so that the download time as well as the decoding time can be greatly reduced.

To design a JPEG/JPEG2000 encoder, we must not only incorporate the automatic color reduction mechanism, but also take full advantage of human vision system to allow color compression to be image-specific. The knowledge such that human perception is less focused on image background and likely to be deceived in texture instead of edge, can help design an image-specific color down-sampling method. This chapter is intended to apply artificial intelligence to the JPEG encoder for high color compression rates while the original perception quality is preserved.

The selection of color quantization algorithm is an important issue for color image quality. A large amount of color quantization methods have been proposed. Wu's

dynamic programming algorithm [Wu, 1992] is optimal in terms of image quality and display time. The color quantization modules are not part of the JPEG decoding, but are essential for output images to be color-mapped file formats in display hardware. The JPEG software developed by the Independent JPEG Group (IJG) applies two-pass color quantization algorithm. It consumes lots of memory and often results in memory short. Many codecs have been built on top of the IJG JPEG library. However, human perceptual optimization has not yet been considered. Although the NASA DCTune method [Watson, 1994] provides YOZ color space and transfers the quantization error into the perceptual color space, it only creates perceptual quantization matrix for the quantization step. It does not include the optional color down-sampling on chrominance channels, because this greatly complicates the optimization process. In this chapter, we incorporate perceptual knowledge into the down-sampling procedure in order to make color quantization more image-specific.

The rest of this chapter is organized as follows. In Section 7.2 we briefly review literature. Our methods are presented in Section 7.3. Section 7.4 describes experimental design. Section 7.5 provides experimental results and discussions. Section 7.6 draws conclusions.

7.2 A Brief Literature Survey

Programs using screen resolution of 256 colors are popular, because they offer a good compromise between memory requirement and processing speed on one hand and image quality on the other hand. The reduction process, from JPEG 24-bit true color RGB mode to 16-bit or even 8-bit color RGB mode, considers the tradeoff between image quality

and communication speed. It is called "palette design". After that, it is essential to perform "pixel mapping" to map the original pixel color to the color from the color palette. The goal of palette design is to come up with the least sufficient colors capable of representing the original true colors; whereas the goal of pixel mapping is to minimize the mean square error (MSE) between the original and the color reduced images.

7.2.1 Palette Design

Two design paradigms have been widely utilized in palette design. One is the universal color palette [Kolpatzik, 1995], which is capable of rendering multiple color images simultaneously, and the other is adaptive color palette [Murching, 1994], which is image dependent and time consuming in the design process, but usually provides images with better quality. Based on the knowledge and techniques used, the adaptive color palette can be divided into three groups:

(1) Pre-clustering scheme. It applies all sorts of statistical analysis of the pixels in the color space. Often-used methods include the popularity, median cut and octree methods [Heckbert, 1980;Gervautz, 1990].

(2) Post-clustering scheme. It contains initial condition and selection of one or combined data mining and pattern recognition methods. The initial condition is the selection of the initial color from which an iteration process starts. Pattern recognition method is applied to conduct the pattern finding and promise of the ending condition, where mean square error (MSE) is minimized. The k-mean algorithm, self-organization map, fuzzy logic, neural network and principal component analysis (PCA) are the methods fallen into this group [Linde,

1980;Dekker, 1994;Hwang, 2002;Ozdemir, 2002;Manian, 2002;Li, 2003; Paramarkos, 2002;Kanjanawanishkul, 2003;Tasdizen, 1998].

(3) Scalar quantization scheme. It is simple and against vector quantization. For example, the sequential scalar quantization (SSQ) approach [Balasubramanian, 1995] reduces computation by allowing vector quantization to be performed with a series of scalar quantizers. The resulting color palette is then combined with a previously developed multi-level error diffusion algorithm to give the best overall result.

It is worth noting that all the above methods fall into the category of "Gamut Compression." There is yet another category called "Gamut Clipping," which in general exploits a clipping function, either linear or nonlinear, to clip the input gamut into a subset of the original gamut.

7.2.2 Pixel Mapping

Pixel mapping (Gamut mapping) [Morovic, 2001] performs the mapping of the color range from the input image to the palette. It can be categorized into 5 methods:

(1) Reduction of color space dimensionality: PMRC method [Cheng, 2001].

(2) Brute-force search: Exhausted search the best color which gives the least mean square error.

(3) Globally or locally sorted search: Heckberg method [Heckbert, 1980].

(4) K-d trees: Friedman method [Friedman, 1997].

(5) Cluster mapping: It maps a given pixel to the center of its belonging cluster.

7.3 The Proposed Methods (An Image Specific Method)

Two methods are presented in this chapter. One is the DCT Classification (dctC) method, and the other is the Modified Median Cut (MMC) method.

7.3.1 DCT Classification Method (dctC)

For the luminance channel, we propose a classification method to reduce more colors in the flat and texture blocks (similar idea yet different application can be found in [Shen, 1998]). Since luminance is more sensitive to human vision, JPEG standard does not suggest any modification on luminance and only 4:1 ratio chrominance channels color quantization is suggested to be an optional feature for the baseline JPEG. Recently more and more papers have discussed the issues on color quantization for human perception without Just Noticeable Difference (JND) [Zhao, 2004;Zhou, 2004;Tremeau, 1994;Sun, 2002;Ebner, 1997;Morovic, 1997;Morovic, 1998;Braun, 1999]. For two chrominance channels, we propose a golden-ratio method to reduce colors based on the factor of golden ratio [Vajda, 1989].

Phase 1: The Color Space Selection Phase

Since JPEG/JPEG2000 uses YC_bC_r as the computational color space, we continue this convention. Another reason is that YC_bC_r outperforms RGB space in terms of image quality according to experiments.

Phase 2: The Deterministic Classification Phase: (for luminance)

1. Divide the luminance channel (i.e., Y component) into blocks of 8×8.
2. Compute DCT on each block. Let $C_{m,n}$ denote the DCT coefficient at location

(m, n), where $0 \le m \le 7, 0 \le n \le 7$; note that the origin is placed at the upper-left corner. Let V be the sum of two vertical columns along the left border of the block, H be the sum of two horizontal rows along the upper border, and D be the sum of the diagonal direction index. The detailed definitions are as follows [21]:

$$V = \sum_{m=2}^{7} \sum_{n=0}^{1} C_{m,n};$$

$$H = \sum_{m=0}^{1} \sum_{n=2}^{7} C_{m,n};$$

$$D = \sum_{k=2}^{5} C_{k,k} + \sum_{k=2}^{5} C_{k,k-1} + \sum_{k=2}^{5} C_{k-1,k}$$

Compare the horizontal, vertical, and diagonal energies and select their maximum as $M = \max\{V, H, D\}$.

Let α and β be the separating constants such that each block can be divided into the following categories:

Flat block,	if	$M \le \alpha$;
Texture block,	if	$\alpha < M < \beta$;
Edge block	if	$M \ge \beta$;

We adopt $\alpha = 20$ and $\beta = 30$ according to our experiments.

We define Color Reduction Ratio (CRR) as

4:1	if it is a flat block;
2:1	if it is a texture block;
1:1	if it is an edge block.

3. Based on the block category and CRR, each 8×8 block is further divided into sub-blocks and the mean value will be assigned to each pixel in the sub-block. For example, if it is an edge block, keep the block pixels as they are; if it is a texture block, an 8×8 block will be divided into four 4×4 sub-blocks; if it is a flat block, the block will be divided into sixteen 2×2 sub-blocks.
4. Further color reduction is performed by designing a color palette according to the Median Cut method [5] using the allowable number of colors. For instance, if a 16-bit display device is used, then 6 bits (i.e., up to 64 colors) are selected for the luminance channel, and if an 8-bit display device is used, then 3 bits (i.e., up to 8 colors) are selected.
5. Perform pixel mapping according to its nearest neighbor in the color palette.

Phase 3: The Golden Ratio Phase (for chrominance)

1. For each chrominance channel, apply 8:1 down-sampling for each 8×8 block. That is, each pixel in the 8×8 block is assigned by the mean value of this 8×8 block.
2. Build its corresponding color histogram.
3. For the color palette design, we apply the Median Cut method to select the total number of colors to be up to $\mathbf{(0.380859375)^N}$ of the total colors [8]. For instance, for 256 colors the range of color palette (RCP) is 97 colors when N=1; RCP is 37 when N=2.

