# A Study of Color Quantization and its Quality Assessment

A thesis submitted to the University of Hyderabad in partial fulfillment of the requirements for the award of

**Doctor of Philosophy**

in

**Computer Science**

by

**Mohammed Ahmed Hassan**


**School of Computer & Information Sciences**

**University of Hyderabad**

**Hyderabad – 500 046, India**

**July, 2013**

# CERTIFICATE

This is to certify that the thesis entitled **A Study of Color Quantization and its Quality Assessment** submitted by **Mohammed Ahmed Hassan** bearing Reg. No. **08MCPC18** in partial fulfillment of the requirements for the award of **Doctor of Philosophy** in **Computer Science** is a bonafide work carried out by him under our supervision and guidance.

The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.


Chakravarthy Bhagvati

Professor

Deparment of CIS

University of Hyderabad

Hyderabad – 500046

Prof. Arun K. Pujari

Dean

School of CIS

University of Hyderabad

Hyderabad – 500046

# DECLARATION

I, **Mohammed Ahmed Hassan**, hereby declare that this thesis entitled **A Study of Color Quantization and its Quality Assessment** submitted by me under the guidance and supervision of **Prof. Chakravarthy Bhagvati** is a bonafide research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Date:

Name: Mohammed Ahmed Hassan

Reg. No: 08MCPC18

Signature of the student

# Acknowledgments

First of all I would like to thank Allah, the most Gracious and ever Merciful, for giving me this great opportunity to pursue Ph.D. at this university.

I wish to express my sincere gratitude to my supervisor, Prof. Chakravarthy Bhagvati, for his support and guidance during the completion of this dissertation. Most of my basic knowledge and skills in the academic research were learned from him. I could not have wished for a better advisor for my Ph.D. study.

I would like to thank all my friends at University of Hyderabad, especially Pavan Kumar, for their support.

Last, but not least, I would like to thank my wife for her patience, support, caring, continual encouragement, compassion and love. For enriching my life and being a constant source of motivation and happiness.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1

# Introduction

## 1.1 What is Color Quantization?

The process of reducing the number of colors presented in a digital full color image is called color image quantization or simply color quantization. Consider a true color image $I$ with $n$ colors which can be defined as [1]:

$$f : N \times N \rightarrow C \subseteq \Omega \tag{1.1}$$

where $\Omega = \{(r, g, b) \,|0 \leq r, g, b \leq 255\}$ is the RGB color space, $(x, y) \in N \times N$ are the spacial coordinates of a pixel, $N$ is the integer set, and $C = \{c_1, c_2, ..., c_n\}$ is the set of colors used in the image $I$.

The color quantization of the image $I$ into $k$ colors, with $k < n$ (usually $k << n$), can be defined as a mapping :

$$q : N \times N \rightarrow R \subseteq \Omega \tag{1.2}$$

where $R = \{r_1, r_2, ..., r_k\}$ is the set of representative colors used in the quantized image. This mapping $q$ should minimize the perceived difference between the original and the quantized images.

From the above definition, we conclude that the color quantization process

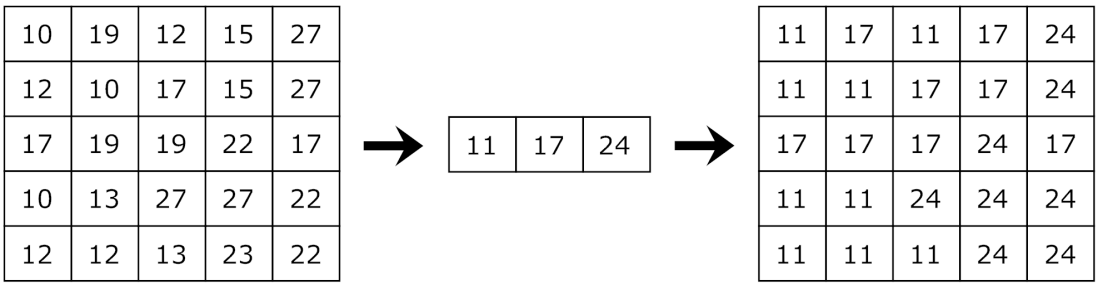



Figure 1.1: Color quantization example. The original image in the left, the color palette in the middle, and the color quantized image in the right.

consists of two main steps: the first step called color palette design in which a small set of representative colors or palette (typically 8-256 colors) is chosen from the $2^{24}$ possible colors. The second step called pixel mapping in which each pixel in the full color image is mapped to the closest color in the color palette. An example of color quantization is shown in Fig. 1.1, a color palette with three colors is first generated from the original image using quantization algorithm, then the pixels in the original image are replaced by their nearest color in the palette to form the quantized image.

## 1.2 Why Color Quantization?

Originally, color quantization has been used to satisfy the hardware constraints of the display devices that allow only a limited number of colors to be displayed simultaneously [2]. Today, the original motivation of color quantization has changed due to the availability of inexpensive full color displays. However, color quantization is still an important problem in the fields of image processing and computer graphics [3, 4].

The memory requirement for 24 bits full color image of size $512 \times 512$ is about

768 Kbytes that may seem trivial for the high-end computer system, but it is even today significantly large for mobiles and hand-held devices where memory is usually small. In addition, these devices typically have insufficient capability to display a full color images due to the limited amount of memory devoted for that purpose. Therefore, color quantization can be used to reduce the memory requirements in mobiles and hand-held devices and to satisfy their limited color display. Instead of each pixel in 24 bits full color image contains its own red, green and blue values; each pixel holds an 8 bit value, which is an index number into a user defined table of color palette. Each entry into the table contains a 24 bit value which specifies the color's red, green and blue components. In this way, the size will be reduced to about one third of the original size.

Since color quantization is considered as a form of data compression, multimedia applications such as the transmission and viewing of color images and video streams over the World Wide Web is another area that can benefit from color quantization. The number of colors used in the individual video frames can be reduced while maintaining a good quality image. By reducing the number of colors in each frame, a higher-level of compression is possible resulting in faster transmission times.

Color quantization also has been used as a preprocessing step for many applications such as object recognition, image compression, and content-based image retrieval (CBIR). In object recognition [5], prior to the process of object recognition the input image is quantized according to the authors' claim that the entire range of colors need only be quantized into about 200 different discrete colors to distinguish a large number of objects. Compressing an image starts with quantizing the 24 bits input image into 8 bits one. The quantized image is then divided

into $n \times n$ non-overlapping square blocks. Each block is then coded separately [6]. In content-based image retrieval, images are retrieved based on the similarity of some features. Color histogram of an image is rotation, translation, and scale-invariant; therefore, it is widely used color feature in CBIR [7]. Color histogram is derived by first quantizing colors in the image into a number of bins in a specific color space, and then counting the number of image pixels in each bin.

Another aspect of importance of color quantization is that the human visual system cannot perceive a large number of colors at the same time, nor is it able to distinguish close colors well [8] even though under appropriate adaptation. A full color image may contain up to 16 million different colors; this large number of colors makes it difficult to handle a variety of color-based tasks such as edge detection, enhancement, computing histograms, and color adjustment. These tasks are more efficiently carried out on a small set of colors. Further, the ability of the human visual system to resolve multiple shades of color is also dependent on the spatial resolution of the images. In mobiles and other hand-held devices with their small-sized displays, it may not really be necessary for $2^{24}$ shades of colors

## 1.3 Contributions

In this dissertation, we address three problems related to color image quantization. The first problem is the creation of a ground truth database for human perception of the quality of color quantized images, the second is the objective quality assessment of color quantized images, and the third is the development of new algorithms that combine color quantization and dithering techniques. The major contributions of our work are summarized below:

- Conducted extensive psychometric experiments to collect subjective quality scores for a diverse set of color quantized images. The subjective mean opinion scores (MOS) from the subjective study were used for training and testing the algorithms presented in this dissertation and also for validating the correlation of the existing image quality metrics with human perception in evaluating the quality of color quantized images. There were 25 reference images, each of them was quantized into seven quantization levels using five popular quantization algorithms. The perceived quality of each of the quantized images was evaluated by 22 human subjects and more than 19200 individual human quality judgments were carried out to obtain the final mean opinion scores.

- A comparative analysis of state-of-the-art image quality assessment metrics against the subjective mean opinion scores (MOS) from the subjective study is presented and their correlation with the human opinion scores in estimating the perceived quality of color quantized images is explored.

- Proposed an extension to the well-known multi-scale structural similarity index (MSSIM). This extension adds a color comparison to the criteria of the grayscale MSSIM in order to improve its ability of measuring the quality of color quantized images. The performance analysis of the proposed metric against the subjective mean opinion scores (MOS) from the subjective study shows that the new metric outperforms state-of-the-art image quality metrics in the quality assessment of color quantization distortion.

- The performance of the extended MSSIM has been analyzed also on the Tampere Image Database (TID2008), which is largest database so far available with 17 types of distortions, to further test the extended metric for differ-

ent distortion types. The results show that the proposed extension greatly improves the performance of the original MSSIM for these other distortion types also.

- Proposed a novel objective image quality metric for quality assessment of color quantized images. This proposed metric models any color quantization distortion as a combination of three similarities: color similarity, edge similarity, and structure similarity. The performance validation of the proposed metric against the subjective mean opinion scores (MOS) from the subjective study shows that the new metric remarkably outperforms state-of-the-art image quality metrics in the quality assessment of color quantization distortion.

- Proposed a color quantization method for use with dithering techniques. The palette colors are selected in a way that they are distributed around few initially selected palette colors based on the Just Noticeable Color Difference (JNCD) threshold in the CIELAB color space. The results show that our method noticeably removes false contours as well as prevents color impulses better than existing color quantization methods when applying dithering techniques to color quantized images.

- Proposed color quantization method for use with dithering techniques based on local fractal dimension. The local fractal dimension is used to assign different weights to image pixels according to their spatial distribution. Simulation results show that the proposed method gives a remarkable improvement in the quality of joint color quantization and dithering both subjectively and objectively.