4 Use the selected colors to reconstruct the chrominance channel.

5. Let each pixel in the unselected color block be assigned by its nearest neighbor from the color palette.
6. Repeat steps 3-5 until the total number is close to the allowable colors in the display device. For instance, if a 16-bit display device is used, then altogether 10 bits up to 1024 colors can be selected for the two chrominance channels, and if an 8-bit display device is used, then altogether 5 bits up to 32 colors can be selected for the two chrominance channels.

7.3.2 Modified Median Cut Method (MMC)

Our proposed MMC method modifies the cutting of the longest dimension to all three dimensions (channels). The detailed procedure is described as follows:

1. Convert color space from RGB to YC_bC_r.
2. For each channel, a 2-D array will be converted to a 1-D array.
3. Sort the 1-D array by descending the order.
4. Compute the number of columns needed for the reconstructed 2-D array according to the pre-defined bits. For instance, if the bits for three channels are 3-2-3, then the colors for 3 channels are 8-4-8. A floor function is defined as below:

   *Column_reconstruced = floor (column * row / numColors_intheChannel);*
5. Reconstruct a 2-D sorted array by row major based on the following two dimensions: *Column_reconstructed* and *numColors_intheChannel.*
6. For palette design, we find the medians from the reconstructed 2-D sorted array. The total number of medians is the number of colors in the channel.

7. For pixel mapping, if the pixel is not one of the medians, it is mapped to the nearest median; otherwise, the value remains the same.

7.4 Experimental Design

Experimental design aims at understanding the performance of our two proposed methods in the following aspects: (1) Image quality in terms of MSE and PSNR, and (2) under the same PSNR, the performance of the color compression of our method compared to others.

In the first experiment, 100 natural scene JPEG color images are randomly selected from Internet. The language we use is MATLAB and the OS platform is SUN Solaris 2.0. All the images are resized to 256×256×3 resolution. The two proposed methods intend to reduce to (1) 16 bits or (2) 8 bits. We compare the PSNR of our methods with three established known methods: popularity method, min variance method and uniform method. The average PSNR for each method is compared both on 16 bits and 8 bits settings.

In the second experiment, 30 JPEG color images are controlled under the following two quality settings: PSNR = 30 and PSNR = 35. We also compare the average image size of our method with others.

7.5 Experimental Results and Discussions

The first experiment aims at reducing colors from 16.7 million (24 bits) colors to 65536 (16 bits) and 256 (8 bits) colors. The baseline JPEG color down-sampling (chrominance channels 4:1) method is first tested for quality in terms of PSNR. We use

this baseline JPEG method as a naïve approach to compare with other and our methods. Experimental results are listed in Table 1.

For the dctC method in 65536 colors, the average PSNR of the reduction images from 100 testing JPEG image samples is 36.907, which is similar to the Min Variance method of 36.083 and the Popularity method of 36.259. For the MMC method in 65536 colors, the average PSNR is 40.876, which is very close to 41.429 by the JPEG baseline down-sampling method. However, in 256 colors, the PSNR produced by our dctC method is 32.2111, which is far better than the Uniform method and Popularity method, but worse than the Min Variance method. The PSNR produced by our MMC method has the best quality, which is 36.912.

Table 7.1: PSNR Comparison for Different Color Quantization Methods

	Uniform	**Min Variance**	**Popularity**	**DctC**	**MMC**	**JPEG Downsampling**
PSNR_24bit						41.429
PSNR_16bit	36.034	36.083	36.259	36.907	40.876	
PSNR_8bit	25.782	34.777	30.250	32.211	36.912	

Note 1. Experimental results are based upon the average PSNR between 100 JPEG color images and their color reduced images.

Note 2. JPEG baseline down-sampling method applies 4:1 for two chrominance channels.

It is worth noting that the PSNR by the Min Variance method has the slowest decay rate. That is, from 16 bits to 8 bits, the PSNR by the Min Variance method is only dropped by 1.3, which is far smaller than 10.25 by the Uniform method and 6.009. Our

methods also have relatively slow decay rates, which have the PSNR decreases of 4.696 and 3.964, respectively. This slow decay rate is very important factor in color quantization algorithm design, since it can preserve as much quality when only a few colors can be allowed to be displayed due to hardware limitation. Most importantly, this factor allows our proposed method to be one component inserted into the JPEG encoder between color space transform and DCT procedure. With automatic color quantization method embedded in JPEG encoder, not only transmission time, but decoding time and display time can be saved, since now the bottleneck of JPEG display time is mainly restricted in 16.7 million colors.

One of the important observations is that our dctC method has very unstable PSNR for the image with different colors and background. This is due to the fact that dctC method classifies image pixels into three groups: edge, texture, and flat blocks. Therefore, if an image with plenty of objects in it, the color compression can not be high. As a result, the PSNR remains in the high level. On the other hand, if an image with a great deal of flat background, the compression can be very substantial. As a result, the PSNR will go down. Another important observation is that the dctC method can produce the least MSE in some images whose color ranges are small when compared with other methods. In summary, the dctC method is good at performing color quantization tasks for flat background and small color range; whereas the MMC method is good at reducing more colors while still preserving the high PSNR.

The second experiment aims at a comparative study of quantitative compression effect when fewer colors are used under the same PSNR. Detailed experimental results are shown in Table 2. Our dctC method uses in average 8192 colors to achieve PSNR=

35, which is only better than 32768 colors by the Popularity method. But dctC uses only 216 colors to achieve PSNR=30, it is better than 849 by the Uniform method and 256 by the Popularity method. The reason for not so promising performance is that the dctC method does no compression in edge blocks in the Luminance channel. Hence, if one image happens to have numerous objects in vivid colors, then the number of colors may not be reduced more than other known approaches.

Table 7.2 Number of Colors Used in Different Color Quantization Methods Under the Same PSNR

	Uniform	**Min Variance**	**Popularity**	**DctC**	**MMC**
PSNR=35	3375	260	32768	8192	216
PSNR=30	849	36	256	216	80

For MMC methods, it uses only 216 colors under PSNR=35. This is the least color used among all known methods. MMC also uses 80 colors to achieve PSNR=30, which is one of the best performance among all approaches. It is already very close to 36 colors made by the Min Variance method. The reason for this performance is that MMC method already considers all possible color combinations of Luminance and Chrominance channels. Meanwhile, MMC considers medians in three dimensions other than the longest dimension in the original median cut method. Last, the limitation of both dctC and MMC methods are in the time complexity. They consume more time than the Uniform and Min Variance methods provided that they are measured on the same

platform and written in the same language. This tradeoff makes our method difficult for real time application, but once we reduce the number of colors in the encoder by using our methods, the decoder can be implemented fast because of the less number of colors in JPEG. Other application areas are the high quality video encoding and fast decoding.

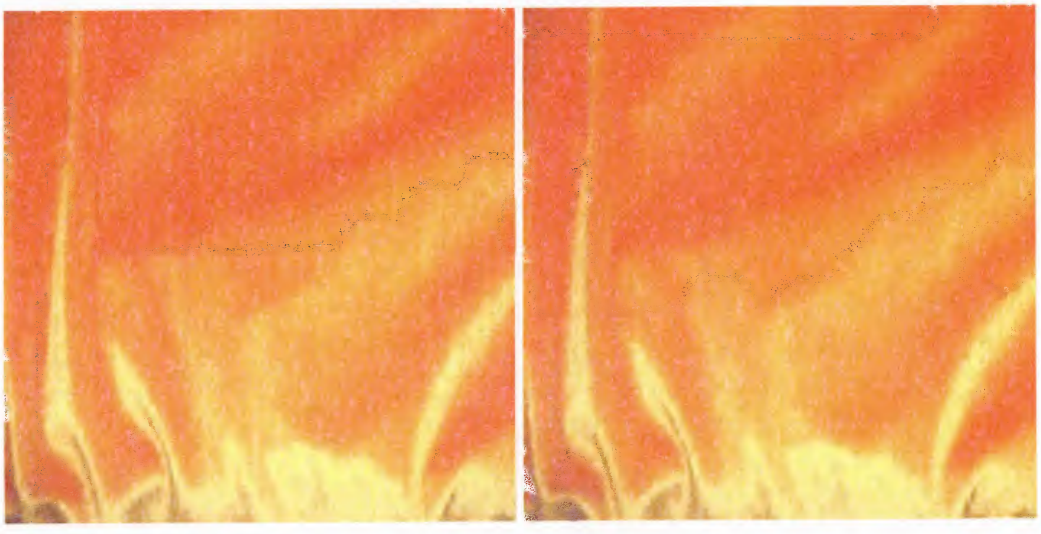
Figure 7.1 Sample experimental result for JPEG color quantization. (from 24 bits to 16 bits)

1st row: Left: Original JPEG color image: Flower
Right: Baseline JPEG down-sampling method

2nd row: Left: Uniform Method
Right: Min Variance method

3rd row: Left: Popularity method
Right: Modified Median Cut method

4th row: Left: DCT Classification method



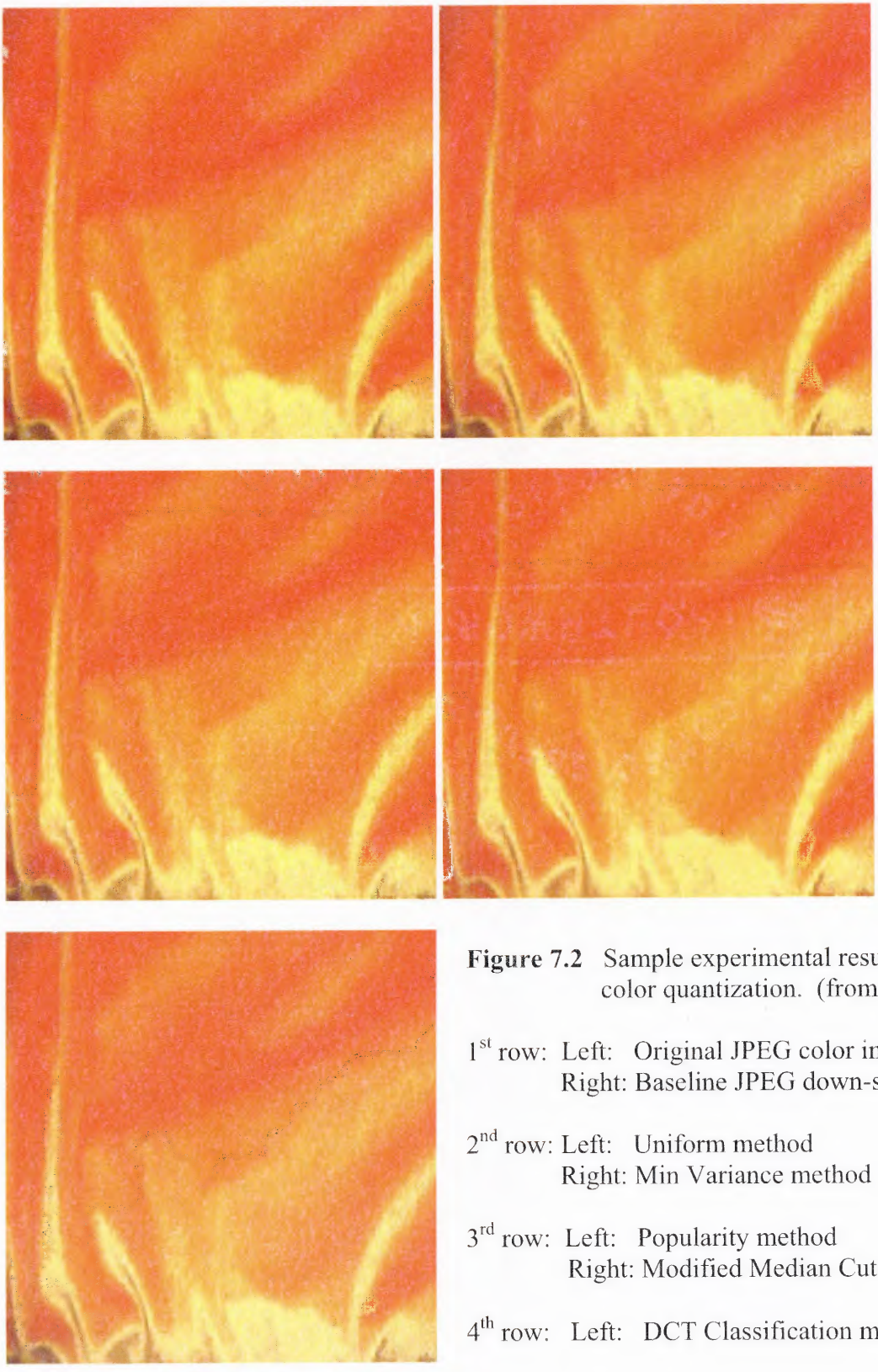


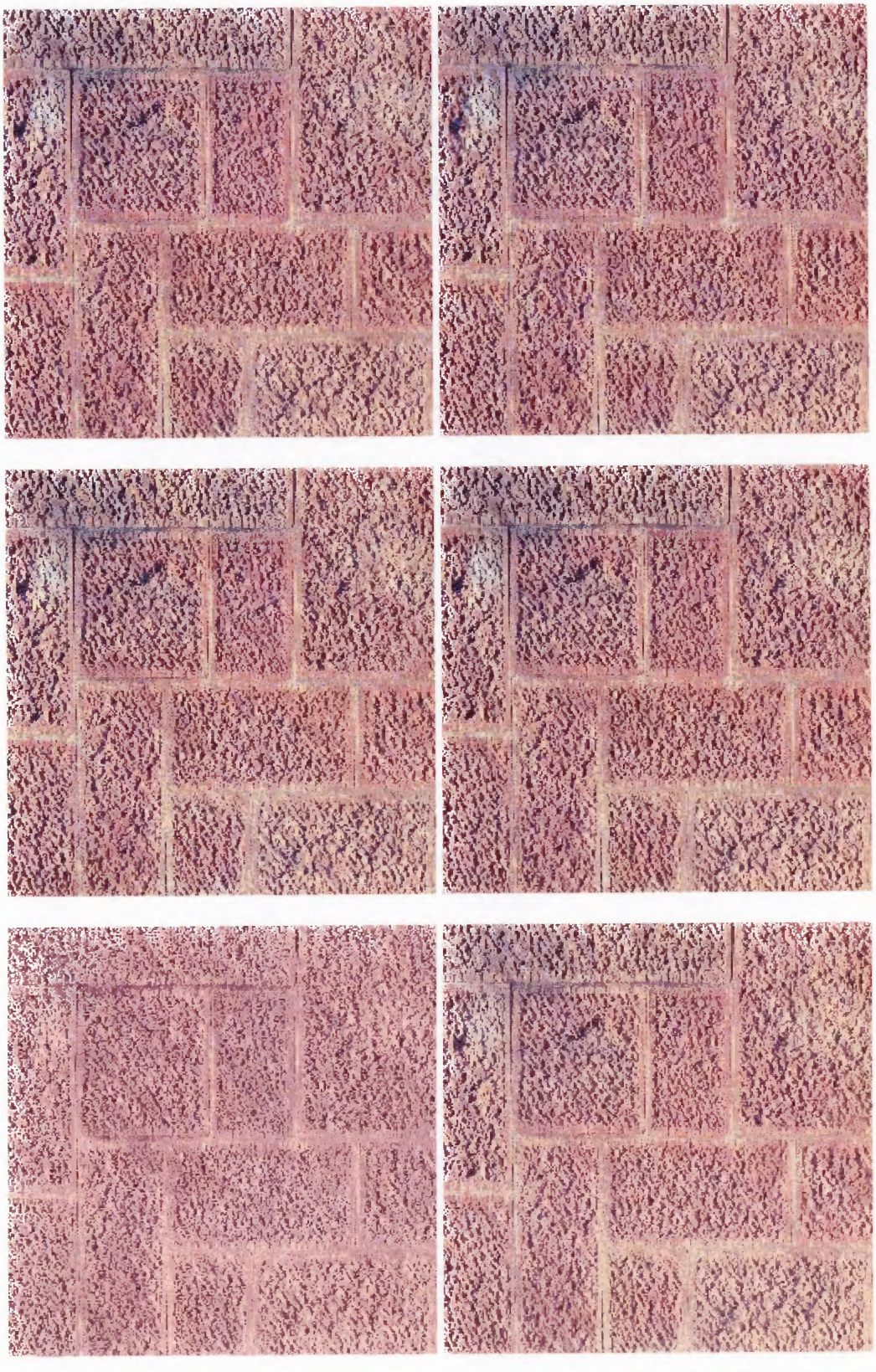
Figure 7.2 Sample experimental result for JPEG color quantization. (from 24 bits to 16 bits)

1st row: Left: Original JPEG color image: Velvet
Right: Baseline JPEG down-sampling method