This dissertation is divided into six chapters. The first Chapter is the Introduction where we describe the problem of color image quantization and present the need for color image quantization. We also enumerate the contributions of the dissertation and show the organization of the dissertation in this chapter. Chapter 2 reviews the literature that is relevant to the topics in this dissertation. The first part of this review explores the color quantization methods. In the second part, a review of the objective image quality assessment metrics is given. Since the dissertation focuses only on full-reference objective metrics, existing full-reference objective metrics are roughly classified and reviewed. Chapter 3 presents an investigation of *the subjective quality assessment of color quantized images* in which we have conducted extensive psychometric experiments to collect subjective quality scores for a diverse set of color quantized images produced by five popular quantization algorithms. The subjective mean opinion scores (MOS) from the subjective study were used for training and testing the algorithms presented in this dissertation and also for validating the correlation of the existing image quality metrics with human perception in evaluating the quality of color quantized images. Chapter 4 addresses the problem of *objective quality assessment of color quantized images*, where we propose two image quality assessment methods for the quality assessment of color quantized images. In Chapter 5, we present the problem of *joint color quantization and dithering techniques* in which we have proposed two color quantization methods that are designed to obtain better results with minimum visual degradation after dithering technique is applied. Finally, conclusions are drawn and future work is suggested in Chapter 6.

# CHAPTER 2

# Review of the Literature

This chapter presents a detailed review of the literature that is relevant to the topics in this dissertation. The first part of this review will explore the color quantization methods. In the second part, a review of the objective image quality assessment metrics is given. Since the dissertation focuses only on full-reference objective metrics, existing full-reference objective metrics are roughly classified and reviewed.

## 2.1 Color Image Quantization

In the literature, many color quantization methods have been discussed. The previous works can be grouped into two main classes: hierarchical and non-hierarchical methods. Hierarchical methods construct a tree depicting the relationship among the colors in the image color space. Hierarchical techniques that construct a tree in a bottom up manner are called *agglomerative*, while hierarchical methods that construct a tree in top down manner are called *divisive*. Non-hierarchical techniques include methods that use iteration, self-organizing neural networks, genetic programming, or other approaches that refine previously selected representative colors.

### 2.1.1 Hierarchical Color Image Quantization

One of the most widely used color quantization algorithms is the Median-cut algorithm developed by Heckbert [2]. The premise of this algorithm is to partition the image color space such that each partition represents approximately the same number of image pixels. To reduce the storage requirements and the computation time, the number of distinct colors is reduced prior to the partitioning step. Heckbert suggests that 5 bits per color channel is sufficient. This reduction is known as bit reduction and is performed by converting the three low order bits of each color component to zero. The algorithm starts with one box that encloses colors of all pixels in the input image. The box is partitioned into two sub-boxes. The cutting plane is chosen perpendicular to the longest side and passes through the median point. The operation is recursively applied to the new sub-boxes and the sub-box with the largest number of colors is chosen for further partitioning. This process continues until the number of boxes is equal to the number of desired palette colors. The palette colors are then generated by averaging the colors in each box.

Joy and Xiang [9] proposed the center-cut algorithm which is largely based on Heckbert's median cut algorithm [2] with three modifications. It uses 3-2-4 bit reduction scheme where emphasis is given to the green component. This bit reduction is performed by converting the three low order bits of red component, the two low order bits of green component, and the four low order bits of blue component to zero. The center-cut algorithm divides the hyper-box along its longest dimension and the cutting plane passes through the center point rather than the median point. The box with the "longest-longest dimension" is chosen for further partitioning instead of selecting the box with the most pixels.

A greedy variance-minimization technique has been described by Heckbert in his B.S. thesis [10] as a variation on the median-cut algorithm. There is only one change: instead of splitting at the median point, he splits spatially. His suggestion considers all possible split planes perpendicular to the dominant dimension, and for each finds the variance of colors in the left and right halves of the box. The split point is chosen in order to minimize that sum. Heckbert did not implement his idea of greedy variance-minimization. However several approximation algorithms on the theme have been investigated including Wan et al. [11, 12], Wu [13, 14], Liu and Chang [15], and Braquelaire and Brun [4].

In octree quantization [16] the input image color space is divided into eight sub-boxes. Each sub-box is then divided into eight new sub-boxes. The subdivision continues until the individual colors are reached. This process of subdivision is represented by octree where the entire color space is treated as a hierarchy of octants and each color in the input image is placed into the leaf of the constructed tree. The pruning process starts with the longest branch. Neighboring leaves in the tree are recursively replaced with their parent node containing an average of the color in the child leaves. The process stops when the number of leaves is equal to the desired number of palette colors. The colors in the leaves are then considered as the representative colors. As the merging process of the leaves grows by a factor of eight, the size of the color palette for octree quantization is always a multiple of eight.

Sirisathitkul et al. [17] proposed color quantization algorithm based on a concept that the sum of distances between adjacent colors along the highest color variance axis indicates the dissimilarity of all colors in each color cell. And so, dividing a cell into two smaller cells by balancing the sum of distances between

adjacent colors and frequencies of each cell could mean that the colors in each cell are similar and gradually change. Consequently, the cell representative is close to every color in each cell and hence the total quantization error becomes minimal. The algorithm starts by creating a 3D histogram of the input full color image to serve as the color space for selecting representative colors. After sorting colors along the principal axis, a dimensional axis with the highest variance of color component distributions, the Euclidean distances between adjacent colors along the principal axis are computed. The cutting plane is selected perpendicular to the principal axis and passes through a color which approximately equalizes the quantization errors of the two sub-cells. The splitting procedure is repeated until the desired number of cells is obtained. Finally, the centroids of all cells are collected to form the representative colors in the color palette.

The partitioning process for all the algorithms reviewed above is performed by orthogonal planes even though in general color image data sets are not distributed orthogonally. To address this, some algorithms in the literature consider the principal axis of the color distribution to split the partition orthogonal to the direction of greatest distortion which may not agree with one of the R, G, or B axis.

Orchard and Bouman proposed the binary splitting algorithm [18]. The binary splitting algorithm partitions the respective cluster of colors along the direction of the maximum color variation around the cluster centroid by a hyperplane which passes through the centroid. Such optimal direction is determined by the principal eigenvector corresponding to the largest eigenvalue. The cluster with the largest eigenvalue is chosen for further partitioning. Balasubramanian et al. [19] used efficient data structures, histogramming, and prequantization to speed up the Orchard-Bouman's binary splitting algorithm.

Yang and Lin [20] proposed a color quantization method that uses the radius weighted mean cut (RWM-cut). This method is a hierarchically divisive method similar to median cut except that the decision boundary used to split the given set into two smaller subsets is taken to be the plane passing through the radius weighted mean (RWM) color vector and perpendicular to the line between the RWM color vector and the centroid of the set. The set with the largest variation is chosen for further partition.

A fast dimensionality reduction technique has been used by Cheng and Yang [1] in their color quantization method. The resolution of each color component is reduced to 5 bits to create a 32×32×32 3D histogram from a 24 bit/pixel color image. This histogram, which gives the relative frequency of occurrences of colors in an image, is then used as the color space for extracting representative colors. The partitioning axis of each class is the line defined by the mean color vector and the color of the largest distance away from the mean color. The vector composed of the projection values, which are projected onto the selected line in a class, is then used to split the color vectors of the class into two principle colors. Partitioning of the color vectors is repeated until the desired number of representative colors is achieved. After the partitioning process is completed, all color vectors in each class are regarded as identical and are represented by the color value computed by averaging the color vector.

Although hierarchical color quantization methods are fast in execution, they suffer from a common drawback that they are not guaranteed to obtain an optimal solution because the decision made at some level cannot be recovered at later level.

### 2.1.2 Non-hierarchical Color Image Quantization

The popularity algorithm [2] is one of the first suggested algorithms in the field of color quantization. The algorithm is based on the assumption that a good color palette should contain the most frequent image colors. In this algorithm, the histogram of the image pixels is first created and sorted using the selection sort and then the most $N$ frequent colors are selected as the color palette. An improvement to the popularity algorithm was proposed by Hsieh and Fan [21] to select the $N$ most frequent colors that are separated by at least a predefined threshold distance. Some other improvements to popularity algorithm have been proposed in [22], and it has been further modified to work in CIELUV color space [23].

Larabi et al. [24] presented an extension to the popularity algorithm that takes into account the most frequent colors without forgetting the least frequent but important colors. This method is built around a matrix of local pallets created from local histograms. An input image is partitioned into a number of windows according to the image variance. The histogram of each local window is then calculated and the most frequent colors are selected from each local histogram to form a matrix of local palettes. The local palette matrix formed is considered as a new image representing the most frequent colors of the input image. The histogram of this new image is calculated and divided into two equifrequent parts, i.e. the total of the appearance frequencies is equal on both sides. Then each part is divided into two equi-frequent parts, until reaching the desired number of palette colors. The final palette selection is carried out on the basis of the frequency of local palettes criterion.

The k-means algorithm [25] is one of the most well-known color quantization algorithms. It is an iterative algorithm which starts with an initial set of centroids. This initial set can be generated by randomly selecting $k$ pixels from the input image. At each iteration, a scan through all pixels of the original image is performed to assign each pixel to the nearest centroid. Then a new set of centroids is generated as the means of the pixels associated to each centroid. These steps are repeated until the algorithm converges or the number of iterations reaches a specified value. A fuzzy version of the k-means algorithm, called fuzzy k-means, was proposed by Bezdek [26]. The fuzzy k-means is based on a fuzzy extension of the least square error criteria where data points are assigned to each cluster with some degree of membership.

The k-means algorithm is known to converge to a locally optimal solution. Its success to converge to a globally optimal solution depends significantly on its initial centroids. Several different initialization approaches for the k-means algorithm have been proposed. Astrahan [27] used initial centroids that are both separated with at least the average pairwise Euclidean distance and have a large number of occurrences within a multidimensional sphere of the initial points. In [28] all of the dataset points are ordered based on their distance to the overall dataset mean and the initial centroids are chosen as quintiles of these ordered points. In [29] the dimension with maximum variance is divided into k groups and the medians within each of the groups are chosen to initialize the k-means algorithm. Mirkin [30] used a "MinMax" procedure for initialization. The initialization method involves the calculation of dataset pairwise Euclidean distances. The two points with the largest pairwise distances are selected as the first and second centroids. The remaining centroids are chosen iteratively by selecting the closest point to each

centroid in the current set of centroids and out of them the most distant point is considered as a new centroid.