2nd row: Left: Uniform method
Right: Min Variance method

3rd row: Left: Popularity method
Right: Modified Median Cut method

4th row: Left: DCT Classification method



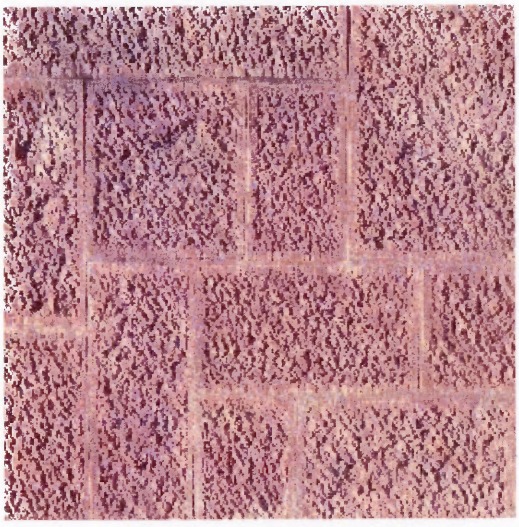


Figure 7.3 Sample experimental result for JPEG color quantization. (from 24 bits to 8 bits)

1st row: Left: Original JPEG color image: Wall
Right: Baseline JPEG down-sampling method

2nd row: Left: Uniform method
Right: Min Variance method

3rd row: Left: Popularity method
Right: Modified Median Cut method

4th row: Left: DCT Classification method

In this chapter, we present two color down-sampling methods. One aims at more color compression while preserving the original color perception. The other aims at good image quality in terms of PSNR. In encoding, the encoder first decides how many colors are sufficient for the specific display device, which includes how many colors are needed in both luminance and chrominance components. With this sufficient numbers in hand, the encoder then selects the appropriate color quantization algorithms between two proposed methods. The selection criterion is that if the input image can be compressed more in colors, then the DCT Classification method is selected; otherwise high quality will be our goal and the Modified Median Cut method will be selected. Other than the procedure for our color down-sampling method, the rest of the steps for encoder will be exactly the same as what the baseline JPEG standard suggests.

In the DCT Classification method, for luminance channel, an automatic deterministic classification method based upon DCT coefficients is proposed; for chrominance channels, a golden-ratio method is presented. The resulting image colors will be image-specific and subject to some properties. In general, the total number of colors for JPEG will be reduced from 16 million to a range from 65536 to 256 colors,

which is between 16 and 8 bits. Experimental results indicate that the DCT Classification method can compress more colors in the setting where the images contain fewer objects.

In the Modified Median Cut method, the color quantization results are quite stable from 24 bits down to 8 bits. On comparative assessment of some popular color quantization methods, we find that our proposed Modified Median Cut Method has relatively high quality with a cost of time complexity.

In summary, the proposed two color reduction methods can be incorporated and inserted into the JPEG/JPEG2000 encoder right after the color space conversion procedure, so that both the download time and decoding time can be reduced. Future research will be on the tuning for the running time of our two methods tailored to industrial needs.

CHAPTER 8

CONCLUDING REMARKS

8.1 Summary of Contributions

As described in the introduction section, color imaging science is a discipline of studying color vision, so as to create visual representation of objects for various digital devices. Since color vision is a combination effect of light, object and observer, it is essential to create visual representation of objects tailoring to a specific device. To this end, in this dissertation we have designed some visual representations as follows:

- A visual representation for color correction on low-bit rate JPEG images;
- An image representation for content-based image retrieval;
- An image-specific 256 colors (8 bits) reproduction for color reduction from 16 millions colors (24 bits). By inserting our color reduction method into the JPEG encoder, the image size will be further reduced and therefore the transmission time is reduced. This smart encoder enables its decoder using less time in decoding.

In addition to object's visual representation, this dissertation focuses on image quality evaluation for both color and grayscale images. The contributions to color vision science are as follows:

- A referenced IQM based upon image representation in very low-dimension; it is worth noting that the larger the image size is, the more efficient the low-dimension representation is.
- A no-reference image blurring metric;

- A no-reference color IQM based upon colorfulness, contrast and sharpness.

In summary, the contributions of this dissertation can be shown and realized by the following table lookup:

Table 8.1: A Check List for All the Proposed Algorithms in the Dissertation

	Apply a cutting-edge algorithm to an application	**Improve** a traditional algorithm in speed or size	**Create** a new algorithm
A Referenced Low-Dimensional Quality Index	**X**		
A No-Reference Metric for Image Blurring			**X**
A No-Reference Quality Index for Color Images			**X**
A No-Reference Color Correction Method			**X**
A Content-based Image Retrieval Method			**X**
An Image Specific Color Quantization Method	**X**		
A Modified Median Cut Method for Color Quantization		**X**	

8.2 Potential Applications

- Indexing and Clustering (Chapter 2)
- Information Retrieval (Chapter 2)
- Machine Learning (Chapter 2)
- Quality Filter for Image Search Engine (Chapter 2, 3 and 4)
- Front end of Automatic Image Enhancement (Chapter 4)
- Front end of JPEG Image Transcoder (Chapter 4)
- Color correction for low bit-rate JPEG and JPEG2000 images (Chapter 5)
- Content-based Image Retrieval (Chapter 6)
- Trademark Retrieval and Registration (Chapter 6)
- Wireless smart decoder for JPEG image from 24 bit to 8 bit (Chapter 7)
- Multimedia Image Streaming and Tele-Conferencing (Chapter 7)

8.3 Future Research Directions

Digital images span into numerous topics. Even as a narrowed range as color image quality evaluation, there are hundreds of metrics. One of the areas that this dissertation left out is the image quality issue on document images, i.e. Optical Character Recognition (OCR). Quantifying document image quality sometimes is important in applications, such as digital archive for image creation of historical paper documents. This will be one of my research directions in the future. The reason why document-imaging topic was not covered in this dissertation is that document images cannot be treated merely as grayscale or binary images. By intuition, we simply have to modify our color IQM for natural scene

images to fit the document images. The parameter of colorfulness is evidently of no use in document imaging. However, contrast and sharpness as variables are not enough for document images. There exist some distortions specific to the document images, for instance, speckles and while noises. After a brief literature review, we have found that the best way to measure the quality of document image is the OCR error rate. The rationale is that the less the OCR error rate, the higher the document image quality is. It will be inappropriate if we only delete the colorfulness variable in order to fit the gray-scale document images.

Yet another aspect of color vision is that artwork digitization, for instance, digitizing a collection of famous manmade artwork, which requires both truthful and device-specific. Meanwhile, with the increasing demand of artwork reproduction (transform artwork from one digital device to another), gamma correction (a nonlinear correction with respect to image values to fit with a particular display device, whose effect is to enhance the contrast) is no longer enough to perform color correction. Since this research area will be closely related to color quantization, color correction, and color reproduction, it will be one of my future research directions too.

There are two types of color reproduction. One is the constrained color reproduction, which is subjective to original artwork's authenticity constraint, and the other is non-constrained color reproduction, which is tailored to human perception. Nonlinear methods are not suitable for the former type of color reproduction. Chapter 5 has shown that the lookup table method is the best choice for the constraint color reproduction. I will design more algorithms based on the lookup table method for different application. As for the non-constrained color reproduction, it can be viewed as

color image enhancement; we think that nonlinear method can do the job. My future research focus is mainly on the constrained color reproduction.

In additional to the short-term focus of document images and artwork color reproduction, color profiling will be the task of long-term research direction. I plan to develop a fast instrumentation method to create color profile (Device color profiles provide color management systems with the information necessary to convert color data between native device color spaces and device independent color spaces.) for a new digital device. Since color profile allows exchanging device characteristics and color spaces in the form of metadata, it is better off to embed the device color profile in hardware or device driver level, and not in the image itself, which may increase the image size. I believe that the impact of the dissertation can be maximized if my long-term direction is color profiling.

APPENDIX A

THE ELEMENTARY PROOF OF JOHNSON-LINDENSTRAUSS LEMMA

The proof below is used in Chapter 2.

Johnson-Lindenstrauss lemma: Let X be an n-point set in a Euclidean space, and let $\varepsilon \in (0,1]$ be given. Then there exists (1+ε) embedding of X into R^k, where $k = O(\varepsilon^{-2} \log n)$.

Proof: Let $X \subseteq R^n$ (if X lies in higher dimensions, we can consider it to be lying in the span of its points, if it is in lower dimensions, we can add zero coordinates). Let $k = 200\varepsilon^{-2} \ln n$. Assume $k < n$, and Let L be a random k-dimensional linear subspace of R^n. Let $p : R^n \to L$ be the orthogonal projection operator. Let m be the number around which $\| p(x) \|$ is concentrated, for $x \in S^{(n-1)}$ as in Lemma 2.1.

Fix two points $x, y \in R^n$, we prove that

$$(1-\frac{\varepsilon}{3})m \| x-y \| \leq \| p(x)-p(y) \| \leq (1+\frac{\varepsilon}{3})m \| x-y \|$$

Holds with probability $\geq 1-n^{-2}$. Since there are $\binom{n}{2}$ pairs of points in X, it follows that with constant probability this holds for all pairs of points of X. In such case, the mapping p is D-embedding of X into R^k with $D \leq \dfrac{1+\frac{\varepsilon}{3}}{1-\frac{\varepsilon}{3}} \leq 1+\varepsilon, for \varepsilon \leq 1.$

Let $u = x - y$, we have $p(u) = p(x) - p(y)$ since $p(\bullet)$ is a linear operator. Thus, the condition becomes $\left(1-\frac{\varepsilon}{3}\right)m \| u \| \leq \| p(u) \| \leq \left(1+\frac{\varepsilon}{3}\right)m \| u \|$. Since this condition is scale independent, we can assume $\| u \| = 1$. Namely, we need to show that

$$\| \| p(u) \| - m \| \leq \frac{\varepsilon}{3} m.$$

Lemma 2.1 (exchanging the random space with the random vector) implies, for $t = \frac{\varepsilon}{3} m$, That this is bounded by

$$4\exp(-t^2 n/2) = 4\exp\left(\frac{-\varepsilon^2 m^2 n}{18}\right) \leq 4\exp\left(-\frac{\varepsilon^2 k}{72}\right) < n^{-2},$$

Since $m \geq \frac{1}{2}\sqrt{k/n}$.

APPENDIX B

SURVEY QUESTIONNAIRE ON COLOR IMAGE QUALITY

The objective of this questionnaire is to understand how human perceives the color image. The outcome of this color image quality survey will be used to construct the image quality measure in machine vision. There are altogether 30 images in this questionnaire.

Please rate each color image with a quality score from 1 to 7, according to the following meaning:

1 Very Poor (with a lot of artifacts, distortions and noises.)

2 Poor (with low contrast, object is hard to discern.)

3 Acceptable, A Little Below Average

4 Average

5 Satisfied, A Little Above Average

6 Good (with sharp edges, high contrast, maybe not so colorful.)

7 Excellent (with sharp edges, high contrast, and rich in colors.)


VINCENT VAN GOGH
The Drawings





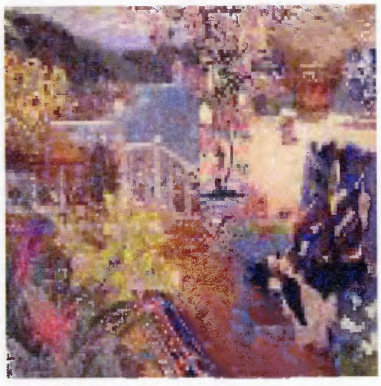








blur



REFERENCES

Abdel-Mottaleb, M., Krishnamachari, S., and Mankovich, N., "Performance evaluation of clustering algorithms for scalable image retrieval," *Proc. of IEEE Workshop on Empirical Evaluation Techniques in Computer Vision*, pp. 45-56, 1998.

Achlioptas, D., "Database friendly random projections," *Proc. of the Twentieth Symposium on Principles of Database Systems*, pp. 274-281, 2001.

Ahumada, A. J., "Computational image quality metrics: A review," *Society for Information Display Symposium Digest*, vol. 24, pp. 305–308, 1993.

Ahumada, A. J. and Null, C. H., "Image quality: A multidimensional problem," in A. B. Watson (ed.), *Digital Images and Human Vision*, pp. 141–148, MIT Press, 1993.

Albuz, E., Kocalar, E., and Khokhar, A., "Scalable color image indexing and retrieval using vector wavelets," *IEEE Trans. on Knowledge Data Engineering*, vol. 13, pp. 851–861, Sept./Oct. 2001.

Arriaga, R. I. and Vempala, S. "An algorithmic theory of learning: robust concepts and random projection," *Proc. of Foundations of Computer Science*, 1999.

Avcibas, I., "Statistical evaluation of image quality measure," *Journal of Electronic Imaging*, vol. 11, no. 2, pp. 206-212, 2002.

Balasubramanian, R., Allebach, J. P. and Bouman, C., "Sequential scalar quantization of vectors: an analysis," *IEEE Trans. on Image Processing*, vol. 4, no. 9, pp. 1282-1285, 1995.

Balasubramanian, R., de Queiroz, R., Fan, Z., "Multiresolution color correction," *SPIE Proceedings 3300, Color Imaging: Device Independent Color, Color Hardcopy and Graphics Arts III*, pp. 163-172, 1998.

Balci, K. and Atalay V., "PCA for gender estimation: which eigenvectors contribute?" *ICPR*, 2002, Quebec Canada, vol. 3, Aug. 2002.

Bast, H., "Dimensionality reduction: a powerful principle for automatically finding concepts in unstructured data," *International Workshop on Self-Properties in Complex Information System*, June 2004.

Beghdadi, A. and Pesquet-Popescu, B., "A new image distortion measure based on wavelet decomposition," *7th International Symposium on Signal Processing and its Application,* Paris, France, July, 2003.

Bi, J., Bennett, K., Embrechts, M., Breneman, C. and Song, M., "Dimensionality reduction via sparse support vector machines," *The Journal of Machine Learning Research,* vol. 3, pp. 1229-1243, March 2003.

Bingham, E. and Mannila, H. "Random projection in dimensionality reduction: applications to image and text data," *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 2001.

Bovik, A. C. and Liu, S., "Efficient DCT-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. Circuits System Video Tech*, vol. 12, no. 12, pp. 1139, 2002.

Braun, G. J., *A paradigm for color gamut mapping of pictorial images*, PHD. Thesis, Rochester Institute of Technology, Rochester, NY, 1999.

Callet, P. L. and Barba, D., "Robust approach for color image quality assessment," *Proc. SPIE - The International Society for Optical Engineering,* vol.5150 III, pp. 1573-1581, 2003.

Cao, L. J., Chua, K. S., Chong, W. K., Lee, H. P. and Gu, Q. M., "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing,* vol. 55, pp. 321-336, 2003.

Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M. and Malik, J., "Blobworld: A system for region-based image indexing and retrieval," *Third International Conference on Visual Information Systems*, 1999.

Caviedes, J. and Oberti, F., "A new sharpness metric based on local kurtosis, edge and energy information," *Signal Processing: Image Communication,* vol. 19, pp. 147-161, 2004.

Chaira, T. and Ray, A. K., "Fuzzy measures for color image retrieval," *Fuzzy Sets and Systems*, vol. 150, pp. 545-560, 2005.

Chang, C., Hwang, M. and Chen, T., "A new encryption algorithm for image cryptosystems," *The Journal of Systems and Software*, vol. 58, pp. 83-91, 2001.

Chang, C. and Yu, T., "Cryptanalysis of an encryption scheme for binary images," *Pattern Recognition Letters,* vol. 23, pp. 1847-1852, 2002.

Chang, E., "Support vector machine for high-dimensional indexing," *Technical Report, http://www-db.stanford.edu/~echang/svmdex.ps*, May 2001.

Chang, H. S., Sull, S. and Lee, S. U., "Efficient video indexing scheme for content-based retrieval," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 9, no. 8, Dec. pp. 1269-1279, 1999.

Cheng, S. and Yang, C., "A fast and novel technique for color quantization using reduction of color space dimensionality," *Pattern Recognition Letter,* vol. 22, pp. 845-856, 2001.

Chew, J. Y., Bouman, C. A. and Dalton, J. C., "Hierarchical browsing and search of large image databases," *IEEE Trans. on Image Processing*, vol. 9, no. 3, pp. 442-455, March 2000.

Chosson, S. and Hersch, R., "Techniques for color-gamut reduction," *SPIE's International Technical Group Newsletter, Electronic Imaging*, vol. 13, no. 2, pp. 5-8, June 2003.

Chou, C. H. and Li, Y. C., "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits System Video Technology,* vol. 5, no. 6, pp. 467-476, 1995.

Cowen, L. J. and Priebe, C. E., "Randomized non-linear projections uncover high-dimensional structure," *Advances in Applied Math.,* vol. 19, pp. 319-331, 1997.

Daly, S., "The visible differences predictor: an algorithm for the assessment of image Fidelity," *Proc. of SPIE*, vol. 1616, pp. 2-15, 1992.

Damera-Venkata, N., Kite, T. D., Geisler, W. S., Evans, B. L. and Bovik, A. C., "Image quality assessment based on a degradation model," *IEEE Trans. on Image Processing,* vol. 9, no. 4, pp. 636-650, April 2000.