Dekker [31] studied the use of the one-dimensional self-organizing map (SOM) [32] for image quantization. The network contains one neuron for each desired cluster. Each neuron acquires a weight vector which is used as a representative. The weight vectors are initialized to positions on the main diagonal of the color space cube as a good first approximation to the input. Then input pixels that are obtained by multiple sampling of the image with large step sizes are repeatedly scanned to find the 'best' weight vector corresponding to the input in terms of Euclidean distance. This vector is then updated by moving it closer to the input. After learning is completed, pixels are mapped to the closest weight vector.

A similar approach to Dekker [31] was proposed by Papamarkos [33] which is based on Kohonen Self-Organized Feature Map (SOM) neural network. The proposed approach uses both the image color values and additional spatial features extracted in the neighborhood of each pixel to feed the SOM neural network. These features describe image local characteristics. Therefore, the color of each pixel is related to the colors and the texture of the neighboring pixels. As a result, the final image has not only the proper colors, but its structure approximates the local characteristics used. A fractal sub-sampling technique can be applied based on the Hilbert's space filling curve to speed up the training process.

Papamarkos et al. [34] proposed an adaptive color reduction technique as an improvement to [33]. The proposed technique achieves color reduction using a tree clustering procedure. In each node of the tree, a self-organized neural network classifier (NNC) is used which is fed by image color values and additional local spatial features. The NNC consists of a principal component analyzer (PCA)

and a Kohonen self-organized feature map (SOFM) neural network. The output neurons of the NNC define the color classes for each node. The PCA is used to manipulate the coordinate axes that the data falls on. The SOM is competitively trained according to the Kohonen's learning algorithm. After training, the output neurons of the SOM define the proper feature classes. Next, each pixel is classified into one of these classes and resumes the color of the class. In that way, the original is converted into a new image, which has a limited number of colors and its spatial characteristics approximate those defined by the features used. In order to reduce storage requirements and cut down computation time, a fractal scanning sub-sampling technique is used.

Zhang and Hu [35] introduced a method that makes essential modifications to the adaptive color reduction method [34] in two aspects: one is using hierarchical frequency sensitive competitive learning (HFSCL) to substitute Kohonen's SOM which can effectively solve the problem of dead units. And the other is using binary tree (2-way division) instead of giving different number of clusters at different level which can effectively solve the difficulties of criteria definition for tree spanning and help to define the general split criteria for various images. In addition, the binary tree structure allows the number of clusters to be incremented by one which is helpful to find out the actual number of input data.

Evolutionary optimization approaches also have been used in the problem of color quantization. A color image quantization algorithm based on Particle Swarm Optimization (PSO) is developed by Omran et al. [36]. Each particle represents a candidate solution to the optimization problem, thus the swarm represents a number of candidates. Each particle in the swarm is represented by three characteristics: the current position of the particle, the current velocity of the particle,

and the personal best position of the particle. The proposed algorithm initializes each particle randomly from the color image to contain $k$ centroids (i.e. color triplets). The set of $k$ color triplets represents the color palette. The k-means clustering algorithm is then applied to each particle in order to refine the chosen colors and to reduce the search space. Each pixel is then assigned to the cluster with the closest centroid. The performance of each particle (i.e. how close the particle is to the global optimum) is measured using a mean squared error fitness function. The PSO position and velocity are then updated. The procedure is repeated until a stopping criterion is satisfied. The color palette of the global best particle after the maximum iterations is chosen as the optimal result.

A new evolutionary optimization approach [37] is proposed based on ant colony clustering algorithm and applied to color image quantization. The algorithm was inspired by the ants picking up or dropping behaviors. The general idea is that an individual ant is able to pick up or drop an object at some location where more objects of the similar type are present. The algorithm starts by considering every coordinate pair $(x, y)$ of the input image pixels as a virtual ant with loaded initial state. When a randomly unloaded moving ant encountered an object, the less similar objects around the object loaded by the ant, the more probabilities for the ant to pick it up. Otherwise, the more similar objects around the object, the more probabilities for an ant to drop it. So it can assure that the objects into the bigger heaps are not destroyed, and some individual objects can be clustered in small heaps. According to the dropping or picking-up behavior, similar data objects can be grouped in a cluster around a clustering center.

## 2.2 Objective Image Quality Metrics

Over the past few decades many metrics for the assessment of image quality have been proposed. Although it is difficult to classify all of these metrics, a rough classification would be good for understanding the fundamental ideas and forward real-world applications as well as future research. Objective image quality assessment metrics can be classified on the basis of their application into general-purpose image quality assessment metrics and application-specific image quality assessment metrics [38]. General-purpose image quality assessment metrics are those metrics that are designed to be used for the quality assessment in a variety of different applications and have no assumptions about a specific distortion type. They are usually designed using common features or hypothesized assumptions about the human visual system. On the other hand, the application-specific image quality assessment metrics are designed for specific applications. For example, image and video compression, transmission, halftoning, watermarking, denoising, restoration and enhancement are some of the largest application areas of image quality assessment techniques.

Another popular classification of objective image quality assessment metrics is based on the availability of the information from reference image. The metrics are classified into full-reference, reduced-reference and no-reference metrics; full information from the reference image is needed for the quality assessment of the test image in the full-reference metrics, partial information from the reference image is needed in the reduced-reference metrics, while no information from the reference image is needed in the no-reference metrics.

In this section, a review of some full-reference image quality assessment metrics

is presented. A more detailed review of image image quality metrics can be found in [39–42].

### 2.2.1 Error Sensitivity Based Metrics

Mean square error (MSE) and its counterparts signal to noise ratio (SNR) and peak signal to noise ratio (PSNR) are among the earliest and most popular image quality metrics. The popularity of these metrics comes from their simplicity in calculation, having clear physical meanings, and they are mathematically convenient in the context of optimization.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[I_r\left(i,j\right) - I_t\left(i,j\right)\right]^2 \tag{2.1}$$

$$SNR = 10\log_{10}\left(\frac{\frac{1}{MN}\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N} I_r\left(i,j\right)^2}{MSE}\right) \tag{2.2}$$

$$PSNR = 10\log_{10}\left(\frac{L^2}{MSE}\right) \tag{2.3}$$

where $I_r$ is the reference image, $I_t$ is the test image, $M$ and $N$ are the dimensions of the images. $L$ is the maximum possible pixel value of the image. When the pixels are represented by 8 bits, this value is 255.

Mitsa and Varkur [43] improved the MSE by incorporating human visual system modeling. The human visual system modeling employed was a contrast sensitivity function (CSF) weighting where the frequency-domain error between the

original and test images is weighted with CSF while computing MSE.

In [44] a just-noticeable-difference (JND) model was incorporated to define a Peak Signal to Perceptible Noise Ratio (PSPNR) quality metric. They model luminance masking and activity masking to obtain a JND profile. The PSPNR has the same definition as given in Eq. (2.3) except that the MSE expression is computed for the error image after removing the JND values.

Lambrecht and Farrell [45] proposed color masked signal to noise ratio (CMP-SNR) metric based on the opponent-colors theory to model color perception and on multi-channel model of spatial vision to account for pattern sensitivity. The opponent-color space chosen was the same as proposed by Poirson and Wandell [46]. The metric transforms the original and distorted images to the opponent color space, a Gabor filter bank is then applied to each component of the original and distorted images. After that a contrast threshold is detected based on sensitivity and masking phenomenon and the error signal in each component is multiplied by the inverse of the detected contrast threshold. Finally, the PSNR metric is computed by pooling the error over the channels.

Zhang and Wandell [47] proposed S-CIELAB metric which is a spatial extension to the CIELAB color metric for measuring color reproduction errors of digital images. Digital color images are transformed into an opponent-colors space. Each opponent-colors image is convolved with a kernel whose shape is determined by the visual spatial sensitivity to that color dimension; the area under each of these kernels integrates to one. The filtered images are then transformed to the CIELAB representation. Finally, the color difference for each pixel is calculated, these values are averaged to achieve one value for overall image difference.

Iordache and Beghdadi [48] presented a new measure, $\mathrm{SNR}_W$, of image degradation based on a joint spatial/spatial-frequency representation using Wigner-Ville Distribution (WVD). The idea behind this approach is to exploit the fact that structured distortions are usually highly concentrated in the spatial/spatial-frequency domain. Between two distortions with the same energy, i.e. same signal-to-noise ratio (SNR), the more disturbing is the one having a peaked energy distribution in spatial/spatial-frequency plane. The high energy components will also induce large interference terms, so that the WVD of the distorted image will be significantly different from the WVD of the original image.

The drawbacks of using the Wigner-Ville Distribution (WVD) in [48] have been addressed by Beghdadi and Popescu [49] and a new image distortion measure was introduced based on a non-redundant wavelet decomposition as a replacement to WVD. The metric is computed in a similar way as Eq. (2.2) for measuring the distortion between the maximum of the absolute value of the difference between the wavelet coefficients of the original and the distorted images. The use of the maximum of the absolute value of the difference over resolution levels is motivated and inspired by some findings on the nonlinear behavior of the HVS.

Samet et al. [50] proposed a new metric, new weighted mean square error (NwMSE), that takes into account a pixel's neighborhood information. The NwMSE metric is based on the assumption that the neighborhood pixels, within the same block, are of good continuity. The weighting function used behaves as follows: in high spatial frequency activity regions where the HVS is less sensitive to the contrast, the partial MSE for a particular pixel will not be completely modified; while for low spatial frequency activity regions where the HVS is more sensitive to the contrast, the partial MSE will be amplified and contributes to the

increase of the NwMSE value.

An image quality metric called multi-exposure peak signal- to-noise ratio (mPSNR) [51] is proposed for the quality assessment of high dynamic range (HDR) images. To account for all normal viewing conditions, the HDR image is tone mapped to a number of different exposures uniformly distributed over the dynamic range of the image. As a tone mapping operator, a simple gamma-adjustment is used after exposure compensation. For each exposure, the mean square error (MSE) is computed on the resulting low dynamic range (LDR) image, and then the peak-signal-to-noise ratio (PSNR) is calculated using the mean of all MSEs.

In [52] a new metric, PSNR-HVS, is proposed based on Discrete Cosine Transform (DCT). The metric uses a scanning window of size 64×64 pixels for mean shifting and the contrast stretching in away similar to that used in [53]. The $\mathrm{MSE}_H$ is calculated between the DCT coefficients of each 8×8 image block in the original image and the DCT coefficients of the corresponding block in the distorted image. The differences between DCT coefficients are corrected with correcting factors. The values of correcting factors have been obtained using the quantization table for the color component Y of JPEG. An improved version of PSNR-HVS, PSNR-HVS-M, is proposed in [54] using a model of visual between-coefficient contrast masking of DCT basis functions and the contrast sensitivity function (CSF).

Chandler and Hemami [55] proposed a metric, the visual signal-to-noise ratio (VSNR), based on nearthreshold and suprathreshold properties of the HVS. The metric consists of two stages: in the first stage, contrast thresholds are used to detect visible distortions in the image. Those thresholds are computed via wavelet-based models of visual masking and visual summation in order to determine whether the distortions in the distorted image are visible. If the distortions

are below the threshold of detection, the distorted image is deemed to be of perfect visual fidelity (VSNR $=\infty$) and no further analysis is required. If the distortions are visible (suprathreshold), a second stage is applied which operates based on the low-level visual property of perceived contrast and the mid-level visual property of global precedence. These two properties are modeled as Euclidean distances in distortion-contrast space of a multi-scale wavelet decomposition, and VSNR is computed in a similar way as Eq. (2.2) on a simple linear sum of these distances.

Wang and Li [56] incorporated the idea of information content weighted pooling and applied it to peak signal to noise ratio (PSNR). The approach is based on the hypothesis that when viewing natural images, the optimal perceptual weights for pooling should be proportional to local information content, which can be estimated in units of bit using advanced statistical models of natural images. The Gaussian scale mixture (GSM) model for natural images is used here as a statistical image model to quantify the image local information.

### 2.2.2 Structural Similarity Based Metrics

As an alternative to the traditional error-sensitivity approaches, the principal idea underlying the structural similarity approach is that the human visual system (HVS) is highly adapted to extract structural information from visual scenes, and therefore, a measurement of structural similarity (or distortion) should provide a good approximation to perceptual image quality. The main question that should be answered to convert the general structural similarity principle into specific image quality assessment algorithms is: how to define structural distortions. The following methods use different definitions for the structural distortions.

The Universal Image Quality Index (UIQ) proposed by Wang and Bovik [53] is a mathematically defined image quality metric that is easy to calculate and applicable to various image processing applications with no HVS model incorporated. The proposed approach models any image distortion as a combination of the factors: loss of correlation, luminance distortion, and contrast distortion. For a reference image $x$ and a distorted version of it $y$, the correlation is estimated by the correlation coefficient between $x$ and $y$ which measures the degree of linear correlation between them, the luminance is estimated by mean, and the standard deviation is considered as an estimation of the contrast. These three statistical features are calculated locally by moving a sliding window of size $8\times8$ pixels starting from the top-left corner of the image to the bottom-right corner through each pixel of the image.

Wang et al. [57] generalized the UQI and introduced the structural similarity index (SSIM) based on the degradation of structural information. The structural information in an image was defined as those attributes that represent the structure of objects in the scene, independent of the average luminance and contrast. SSIM models any image distortion as a combination of three distortions: luminance distortion, contrast distortion, and structural distortion. SSIM index uses an $11\times11$ circular-symmetric Gaussian weighting function, with standard deviation of 1.5 and samples normalized to unit sum, to compute the local statistics in order to avoid undesirable blocking artifacts in the UQI index map that are exhibited due to using local $8\times8$ square window which moves pixel-by-pixel over the entire image.

The single-scale SSIM [57] was extended into a multi-scale version (MSSIM) [58] which provides more flexibility than single-scale approach in incorporating the

variations of image resolution and viewing conditions. The contrast and the structure comparisons are calculated at each scale of the reference and distorted images while the luminance comparison is computed only at highest scale. The overall MSSIM evaluation is obtained by combining the measurement at the different scales.

Wang and Simoncelli [59] addressed a major drawback of the spatial domain SSIM [57] that it is highly sensitive to translation, scaling and rotation of images. They extended the spatial domain SSIM index to the complex wavelet transform domain SSIM (CWSSIM). This extension makes it insensitive to these non-structured image distortions that are typically caused by the movement of the image acquisition devices, rather than the changes of the structures of the objects in the visual scene.

Both the SSIM [57] which is proposed in spatial domain and the CWSSIM [59] which is proposed in complex wavelet transform domain cannot be used directly in Discrete Wavelet Transform (DWT)-based image processing algorithms in order to optimize them. A transformation to/from DWT at each step is required which increases computation load significantly. Yang et al. [60] proposed a Discrete Wavelet Transform based SSIM (DWT-SSIM) that can be easily embedded in DWT-based image processing algorithms. This metric can be summarized as follows: first, the original and the distorted images are carried on five-level DWT using Lifting Daubechies (9, 7) wavelets. Secondly, The SSIM metric is used to calculate each band's similarity. Finally, the similarity of the whole image is obtained by the weighted mean of all DWT-SSIMs.

Brooks and Pappas [61] observed a weakness associated with the single-scale CWSSIM [59] that it is not able to account for the low spatial frequency image

distortions. Therefore, they extended the single-scale CWSSIM to combine the results from multiple wavelet scales using an approach similar to that used in multi-scale SSIM [58] creating a weighted complex wavelet SSIM (WCWSSIM).

Li and Bovik [62] introduced a weighted by region type SSIM called four-component SSIM (4-SSIM) that takes into account image content. The reference and distorted images are segmented into four categories of regions: changed edges, preserved edges, textures and smooth regions. Changed and preserved edge regions are found where a gradient magnitude estimate is large, while smooth regions are determined where the gradient magnitude estimate is small. Textured regions are taken to fall between these two thresholds. These regions are used to apply non-uniform weights to the SSIM values.

In [63] an improved SSIM algorithm called edge-based structural similarity (ESSIM) is proposed which compares edge information between the original image and the distorted one. In this metric, the luminance and contrast comparisons remain the same as in SSIM but the structure comparison is replaced by edge-based structure comparison where the comparison is computed over edge information of the reference image and the distorted one instead of the original image information. The Sobel operator is used to obtain the edge information due to its simplicity and efficiency.

A gradient-based SSIM (GSSIM) [64] was proposed in a way similar to ESSIM [63]. The only difference is that both the structure and contrast comparisons in the SSIM are replaced by edge-based structure and contrast comparisons.

In [65] an extension to SSIM, the content-based metric (CBM), was proposed based on a fuzzy Sugeno integral. In this approach, the luminance and the con-

trast comparisons are similar to SSIM, but the structure comparison is modified by taken the absolute value of the covariance. The modified SSIM index map is calculated and partitioned into three parts: edges, textures and flat regions according to their gradient. Then fuzzy integral is used to fuse the SSIM in the three parts respectively. Finally the overall image quality is evaluated with the weighting average of the three regions. An improved version of CBM, RCBM, is presented in [66] based on fact that the CBM method uses pixel-wised integral, so it is less flexible in the analysis. The RCBM method is supposed to be more flexible by using the rough fuzzy integral instead of the fuzzy integral.

Lam and Loo [67] presented an image quality metric that exploits the quadtree decomposition. Quadtree decomposition is a recursive operation in which the input image is first divided into four quadrants of the same size. The homogeneity of each quadrant is measured with a metric. For each non-homogeneous quadrant, further recursive quadrant decomposition is performed. This recursive operation continues until a quadrant is homogeneous or when the resulting quadrants are too small. The overall metric is evaluated by pooling three comparisons namely: contrast, luminance and structure comparisons. The contrast and luminance comparisons are computed based on the size and location of quadtree segments of the reference image in a similar way to SSIM [57]. The structure comparison is defined as the percentage of the number of non-zero values in the pixel-wise subtraction of the quadtree decomposition boundaries of the reference and distorted images to the total number of image pixels.

Wang and Li [56] incorporated the idea of information content weighted pooling and applied image quality metrics. This approach is based on the hypothesis that the HVS is an optimal information extractor and to achieve such optimality, the

image components that contain more information content would attract more visual attention. The Gaussian scale mixture (GSM) model for natural images is used here as a statistical image model to quantify the image local information. By combining information content weighting with multi-scale SSIM [58], they define an information content weighted SSIM measure (IW-SSIM).

Kolaman and Yadid-Pecht [68] proposed the Quaternion Structural SIMilarity (QSSIM) metric for color images which is a vectorial expansion of structure similarity using quaternion image processing. The QSSIM was composed to be the same as SSIM but with quaternion subparts. After the original and reference images are transformed to Quaternion space, the color mean was done by calculating the mean of each color channel separately over an image area, and the color contrast was done by calculating standard deviation for each color channel as in the case of color mean The expansion to color cross correlation, which is the main reason for using quaternions was defined as in [69] as the mean of the dot products of reference image quaternions and the conjugate of their corresponding distorted image quaternions over an image area.

Shnayderman et al. [70] proposed an image quality assessment metric, M-SVD, in which singular value decomposition (SVD) is used for representation of the structural information of an image. The Singular value decomposition of $m \times n$ matrix $A$ is defined as $A = U\sigma V^T$ where $U, V$ and $\sigma$ represent the left singular vector matrix, the right singular vector matrix, and the diagonal matrix of singular values. In M-SVD, the reference and distorted images are divided into smaller $8 \times 8$ blocks and the SVD is applied to each block separately. The distances between the singular values of each original image block and the singular values of its corresponding distorted image block are computed and pooled to be expressed as

a single numerical value. A second pooling is used to combine the contributions from all blocks to form a quality score corresponds to the perceptual quality.