Dasgupta, S. "Learning a mixture of Gaussians," *Proc. of Foundations of Computer Science*, pp. 634-644, 1999.

Dasgupta, S. and Gupta, A. "An elementary proof of the Johnson-Lindenstrauss Lemma," *Technical Report TR-99-06*, International Computer Science Institute, Berkeley, CA, 1999.

De Backer, S., Naud, A. and Scheunders, P., "Non-linear dimensionality reduction techniques for unsupervised feature extraction," *Pattern Recognition Letters,* vol. 19 , no. 8, pp. 711-720, 1998.

Dekker, A., "Kohonen neural networks for optimal color quantization," *Network: Computation in Neural Systems*, vol. 5, pp. 351-367, 1994.

Del Bimbo, A., *Visual information retrieval*, Morgan Kuffmann Publisher, San Francisco, California, 2001.

Derviaux, C., Coudoux, F., Gazalet, M., Corlay, P. and Gharbi, M., "A post-processing technique for block effect elimination using a perceptual distortion measure", *Proc. IEEE ICASSP*, vol.4, pp. 3001-3004, 1997.

DiGesu, V. and Staravoitov, V. V., "Distance-based functions for image comparison," *Pattern Recognition Letter,* vol. 20, no. 2, pp. 207–213, 1999.

Do, M. N. and Vetterli, M., "The finite ridgelet transform for image representation," *IEEE Trans. on Image Processing*, vol. 12, no. 1, pp. 16-28, 2003.

Donoho, D. L. and Duncan, M. R., "Digital curvelet transform strategy, implementation, and experiments," in H. H. Szu, M. Vetterli, W. Campbell, and J. R. Buss, editors, *Proc. Aerosense 2000, Wavelet Applications VII*, vol. 4056, pp. 12-29, Bellingham Washington, 2000, SPIE.

Donoho, D. L. and Huo, X., "Combined image representation using edgelets and wavelets," *SPIE*, vol. 3813, pp. 468-476, 1999.

Ebner, F. and Fairchild, M. D., "Gamut mapping from below: finding the minimum perceptual distances for colors outside the gamut volume," *Color Research and Application,* vol. 22, pp. 402-413, 1997.

Eckert, M. P. and Bradley A. P., "Perceptual quality metrics applied to still image compression," *Signal Processing*, vol. 70, pp. 177-200, 1998.

Edelman, B., Valentin, D. and Abdi H., "Sex classification of face areas: how well can a linear neural network predict human performance", *Journal of Biological Systems*, vol. 6, no. 3, pp.241-264, 1998.

Emmel, P. and Hersch, R. D., "Colour calibration for colour reproduction," *IEEE International Symposium on Circuits and Systems*, May 28-31, Geneva, Switzerland, pp. 105-108, 2000.

Eskicioglu, M. A. and Fisher, P. S. "Image quality measures and their performance," *IEEE Trans. Communication,* vol. 43, no.12, pp. 2959–2965, 1995.

Eskicioglu, M. A., "Quality measurement for monochrome compressed images in the past 25 years," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 1907-1910, Istanbul, Turkey, 2000.

Evgeniou, T., Pontill, M., Papageorgiou, C. and Poggio, T., "Image representations and feature selection for multimedia database search," *IEEE Trans. on Knowledge and Data Engineering,* vol. 15, no. 4, pp. 911-920, July/August 2003.

Faloutsos, C., Barber, R., Flickner, M. and Hafner, J. et al., "Efficient and effective querying by image content," *Journal of Intelligent Information Systems,* vol. 3, pp. 231-262, 1994.

Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley and Sons, Inc., 1957.

Finlayson, G., Hubel, P., and Hordley, S., "Color by correction," *Proc. of the Fifth IS&T/SID Color Imaging Conference*, pp. 6-9, 1997.

Fodor, I. A. "A survey of dimension reduction technique," *Center of Applied Scientific Computing, Lawrence Livermore National Laboratory*, June 2002.

Franti, P., "Blockwise distortion measure for statistical and structural errors in digital images," *Signal Processing: Image Communication,* vol. 13. no. 2, pp. 89–98, 1998.

Friedman, J. H., Bentley, J. L. and Finkel, R. A., "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. on Math Software*, vol. 3, no. 3, pp. 209-226, 1977.

Fuhrmann, D.R., Baro, J. A., and Cox, J. R., "Experimental evaluation of psychophysical distortion metrics for JPEG-coded images," *Journal of Electronic Imaging*, vol. 4, no. 4, pp. 397–406, 1995.

Gervautz, M. and Purgathofer, W., "A simple method for color quantization: Octree Quantization," *Graphic Germs*, Academic Press, New York, pp. 287-293, 1990.

Girod, B., "What's wrong with mean-squared error?" in *Digital images and human vision*, AB Watson, Ed. MIT: MIT Press, pp. 207-220, 1993.

Grgic, S., Grgic, M., and Mrak, M., "Reliability of objective picture quality measures," *Journal of Electrical Engineering*, vol. 55, no. 1-2, pp. 3-10, 2004.

Gutta, S., Wechsler, H., and Phillips, P., "Gender and ethnic classification," *Prof. of the IEEE International Automatic Face and Gesture Recognition*, pp. 194–199, 1998.

Guyon, I. and Stork, D., "Linear discriminant and support vector machine," *Advances in Large Margin Classifier*. The MIT Press, Cambridge, Massachusetts, pp. 147-169, 2000.

Hasler, D. and Susstrunk, S., "Measuring colorfulness in natural images," *Proc. SPIE Human Vision and Electronic Imaging*, 5007, pp. 87–95, Jan. 21–24, 2003.

Heckbert, P., *Color image quantization for frame buffer display*. B.S. thesis, Architecture Machine Group, MIT, Cambridge, Mass., 1980.

Hosaka, K., "A new picture quality evaluation method," *Proceedings of the International Picture Coding Symposium PCS '86,* Tokyo, Japan, pp. 17-18, April 1986.

Huang, J., Kumar, S. R., Mitra, M., Zhu, W. and Zabih, R., "Image indexing using color correlograms," *Proc. IEEE CVPR '97*, pp. 762–768, June 1997.

Hung, P., "Colorimetric calibration in electronic imaging devices using a look up table model and interpolations," *Journal of Electronic Imaging, vol. 2, no.1,* pp. 53-61, October 1993.

Huo, X., "Sparse image representation via combined transform," Dissertation of Stanford University, August 1999.

Hwang, K. and Chang, C., "A fast pixel mapping algorithm using principal component analysis," *Pattern Recognition Letters,* vol. 23, no. 14, pp. 1747-1753, Dec. 2002.

Indyk, P. and Motwani, R. "Approximate nearest neighbors: towards removing the curse of dimensionality," *Proc. of ACM STOC*, 1998.

Jacobson, R. E., "An evaluation of mage quality metrics," *Journal of Photographic Science,* vol. 43, no. 1, pp. 7–16, 1995.

Janssen, R. *Computational Image Quality*. SPIE press, 2001.

Janssen, T., "Understanding image quality," *Pro. 2001 International Conference on Image Processing*, vol. 2, pp. 7-10 Oct. 2001.

Johnson, W. B. and Lindenstrauss, "Extensions of Lipshitz mapping into Hilbert space," *Contemporary Mathematics,* vol. 26, pp.189-206, 1984.

Juffs, P., Beggs, E. and Deravi, F., ''A multiresolution distance measure for images," *IEEE Signal Processing Letter,* vol. 5, no. 6, pp.138–140, 1998.

Kanamori, K. and Kotera, K., "Color correction technique for hard copies by 4-neighbours interpolation method," *Journal of Imaging Science and Technology,* vol. 36, no. 1, pp. 73-80, Jan. 1992.

Kanjanawanishkul, K. and Uyyanonvara, B., "Fast adaptive algorithm for time-critical color quantization," *Proc. VII Digital Image Computing: Techniques And Applications*, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), pp. 10-12, Dec. 2003, Sydney.

Karunasekera, S. A. and Kingsbury, N. G., "A distortion measure for blocking artifacts in images based on human visual sensitivity," *IEEE Transactions on Image Processing,* vol. 4, no. 6, pp. 713-724, June 1995.

Kaski, S. "Dimensionality reduction by random mapping: fast similarity computation for clustering," *Proc. IEEE Inte'l Joint Conference on Neural Network*, vol. 1, pp. 413-418, 1998.