Narwaria and Lin [71] showed that the singular vectors $U$ and $V$ of the singular value decomposition (SVD) better represent image structure and are more important than singular values $\sigma$ for quality assessment. And so, they proposed an extension of the M-SVD metric [70] that uses both singular vectors and values for comprehensive image quality evaluation.

### 2.2.3 Information Fidelity Based Metrics

Apart from the error sensitivity and structural similarity approaches, information fidelity approach is a new approach to the quality assessment problem. It quantifies the information about the reference image that could ideally be extracted by the receiver (the brain) from the test image.

A novel Information Fidelity Criterion (IFC) [72] was proposed for image quality assessment using natural scene statistics (NSS). In this approach the quality assessment problem is modeled as an information fidelity problem where a natural image source communicates with a receiver through a channel. The channel imposes fundamental limits on how much information could flow from the source (the reference image), through the channel (the image distortion process) to the receiver (the human observer). Using a model for the source and a distortion model, the statistical information shared between the test and the reference images is a good way of quantifying fidelity that could relate well with visual quality. The Gaussian Scale Mixture (GSM) model in the wavelet domain was used as a NSS model for the image source. The distortion model is also described in the

wavelet domain. It is a simple signal attenuation and additive Gaussian noise in each subband. The IFC is then defined as the mutual information between the source and the distorted images.

Sheikh and Bovik [73] presented the Visual Information Fidelity (VIF) index which is an extension of the IFC. The connections between image information and visual quality is further explore where the reference image is modeled as being the output of a stochastic natural source that passes through the HVS channel and is processed later by the brain. Thus the HVS is considered to be a channel as well that imposes limits on how much information could pass through it to the brain (receiver). The information content of the reference image is quantified as being the mutual information between the input and output of the HVS channel. This is the information that the brain could ideally extract from the output of the HVS. The image source and distortion channel are modeled as in the case of the IFC. The HVS is modeled as stationary, zero-mean, additive white Gaussian noise in the wavelet transform domain. Then the ratio between the amount of image information that is visually extracted from the distorted image to the information content present in the reference image used as a measure of image quality.

### 2.2.4 Feature Similarity Based Metrics

The feature similarity is another approach for the assessment of image quality in which the similarity between the test and reference images is computed in a feature domain space. The fundamental idea lying beneath the feature similarity approach is that the human visual system (HVS) understands an image mainly according to its low-level features, and therefore, quantifying of feature similarity

(or distortion) should provide a good approximation to perceptual image quality.

A feature based image quality metric [74], namely Riesz-transform based Feature SIMilarity metric (RFSIM), is proposed based on the fact that the human vision system (HVS) perceives an image mainly according to its low-level features at key locations. The $1^{st}$-order and $2^{nd}$-order Riesz transform coefficients of the image are taken as image features, while a feature mask is defined as the edge locations of the image formed by the Canny operator (without thinning operation). The RFSIM index between the reference and distorted images is measured by comparing their Riesz transform feature maps at key locations marked by the feature mask.

Zhang et al. [75] introduced a Feature SIMilarity (FSIM) index. The phase congruency (PC), which is a dimensionless measure of the significance of a local structure, is used as the primary feature in FSIM. Considering that PC is contrast invariant while the contrast information does affect HVS' perception of image quality, the image gradient magnitude (GM) is employed as the secondary feature in FSIM. Both PC and GM play complementary roles in characterizing the image local quality. FSIM measurement between two images is separated into two components, each for PC or GM. The PC values of the two images are computed and compared to get a PC similarity; the same thing is done with GM values. The PC and GM similarities are then combined together to get a single local FSIM map. The FSIM index is defined as a weighted average of FSIM map values. The maximum of PC values of the two input images were used to weight pooling the FSIM map.

An image quality metric based on the Harris response (HR) feature which is computed from the gradient information matrix and its eigenvalues was introduced

by Kim and Park [76]. This metric was improved in [77] by computing HR values only at pixels having large differences between the reference and distorted images. Those pixels are detected by using the edge suppression technique, which is based on the cross projection tensor[78].

Unlike the conventional feature based image quality metrics that measure the difference between feature values of the reference and distorted images, Kim and Park [79] proposed a feature based image quality metric which evaluates the image quality by measuring the change of the sign of feature values. The feature values are obtained by filters such as gradient and Gabor filters. Finally, the proposed image quality metric is computed as the Hamming distance between two binary codes, in which they use the phase quantization code (PQC) also known as phase quadrant demodulation code, that represents the sign of the feature values is used to measure the image quality. This phase quantization code-based image quality metric was extended later to amplitude/phase quantization code-based image quality metric [80].

Zhai et al. [81] presented the Log Gabor Phase Similarity (LGPS) image quality assessment metric based on measuring of similarities between phases in log Gabor transform domain. An image is firstly decomposed by a filter bank consisting of a pair of log Gabor filters. The phase maps are then computed from the responses of each filter pair. A local normalized cross-correlation is used to gauge the resemblance between the two phase maps, and this correlation map is then averaged for a numerical measurement of the predicted quality of the image.

Inspired by the intuition that HVS is quite sensitive to image local orientation features, Wang et al. [82] proposed a new image quality metric. Considering that the distributions of local orientations in one image reflect the high-order

statistics of image primitives well, Histograms of Oriented Gradients (HOG) [83] is used as an image structure feature to evaluate image distortions. A distortion map is generated by computing the Euclidean distance between histograms of the reference and distorted images. Finally, the numerical expression for the proposed image quality metric is computed as the average of the distortion map.

It has been observed that most of the image quality metrics proposed over the years are designed to perform best on many image distortions. Although color quantization noise is frequently met in practice, it has not been given too much attention in color image visual quality assessment compared with other image distortions such as image compression, transmission, acquisition, watermarking distortions. However, more research should be concentrated on quality evaluation of color quantized images. Furthermore, the correlation of the state-of-the-art image quality assessment metrics with the human opinion scores in estimating the perceived quality of color quantized images should be explored.

# CHAPTER 3

# Subjective Image Quality Assessment

## 3.1 Introduction

Image quality assessment is an important tool in image processing systems. Image quality assessment methods can be classified into two categories: objective and subjective methods. The objective image quality assessment methods are computer based methods that can automatically predict the perceived image quality. On the other hand, the subjective image quality assessment methods are human based methods in which the opinions are obtained through the subjective testing where human subjects judge the quality of a sequence of images. The opinions for the subjects are then processed to obtain the final scores that indicate the perceptual quality of each image. As the ultimate receivers of images are human eyes, the human subjective opinion is the most reliable value for indicating the image perceptual quality.

Although subjective image quality assessments are very difficult to carry out, expensive, and time consuming which makes them impractical for most image applications. However, the obtained subjective scores can be considered as a ground truth of the image perceptual quality. They can be used to reliably evaluate the correlation of the objective image quality assessment algorithms with the human perception. Moreover, subjective studies can also improve the development of new objective image quality assessment methods that are well correlated with human

perception which is the ultimate goal of quality assessment methods. Then these developed quality assessment methods can be utilized for guiding the corresponding application. This necessitates the creation of image databases with subjective opinion scores defining the human perception of quality.

In this chapter I will describe the details of the extensive psychometric experiments done to collect subjective quality scores for a diverse set of color quantized images. The subjective mean opinion scores (MOS) from the study were used for training and testing the algorithms presented in this dissertation and also for validating the correlation of the existing image quality metrics with human perception in evaluating the quality of color quantized images.

## 3.2 Methodologies for Subjective Image Quality Assessment

International recommendations for subjective video/image quality assessment have been standardized by the International Telecommunication Union (ITU) in the ITU-R Recommendation [84] which suggests standard viewing conditions, criteria for the selection of observers and test material, assessment procedures, and data analysis methods. These recommendations define some of the most commonly used methodologies for subjective quality assessment. Mainly, there are three types of such experimental methodologies for conducting subjective quality assessment studies: double-stimulus methods, single-stimulus methods, and stimulus-comparison methods. These methods have generally different applications and the choice of a specific method depends on the context, the purpose,

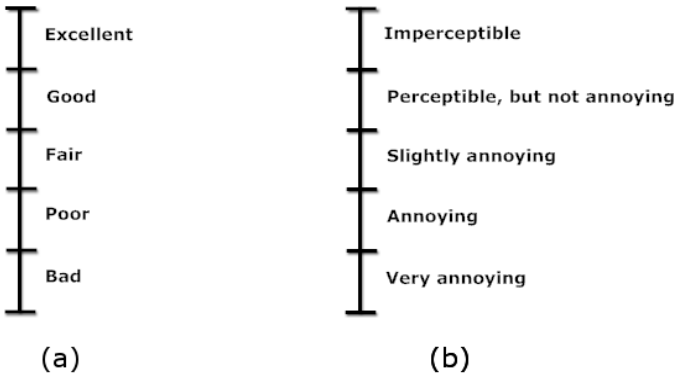



Figure 3.1: Stimulus categorical scale. a) Continuous scale, b) Impairment scale.

and where in the development process of the data the test is to be performed.

### 3.2.1 Double Stimulus Methods

Double stimulus methods are cyclic in that a pair of images is shown in each assessment trial. One image in the pair is an undistorted reference image, and the other is the same image that may or not be distorted. The images and distortions should be presented in a pseudo-random sequence and, preferably, in a different sequence for each test session. Double stimulus methods have the advantage of being more robust to contextual effects, and the results are considered to be more stable and reproducible. There are two main types of double stimulus methods: double stimulus impairment scale method and double stimulus continuous quality scale method.

#### Double Stimulus Impairment Scale Method

The assessor is first presented with an unimpaired reference image, then with the same image impaired. Following this, he is asked to vote on the second, keeping

in mind the first. The method uses the five-grade impairment scale as shown in Fig. 3.1(b), the range of impairments should be chosen so that all grades are used by the majority of observers. At the end of the series of sessions, which last up to half an hour, the mean score for each test image is calculated.

There are two variants to the structure of presentations of the double stimulus impairment scale method, I and II described below:

**Variant I:** Each pair of images, reference and test, are presented only once. An assessment trial consists of four displays: a reference image of 10 seconds, a mid-gray field of 3 seconds, a test image of 10 seconds, and a mid-grey field of 5-11 seconds as shown in Fig. 3.2(a). The assessor should look at the images for the whole of the duration of T1 and T3. Voting should be given only during T4.

**Variant II:** Each pair of images, reference and test, are presented twice. The structure of each of the two presentations is similar to that of variant I. The assessor should look at the images for the whole of the duration of the first presentation. Voting should be given only during the second presentation (see Fig. 3.2(b)).

Although Variant II is more time consuming than variant I, it may be used when discrimination of very small impairments is required.

### Double Stimulus Continuous Quality Scale Method

In this method, the assessors vote the quality of two versions of the same test image in each assessment trial. One version is unimpaired while the other might or might not contain an impairment. The unimpaired image is included to serve as

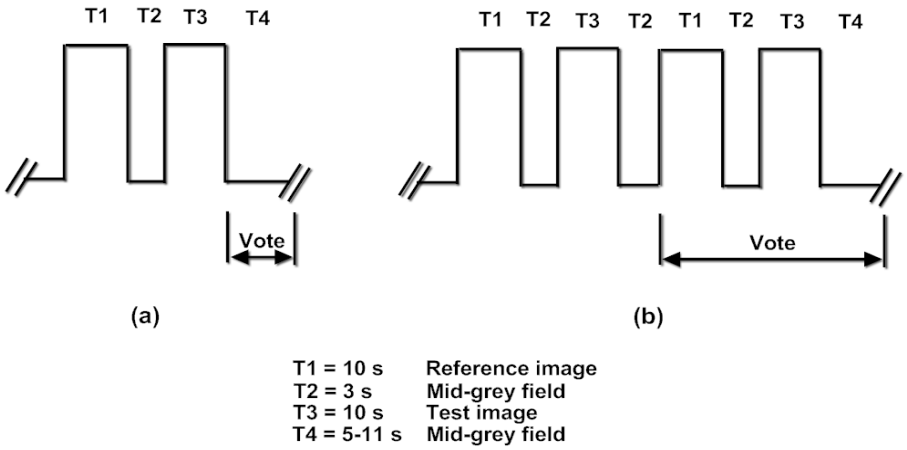



Figure 3.2: Presentation structure of double stimulus impairment scale method. a) Variant I, b) Variant II.

a reference image, but the assessors are not told which is the reference image. The position of the reference image is changed during the test session in pseudo-random fashion.

In each session, which last up to half an hour, the assessors assess the quality of test images of each assessment trial by simply putting a mark on a vertical scale. The vertical scales are printed in pairs to accommodate the double presentation of each test image. The scales provide a continuous rating system to avoid quantizing errors, but they are divided into five categories that are defined in semantic terms for general guidance (see Fig. 3.1(a)). There are two variants for the presentation of the test images:

**Variant I:** There is a single assessor for each presentation; he free to switch between the pair of test images for a period of up to 10 seconds until he is satisfied that he has established his opinion of each test image.

**Variant II:** All the assessors are shown, before to recording results, one or more times the pair of images (reference and test) simultaneously to establish their

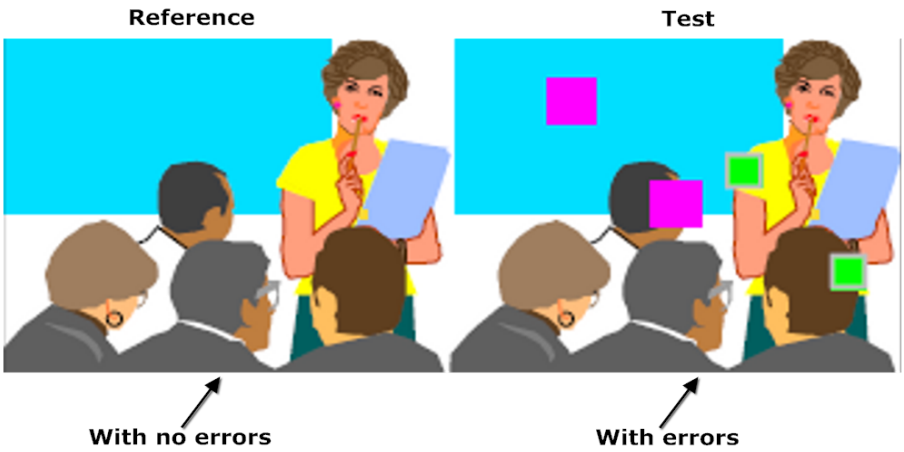



Figure 3.3: Display format for simultaneous double stimulus for continuous evaluation method

opinion of each test image. Then the pair is shown again one or more times while the results are recorded. For still images, a 3-4 s sequence and five repetitions (voting during the last two) may be appropriate.

The pair of assessments (reference and test) for each assessment trail is converted from measurements of length on the score sheet to normalized scores in the range 0 to 100. Then, the difference between the assessment of the reference and the test images are calculated.

### Simultaneous Double Stimulus for Continuous Evaluation

This method is a special case of double stimulus method with a slight modification concerning the way of presenting the images to the assessors. Simultaneous double stimulus for continuous evaluation method is used when the fidelity of visual information has to be evaluated. Therefore the assessors are shown a pair of images at the same time side by side on the same monitor or if not possible on two aligned monitors. One image is the undistorted reference image, and the

other one is the distorted test image (see Fig. 3.3). The assessors, who are aware of which image is the reference and which image is the test, are asked to check the differences between the two images and to judge the fidelity of them by moving the slider. When the fidelity is perfect, the slider should be at the top of the scale range (coded 100); when the fidelity is the worst, the slider should be at the bottom of the scale (coded 0).

## 3.3 Designing the Image Database

### 3.3.1 Choice of Input Images

Our image database consists of 25 reference images collected from the Internet. Those images reflect a variety of image contents and include visually significant objects, uniform regions, slowly varying color gradients, edges, and high level of details. Fig. 3.4 shows the reference images used in our subjective study.

### 3.3.2 Color Quantization Algorithms

All images in our database are of size $512 \times 512$ pixels for the purpose of carrying out subjective experiments. Each of the reference images has been quantized into seven levels (4, 8, 16, 32, 64, 128, and 256 colors) using five color image quantization algorithms that are popular in literature. These algorithms represent different approaches: dividing approach, merging approach, clustering approach, and neural networks approach. The color quantization algorithms used are: *K-means algorithm* [25], *Median Cut algorithm* [2], *Wu's algorithm* [13], *Octree algorithm*

Figure 3.4: The reference images used in the study

[16], *Dekkers SOM* [31] .

### 3.3.3 Number, Selection, and Training of Subjects

A group of twenty two undergraduate students participated in our psychometric experiment. The majority of the subjects were males and they were non-experts with image quality assessment. The reliability of the assessors was qualitatively evaluated by checking their behavior when reference/reference pairs where reliable subjects are expected to give evaluations very close to the maximum point in the quality scale.

Each subject was individually briefed about the goal of the experiment, what they are going to see, what they have to evaluate and how they express their opinion, the grading scale, the sequence, and timing. The subjects also have been shown some examples in how to evaluate the quality of quantized images. Those examples approximate the range of quality of the images for different quantization levels. Images in the training phase were different from those used in the actual experiment.

### 3.3.4 Display Equipments

The psychometric experiments were conducted in a lab with normal indoor illumination environment using Microsoft Windows workstations. The display monitors were all 19-inch CRT and were all approximately the same age. Although the monitors were not calibrated, they were set to the same display settings. A java-based interface was used to show the images and to enable the observers to vote

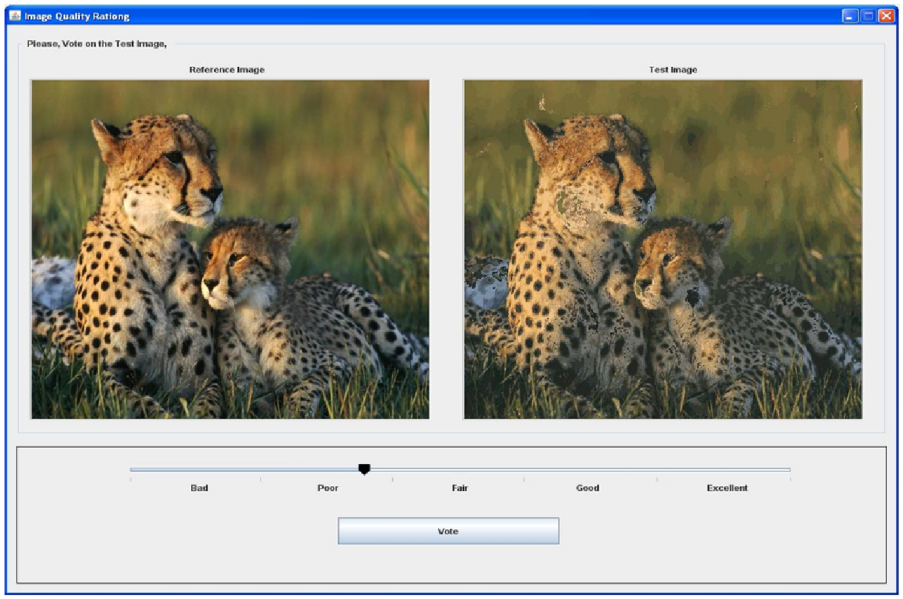



Figure 3.5: Interface for the subjective study.

on the quality of the test images. The software was designed in a way that the observers assess the quality of a quantized image compared with a reference image by simply dragging a slider on a quality scale. The quality scale which is of range [0,100] was unmarked numerically but labeled and divided into five equal categories: "Bad", "Poor", "Fair", "Good", and "Excellent" to be used as general guidance. The position of the slider reflects the score given by the observer for that image and its position was reset after each presentation. Fig. 3.5 shows the user interface of the quality assessment software.