Kayargadde, V. and Martens, J., "Perceptual characterization of images degraded by blur and noise", *Journal of Optical Society American*, vol. 13, no. 6, June 1996.

Kerre, E. and Nachtegael, M., *Fuzzy Techniques in Image Processing*, Series Studies in Fuzziness and Soft Computing, vol. 52, Springer-Verlag, July 2000.

Kolpatzik, B. and Bouman, C., "Optimized universal color palette design for error diffusion," *Journal of Electronic Imaging,* vol. 4, pp. 131-143, 1995.

Kong, H., Vetro, A. and Sun, H., "Edge map guided adaptive post-filter for blocking and ringing artifact removal," *IEEE International Symposium on Circuit and Systems,* vol. 3, pp. 929-932, May 2004.

Koren,Y. and Carmel, L., "Robust linear dimensionality reduction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 4, pp. 459-470, 2004.

Krishnamachari, S. and Abdel-Mottaleb, M., "Hierarchical clustering algorithm for fast image retrieval", *IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases VII*, 1999.

Lazzaro, J. P. and Wawrzynek, J., "JPEG quality transcoding using neural network trained with a perceptual error measure," *Neural Computation,* vol. 11, no. 1, pp. 267-296, Jan. 1999.

Lew, M., "Next generation web searches for visual content," *IEEE Trans. on Computer*, vol. 33, pp. 46-53, Nov. 2000.

Li, F., Song, J., and Zhan, Y., "A fast color correction method based on image analysis," *Proc. of 3rd International Conference on Image and Graphics*, pp. 108-111, 2004.

Li, J., Chen, G. and Chi, Z., "A fuzzy image metric with application to fractal coding," *IEEE Trans. on Image Processing,* vol. 11, no. 6, pp. 636-643, 2002.

Li, J. and Kuo, C. J., "Coding artifact removal with multi-scale post-processing, in *IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997.

Li, X., Yuan, T., Yu, N. and Yuan, Y., "Adaptive color quantization based on perceptive edge protection," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3165-3176, Dec. 2003.

Linde, Y., Buzo, A. and Gray, R., "An algorithm for vector quantizer design," *IEEE Trans. on Communication*, vol. 28, no. 1, pp. 84-95, 1980.

Liu, C. M., Lin, J. Y., Wu, K. G. and Wang, C. N., "Objective image quality measure for block-based DCT coding," *IEEE Trans. on Consumer Electronics,* Aug., vol. 43, no. 3, pp. 511-516, 1997.

Lui, T. Y. and Izquierdo, E., "Scalable object-based image retrieval," *IEEE International Conference on Image Processing, ICIP 2003*, vol. 3, pp. 501-504, Sep. 2003.

MacLennan-Brown, K. and Jacobson, R. E., "Visual significance of digital artifacts," *Society for Imaging Science and Technology: Image Processing, Image Quality, Image Capture, Systems Conference*, pp. 119-123, 2001.

Magen, A., "Dimensionality reductions that preserve volumes and distance to affine spaces, and their algorithmic applications," *Proc. of the 6th International Workshop on Randomization and Approximation Techniques*, 2002.

Manian, V., Diaz, M. and Vásquez, R., "Wavelet features for color image classification," *Imaging and Geospatial Information Society*, Annual Conference, 2000.

Manjunath, B. S. and Ma, W. Y., "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 837–842, 1996.

Martens, J. B. and Meesters, L., "Image dissimilarity," *Signal Processing*, vol. 70, pp. 155-176, 1998.

Marziliano, P., Dufaux, F., Winkler, S. and Ebrahimi, T., "A no-reference perceptual blur metric," IEEE *Proc. International Conference on Image Processing*, pp. 57-60, Rochester, NY, Sept. 22-25, 2002.

Marziliano, P., Dufaux, F., Winkler, S. and Ebrahimi, T., "Perceptual blur and ringing metrics: application to JPEG2000," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 163-172, February 2004.

Matkovic, K., Neumann, L., and Purgathofer, W., "A survey of tone mapping techniques," *13th Spring Conference on Computer Graphics*, pp.163–170, 1997.

Matousek, J., "On the distortion required for embedding finite metric spaces into normed spaces," *Israel Journal of Mathematics*, vol. 93, pp 333-344, 1996.

Mayache, A. and Eude, T., "An evaluation of quality metrics for compressed images based on human visual sensitivity", *Fourth International Conference on Signal Processing (ICSP'98)*, pp.779-782, Beijing, Chine, 12-16 Oct. 1998

Moghaddam B. and Yang, M., "Gender classification with support vector machine," *Proc. of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2000)*, oral presentation, pp. 306-311, Grenoble, France, March 2000.

Morovic, J., *To develop a universal gamut mapping algorithm*, PHD. Thesis, University of Derby, Derby, UK, 1998.

Morovic, J. and Luo, M. R., "Gamut mapping algorithms based on psychophysical experiment," *Proc. of the 5th IS&T/SID Color Imaging Conference*, pp. 44-49, 1997.

Morovic, J. and Luo, M. R., "Cross-media psychophysical evaluation of gamut mapping algorithms," *Proc. AIC Color 97 Kyoto*, vol. 2, pp. 594-597, 1997.

Muller, H., Muller, W., Squire, D. and Pun, T., "Performance evaluation in content-based image retrieval: overview and proposals," *Technical report* in Computer Vision Group, University of Geneva, pp. 1-11, Dec. 1999.

Murching, A. and Woods, J. W., "Adaptive subsampling of color images," *Proc. of International Conference on Image Processing, ICIP 1994*, vol. 3, pp. 963-966, 1994.

Neelamani, R., de Queiroz, R., Fan, Z., and Baraniuk, Z., "JPEG compression history estimation for color images," *2003 International Conference on Image Processing*, vol. 3, pp. III- 245-248, 2003.

Nill, N. B., ''A visual model weighted cosine transform for image compression and quality assessment,'' *IEEE Trans. Communication,* vol. 33, no. 6., pp. 551–557, 1985.

Nill, N. B. and Bouzas, B. H., "Objective image quality measures derived from digital image power spectra," *Optical Engineering*, vol. 31, no. 4, pp. 813-825, 1992.

Ohtsubo, J. and Fujimoto, A., "Practical image encryption and decryption by phase-coding technique for optical security systems," *Applied Optics*, vol. 41, no. 23, pp. 4848-4855, Aug. 2002.

Omer, I. and Werman, M., "Color line: Image specific color representation," *CVPR*, 2004.

Osberger, W., Bergermann, N., and Maeder, A., "Automatic image quality assessment technique incorporating higher level perceptual factors," *IEEE International Conference on Image Processing*, vol. 3, pp. 414-418, 1998.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., and Huff, D., "The perception of face gender: the role of stimulus structure in recognition and classification," *Memory and Cognition*, vol. 26, no. 1, pp. 146-160, 1997.

Ozdemir, D. and Akarun, L., "A fuzzy algorithm for color quantization of images," *Pattern Recognition*, vol. 35, pp. 1785-1792, 2002.

Papamarkos, N., Atsalakis, A. E. and Strouthopoulos, C. P., "Adaptive color reduction," *IEEE Trans on Systems, Man, and Cybernetics-Part B*, vol. 32, no. 1, pp. 44-56, Feb. 2002.

Raymer, M. L. "Dimensionality reduction using genetic algorithms," *IEEE Transaction on Evolutionary Computation,* vol. 4, no. 2, pp. 164-171, July 2000.

Reed, T. R., and Hans, J., "A review of recent texture segmentation and feature extraction techniques," *CVGIP: Image Understanding*, vol. 57, no. 3, pp. 359-372, May 1993.

Remias, E., Sheikholeslami, G., Zhang, A. and Syeda-Mahmood, T. F., "Supporting content-based retrieval in large image database systems," *The International Journal on Multimedia Tools and Applications,* vol. 4, no. 2, pp. 153-170, March 1997.

Roweis, S. and Saul, L., "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, Dec.22, pp.2323-2326, 2000.

Rui, Y., and Huang, T. S., "Image retrieval: current techniques, promising directions, and open issues," *Journal of Visual Communication and Image Representation*, vol. 10, pp.39-62, 1999.

Santini, S. and Jain, R., "Similarity measures," *IEEE Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 871-883, 1999.

Shen, D. F. and Wang, S. C., "Measurements of JND property of HVS and its applications to image segmentation, coding, and requantization," *Proc. of SPIE - The International Society for Optical Engineering*, vol. 2952, pp. 113-121, 1996.

Shen, M. and Kuo, C., "Review of post-processing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation,* vol. 9. no. 1, pp. 2-14, March, 1998.