### 3.3.5 Subjective Quality Tests

Since the fidelity of the quantized images to the reference images has to be evaluated, simultaneous double stimulus for continuous evaluation (SDSCE) method was used in conducting the psychometric experiments. In SDSCE a set of subjects is watching a pair of images (reference and test) at the same time and they are

asked to assess the quality of the test image compared with the reference image. In our image database there are 875 test images (25 reference images $\times$ 5 quantization algorithms $\times$ 7 quantization levels) to be assessed and as recommended by ITU [84] each session should not last more than 30 minutes; therefore, the overall subjective tests were divided into five sessions (175 test images for each session). Five dummy images were added at the beginning of the first session and not considered in the calculation; their purpose is to stabilize the subjects to the rating process. Subjects were shown images in a random order and this order is unique for each subject.

### 3.3.6 Outliers Detection and Subject Rejection

Before starting analysis of the data, a screening of the subjective raw scores was conducted to eliminate observers with unstable scores [84]. The generalized ESD many-outlier procedure [85] was run twice to detect outliers within the subjective raw data. The generalized ESD many-outlier procedure selects the maximum $k$ deviations from the mean and compares them with their corresponding critical values $\lambda_i, i = 1, .., k$ that define cut points to decide whether an observation is an outlier. The values of $\lambda_i s$ are computed based on the percentage points from the Student's $t$ distribution. If at any step $i$ a maximum deviation is greater than its corresponding critical value $\lambda_i$ then the extreme observations for the first $i^{th}$ maximum deviations are all considered to be outliers even if some of them are smaller than or equal to their corresponding critical values. About 2.66 % of the subjective raw data was deleted as being outliers. All quality evaluations of a subject were rejected if more than 5 % of his evaluations were outliers. Only one of the observers was rejected based on this criterion.

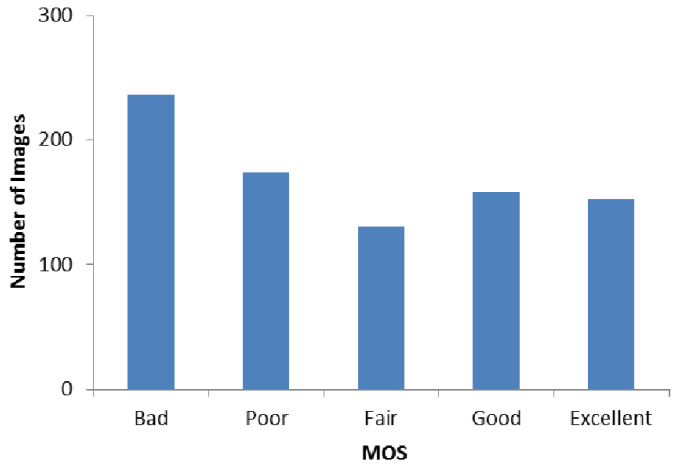



Figure 3.6: Histogram of the MOS values for the perceptual quality range.

### 3.3.7 Calculation of Mean Opinion Scores

To calculate Mean Opinion Scores (MOS), the subjective raw data is first converted to $Z$-score (after outliers removal) to minimize the variation between individual subjective values due to not using the full range of quality scale by the different subjects during the image quality rating process [86]:

$$z_{ij} = \frac{(v_{ij} - \overline{v_i})}{\sigma_i} \tag{3.1}$$

where $v_{ij}$ is the raw scores given by the $i^{th}$ subject to $j^{th}$ test image, $\overline{v_i}$ and $\sigma_i$ are the mean and the standard deviation of raw scores over all test images evaluated by the $i^{th}$ subject. The final MOS for each test image $j$ is obtained by averaging all $Z$-scores $z_{ij}$ given to that image by all subjects.

As the perceptual qualities of the images in the database should span the entire range of visual quality and exhibit good perceptual quality separation [87, 88], Fig. 3.6 depicts the histogram of the subjective MOS. Notice how the distribution of the subjective MOS is uniformly distributed demonstrating a good perceptual

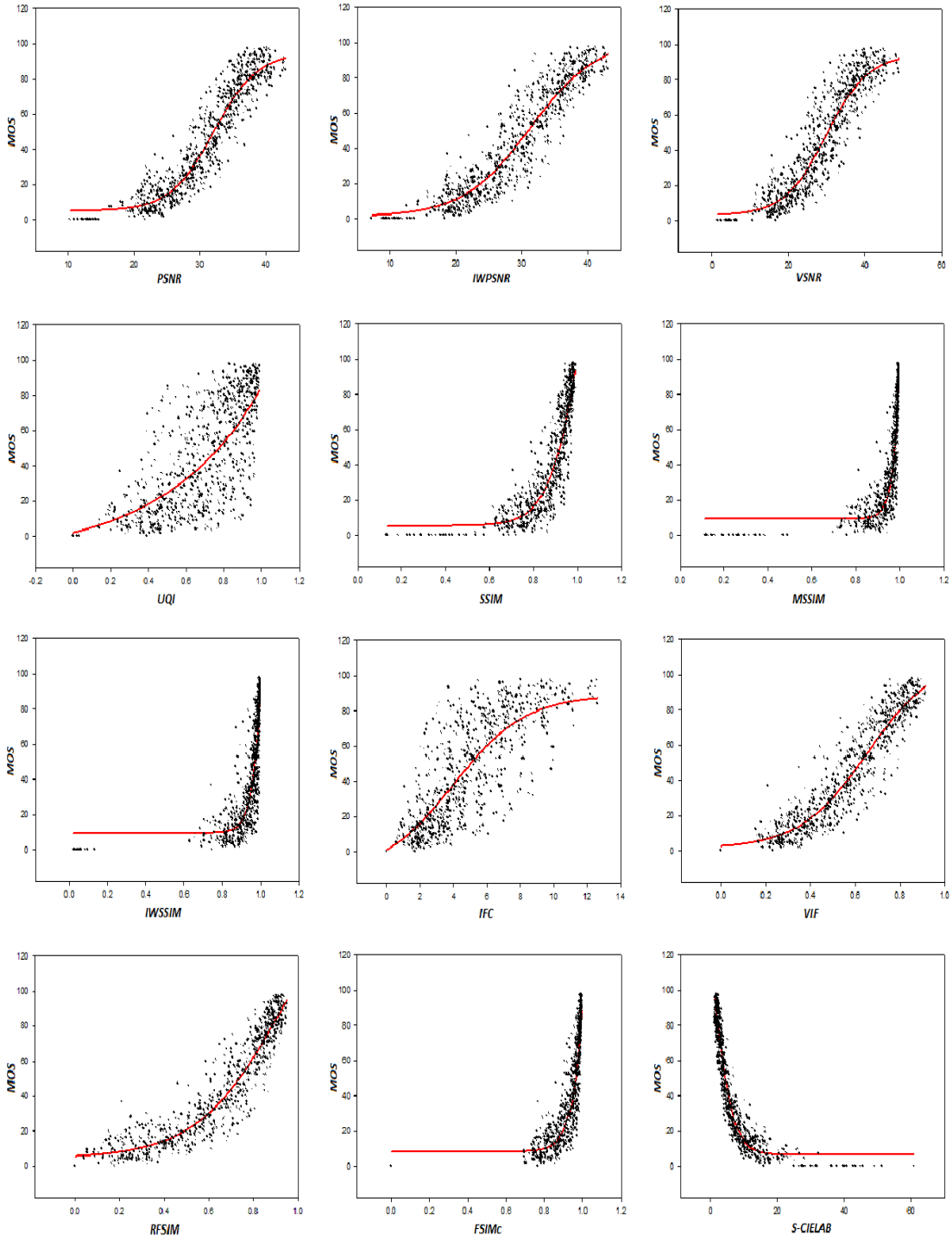



Figure 3.7: Scatter plots for the subjective MOS versus the quality scores form different image quality metrics: PSNR, IW-PSNR, VSNR, UQI, SSIM, MSSIM, IW-SSIM, IFC, VIF, RFSIM, FSIMc, S-CIELAB.

separation. It can be observed that the scores span the entire range of perceptual qualities from low to high values.

## 3.4 Image Quality assessment Metrics and Human Perception

The performance of several well-known objective image quality assessment metrics is evaluated against the MOS from the subjective study. These objective image quality metrics are commonly used and their implementations are publicly available on the Internet. The metrics are selected to represent different categories. The error sensitivity based metrics: peak signal to noise ratio (PSNR), information content weighted PSNR (IW-PSNR) [56], visual signal to noise ratio (VSNR) [55], and S-CIELAB [47]. The structural similarity based metrics: universal image quality index (UQI) [53], structural similarity index (SSIM) [57], multi-scale structural similarity index (MSSIM) [58], and information content weighted SSIM (IW-SSIM) [56]. The information fidelity based metrics: information fidelity criterion (IFC) [72], and visual information fidelity (VIF) [73]. The feature similarity based metrics: Riesz transform based feature similarity (RFSIM) index [74], and feature similarity (FSIMc) [75].

The scores given by an objective image quality metric are transferred into a predicted MOS in order to map the scores of the objective image quality metric into the range of the subjective MOS and to remove any nonlinearity between them using non-linear regression [84]. Fig. 3.7 shows the scatter plots of the scores given by the different objective image quality metrics versus the subjective

MOS before non-linear regression. The function chosen for regression is a four parameters logistic function [89]:

$$MOS_p(Q) = \frac{p_1 - p_2}{1 + \exp\left(\frac{Q - p_3}{p_4}\right)} + p_2 \tag{3.2}$$

where $MOS_p$ is the predicted MOS, $Q$ is the quality rating given by an objective image quality metric. The parameters $p_1, p_2, p_3$, and $p_4$ are chosen to minimize the mean square error between the quality scores given by the objective image quality metric and the subjective MOS.

Tables 3.1-3.4 show Pearsons correlation coefficient, root mean square error, Spearmans rank order correlation coefficient, and outlier ratio of the scores given by different quality assessment methods against MOS from the subjective study after logistic transformation for individual datasets as well as for the whole data. In general, the performance of error sensitivity based metrics is slightly better than feature similarity based metrics and among them S-CIELAB is the best performing metric. In the meantime, feature similarity based metrics are performing better than structural similarity based metrics. The information fidelity based metric, IFC, is second worst performing metric among all metrics while its improved version, VIF, is one of the best performing metrics.