Smeulders, A., Worring, M., Santini, S., Gupta, A. and Jain, R., "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, vol. 22, pp. 1349–1380, Dec. 2000.

Smith, J. R., "Color for image retrieval," in *Image database: search and retrieval of digital imagery*, edited by V. Castelli and L. D. Bergman, pp. 285-309, John Wiley & Sons, New York, 2002.

Smith, J. R. and Chang, S., "Tools and techniques for color image retrieval," *SPIE Proceedings*, vol. 2670, pp. 1-12, 1996.

Spierenburg, J. A. "Dimension reduction of images using neural networks," *Master*

Thesis, Leiden University, 1997.

Squire, D. M., Müller, H. and Müller, W., "Improving response time by search pruning in a content-based image retrieval system, using inverted file techniques," *Proc. of IEEE workshop on CBAIVL*, June 1999.

Srihari, R. K., Zhang, Z. and Rao, A., "Intelligent indexing and semantic retrieval of multimodal documents," *Information Retrieval*, vol. 2, no. 2-3, pp. 245–275, 2000.

Stinson, D. *Cryptography: Theory and Practice*, CRC Press, 1995.

Sun, P. L. and Morrovic, J., "What difference do observers see in colour image reproduction experiments," *CGIV 2002,* Poitiers, France, 2002.

Süsstrunk, S. and Winkler, S., "Color image quality on the Internet," *Proc. SPIE/IS&T Internet Imaging*, vol. 5304, pp. 118-131, San Jose, CA, January 18-22, 2004.

Tang, E. K., Suganthan, P. N., Yao, X. and Qin, A. K., "Linear dimensionality reduction using relevance weighted LDS," *Pattern Recognition*, vol. 38, pp. 485-493, 2005.

Tang, J., Peli, E., and Acton, S., "Image enhancement using a contrast measure in the compressed domain," *IEEE Signal Processing Letters*, vol. 10, no. 10, pp. 289-292, October 2003.

Tasdizen, T., Akarun, L. and Ersoy, C., "Color quantization with genetic algorithm," *Signal Processing: Image Communication*, vol. 12, pp. 49-57, 1998.

Tenenbaum, J. B., Vin de Silva, and Langford, J. C., "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319-2323, Dec. 2000.

Tieu, K. and Viola, P., "Boosting image retrieval", *IEEE Conf Computer Vision and Pattern Recognition (CVPR'OO)*, vol. 1, pp. 228-235, 2000.

Tong, H, Li, M., Zhang, H. and Zhang, C., "No-reference quality assessment for JPEG2000 compressed images," *IEEE International Conference on Image Processing*, pp. 3539-3542, 2004.

Tremeau, A., Calonnier, M. and Laget, B., "Color quantization error in terms of perceived image quality," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April, vol. 5, pp. 93-96, 1994.

Turaga, D. S., Chen, Y. and Caviedes, J., "No-reference PSNR estimate for compressed pictures," *International Conference on Image Processing,* vol. 3, pp. III 61-64, 2002.

Vajda, S., *Fibonacci and Lucas numbers, and the golden section: theory and application,* New York: Halsted Press, John Wiley & Sons, 1989.

Van Droogenbroeck, M. and Benedett, R., "Techniques for a selective encryption of uncompressed and compressed images," *Proc. of ACIVS (Advanced Concepts for Intelligent Vision Systems)* Ghent, Belgium, pp. 90-97, Sep. 2002.

Vapnik, V., *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.

Vempala, S. "A random sampling based algorithm for learning the intersection of half-spaces," *Proc. of Foundations of Computer Science*, pp. 508-513, 1997.

Vin de Silva and Tenenbaum, J. B., "Global versus local methods for nonlinear dimensionality reduction," pp. 705–712 in S. Becker, S. Thrun,, K. Obermayer (Eds.), *Advances in Neural Information Processing Systems*, 15, MIT Press, Cambridge, MA, 2003.

Wang, C., Yang, Y., Li, W., and Chen, S., "Image texture representation and retrieval based on power spectral histograms," *16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'04)*, pp. 491-495, 2004.

Wang, Z. and Bovik, A. C. "A universal image quality index," *IEEE Signal Processing Letters,* vol. 9, no. 3, pp. 81-84, March 2002.

Wang, Z., Bovik, A.C. and Evans, B. L., "Blind measurement of blocking artifacts in Images," *Proc. ICIP*, vol. 3, pp.981–984, Vancouver, Canada, 2000.

Wang, Z., Bovik, A. C. and Lu, L., "Why is image quality assessment so difficult?" *IEEE Int'l Conference on Acoustics, Speech, & Signal Processing*, May 2002.

Wang, Z. and Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing,* vol. 13, no. 4, April 2004.

Wang, Z., Sheikh, H. R., and Bovik, A. C., "No-reference perceptual quality assessment of JPEG compressed images," *IEEE International Conference on Image Processing*, Sept. 2002.

Watson, A. B., "Perceptual optimization of DCT color quantization matrices," *IEEE International Conference on Image Processing*, pp. 100–104, Austin, Texas, November 1994.

Winkler, S., "A perceptual distortion metric for digital color images," *Proc. International Conference on Image Processing*, 1998.

Winkler, S., "Issues in vision modeling for perceptual video quality assessment," *Signal Processing,* vol. 78, no. 2, pp. 231-252, October 1999.

Winkler, S., "Visual fidelity and perceived quality: toward comprehensive metrics," *Proc. SPIE Human Vision and Electronic Imaging*, vol. 4299, pp. 114-125, San Jose, California, January 21-26, 2001.

Winkler, S., Sharma, A. and McNally D., "Perceptual video quality and blockiness metrics for multimedia streaming applications," *Proc. of the International Symposium on Wireless Personal Multimedia Communications*, Aalborg, Denmark, pp. 547-552, 2001.

Wu, H. R. and Yuen, M. "A generalized block-edge impairment metric for video coding." *IEEE Signal Processing Letters,* vol. 4, no. 11, pp. 317–320, 1997.

Wu, P., Manjunath, B. S., and Shin, H. D., "Dimensionality reduction for image retrieval," *Proc. of IEEE International Conference on Image Processing (ICIP 2000)*, vol. 3, Vancouver, Canada, pp. 726–729, 2000.

Wu, X., "Color quantization by dynamic programming and principal analysis," *ACM Transactions on Graph, vol. 11, no. 4,* pp. 348-372, 1992.

Yang, S., Hu, Y., Nguyen, T. and Tull, D., "Maximum-likelihood parameter estimation for image ringing-artifact removal," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 11, no. 8, pp. 963-973, August 2001.

Yin, J., "Color correction methods with applications to digital projection environments," *WSCG'2004*, pp. 102-120, 2004.

Yoshitaka, A. and Ichikawa, T., "A survey on content-based retrieval for multimedia database," *IEEE Trans. on Knowledge and Data Engineering,* vol. 11, no. 1, pp. 81-93, 1999.

Youla, D. C., "Generalized image restoration by the method of alternating orthogonal projections," *IEEE Trans. Circuits Systems*, vol. 25, pp. 694–702, 1978.

Yu, Z., Wu, H. R., Winkler S., and Chen, T., "Vision model based impairment metric to evaluate blocking artifacts in digital video." *Proc. of IEEE*, vol. 90, no. 1, pp. 154-169, Jan. 2002.

Zhang, X., Lin, W., and Xue, P., "Improved estimation for just-noticeable visual distortion," *Signal Processing*, vol. 85, no. 4, pp. 795-808, April 2005.

Zhang, X. and Wandell, B., "A spatial extension of CELAB for digital color image reproduction," *Journal of the Society for Information Display,* vol.5, no.1, pp.61-63, 1997.

Zhang, M., Xie, J., Li, Y., and Wu, D., "Color histogram correction for panoramic images," *Seventh International Conference on Virtual Systems and Multimedia (VSMM'01)*, p. 328, 2001.

Zhao, M., Hofman, P. M., and de Haan, D., "Content-adaptive up-scaling of chrominance using classification of luminance and chrominance data," *Conference* of *Visual Communications and Image Processing,* Jan. 2004, *Proc. of SPIE*, vol. 5308, 2004.

Zhou, B., Shen, J. and Peng, Q., "An adjustable algorithm for color quantization," *Pattern Recognition Letters,* vol. 25, no. 16, pp. 1787-1797, Dec. 2004.

Zhu, X. and Huang, T. S., "Unifying keywords and visual contents in image retrieval," *IEEE Multimedia*, no. 2, pp. 23–33, April and June 2002.