The statistical significance of each metrics performance relative to other metrics was evaluated by performing an $F$-test in the set of residuals (prediction errors). For variances $\sigma_A^2$ and $\sigma_B^2$ of two sets of residuals from metrics $A$ and $B$ respectively; the $F$ statistic is defined as $F = \frac{\sigma_A^2}{\sigma_B^2}$. If $F > F_{critical}$ ($F < 1/F_{critical}$) then it signifies that at a given confidence level, metric $A$ has significantly larger (smaller) residuals than metric $B$. The $F_{critical}$ is computed based on the number of residuals

and the confidence level [55]. In our study, we used 95% confidence level. The $F$-test assumes that the set of residuals (prediction errors) are normally distributed. We used a simple normality test based on the rule of thumb that a set of values is normally distributed if its kurtosis and skewness values lie between 2 and 4, and -1 and 1 respectively [89] (Normal distribution has a kurtosis of 3 and a skewness of zero). The results of the normality test are given in Table 3.5.

Table 3.6 lists the $F$-test statistics results carried out on the set of residuals of each objective image quality metric for the individual subsets as well as for the full dataset. Each entry in the Table 3.6 is a codeword of six symbols. The position of the symbol in the code word represents the following datasets (from left to right): Dekker SOM, Median Cut, Kmeans, Octree, Wus algorithm, and all data. Each symbol gives the result of the $F$-test on the dataset represented by the symbols position. "1" means that the image quality metric from the row is statistically better than the image quality metric from the column, "0 means that it is statistically worse and "-" means that it is statistically indistinguishable. Thus in terms of statistical significance, as expected, S-CIELAB is statistically the best performing metric as it is a color reproduction metric followed by the PSNR. The results of the PSNR metric are qualitatively similar to those reported in [75] for the quality assessment of color quantized images. We can see also that VIF is significantly better than the structure similarity based metrics while IFC and UQI metrics are statistically the worst performing metrics.

Several interesting points came out from the results presented in this chapter. Error sensitivity and feature similarity based metrics are closer to human perception than structural similarity based metrics. Quantization with large number of levels (greater than 16) does not seem to alter significantly the large-scale struc-

ture of the image and therefore the perception of quality is dominated by the error in color information. Such errors are best captured by measures such as S-CIELAB and PSNR. On the other hand when the number of levels is small, structural errors are also present as the quantized image is dominated by spurious and missing edges. In such cases, structural similarity based metrics as well as other metrics are all correlate well with human perception.

Another observation is that structural similarity metrics correlate well with human perception when the original image shows a clearly defined, perceptually significant object such as a large flower, a parrot, or a close-up of a childs face.

Table 3.1: Pearson's correlation coefficient of the scores given by different quality assessment methods against MOS from the subjective study after a non-linear mapping.

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.956 | 0.965 | 0.960 | 0.970 | 0.957 | 0.945 |
| ***IWPSNR*** | 0.959 | 0.935 | 0.957 | 0.969 | 0.963 | 0.934 |
| ***S-CIELAB*** | 0.963 | 0.966 | 0.969 | 0.977 | 0.961 | 0.961 |
| ***VSNR*** | 0.949 | 0.929 | 0.943 | 0.955 | 0.953 | 0.926 |
| ***UQI*** | 0.732 | 0.772 | 0.662 | 0.804 | 0.720 | 0.728 |
| ***SSIM*** | 0.929 | 0.940 | 0.911 | 0.935 | 0.930 | 0.913 |
| ***MSSIM*** | 0.935 | 0.934 | 0.910 | 0.944 | 0.940 | 0.917 |
| ***IWSSIM*** | 0.910 | 0.897 | 0.868 | 0.930 | 0.913 | 0.888 |
| ***IFC*** | 0.806 | 0.783 | 0.791 | 0.869 | 0.812 | 0.805 |
| ***VIF*** | 0.950 | 0.938 | 0.951 | 0.967 | 0.957 | 0.942 |
| ***RFSIM*** | 0.957 | 0.955 | 0.954 | 0.964 | 0.963 | 0.946 |
| ***FSIMc*** | 0.943 | 0.928 | 0.924 | 0.958 | 0.949 | 0.924 |

Table 3.2: Root Mean Square Error

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 8.540 | 8.004 | 8.678 | 7.008 | 8.815 | 9.774 |
| ***IWPSNR*** | 8.281 | 10.860 | 9.029 | 7.216 | 8.234 | 10.715 |
| ***S-CIELAB*** | 7.885 | 7.920 | 7.679 | 6.158 | 8.400 | 8.312 |
| ***VSNR*** | 9.168 | 11.372 | 10.304 | 8.629 | 9.187 | 11.289 |
| ***UQI*** | 19.871 | 19.511 | 23.205 | 17.243 | 21.147 | 20.574 |
| ***SSIM*** | 10.831 | 10.490 | 12.787 | 10.302 | 11.176 | 12.209 |
| ***MSSIM*** | 10.334 | 10.970 | 12.862 | 9.532 | 10.390 | 11.933 |
| ***IWSSIM*** | 12.076 | 13.596 | 15.375 | 10.672 | 12.429 | 13.771 |
| ***IFC*** | 17.283 | 19.099 | 18.938 | 14.332 | 17.774 | 17.811 |
| ***VIF*** | 9.127 | 10.609 | 9.586 | 7.409 | 8.815 | 10.100 |
| ***RFSIM*** | 8.508 | 9.102 | 9.291 | 7.760 | 8.231 | 9.751 |
| ***FSIMc*** | 9.692 | 11.403 | 11.823 | 8.298 | 9.647 | 11.474 |

Table 3.3: Spearman's Rank Order Correlation Coefficient of the scores given by different quality assessment methods against MOS from the subjective study after a non-linear mapping.

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.950 | 0.961 | 0.952 | 0.965 | 0.953 | 0.939 |
| ***IWPSNR*** | 0.954 | 0.938 | 0.950 | 0.961 | 0.960 | 0.931 |
| ***S-CIELAB*** | 0.954 | 0.958 | 0.960 | 0.972 | 0.957 | 0.956 |
| ***VSNR*** | 0.944 | 0.925 | 0.938 | 0.954 | 0.952 | 0.923 |
| ***UQI*** | 0.742 | 0.774 | 0.679 | 0.802 | 0.728 | 0.735 |
| ***SSIM*** | 0.921 | 0.938 | 0.909 | 0.934 | 0.929 | 0.911 |
| ***MSSIM*** | 0.934 | 0.931 | 0.912 | 0.950 | 0.939 | 0.918 |
| ***IWSSIM*** | 0.911 | 0.897 | 0.872 | 0.935 | 0.913 | 0.890 |
| ***IFC*** | 0.812 | 0.790 | 0.798 | 0.873 | 0.816 | 0.810 |
| ***VIF*** | 0.945 | 0.936 | 0.946 | 0.962 | 0.955 | 0.938 |
| ***RFSIM*** | 0.950 | 0.950 | 0.948 | 0.960 | 0.959 | 0.941 |
| ***FSIMc*** | 0.942 | 0.930 | 0.928 | 0.959 | 0.949 | 0.926 |

Table 3.4: Outlier Ratio (the percentage of the number of predictions outside the range of two times the standard deviations of the subjective MOS [84])

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.206 | 0.143 | 0.189 | 0.286 | 0.206 | 0.280 |
| ***IWPSNR*** | 0.183 | 0.314 | 0.257 | 0.280 | 0.171 | 0.317 |
| ***S-CIELAB*** | 0.183 | 0.160 | 0.194 | 0.251 | 0.234 | 0.240 |
| ***VSNR*** | 0.229 | 0.337 | 0.240 | 0.349 | 0.194 | 0.335 |
| ***UQI*** | 0.589 | 0.606 | 0.686 | 0.617 | 0.611 | 0.647 |
| ***SSIM*** | 0.274 | 0.251 | 0.349 | 0.417 | 0.269 | 0.376 |
| ***MSSIM*** | 0.257 | 0.297 | 0.320 | 0.349 | 0.240 | 0.329 |
| ***IWSSIM*** | 0.303 | 0.394 | 0.383 | 0.377 | 0.326 | 0.383 |
| ***IFC*** | 0.469 | 0.577 | 0.537 | 0.520 | 0.497 | 0.542 |
| ***VIF*** | 0.217 | 0.240 | 0.223 | 0.280 | 0.194 | 0.282 |
| ***RFSIM*** | 0.217 | 0.246 | 0.234 | 0.314 | 0.194 | 0.280 |
| ***FSIMc*** | 0.200 | 0.326 | 0.337 | 0.331 | 0.229 | 0.328 |

Table 3.5: Normality test for the set residuals (Skewness/Kurtosis)

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | -0.08/4.14 | 0.25/4.0 | -0.13/4.16 | 0.02/3.6 | 0.48/4.3 | 0.13/3.6 |
| ***IWPSNR*** | -0.25/3.88 | 0.12/3.64 | -0.05/4.26 | 0.28/4.85 | 0.07/4.20 | 0.25/4.04 |
| ***S-CIELAB*** | 0.00/2.78 | 0.10/3.43 | 0.16/2.76 | 0.22/3.72 | 0.23/2.61 | 0.11/3.03 |
| ***VSNR*** | -0.17/3.50 | 0.41/3.9 | 0.14/4.7 | -0.20/4.50 | 0.17/4.9 | 0.28/3.9 |
| ***UQI*** | -0.21/2.71 | -0.16/3.05 | 0.10/2.7 | 0.39/3.4 | -0.13/3.02 | -0.01/3.01 |
| ***SSIM*** | -0.30/3.81 | -0.47/4.80 | -0.50/4.12 | -0.15/3.67 | -0.06/4.81 | -0.11/4.01 |
| ***MSSIM*** | -0.66/4.50 | -0.64/4.49 | -0.39/5.80 | 0.14/4.1 | -0.96/6.05 | -0.35/4.94 |
| ***IWSSIM*** | -0.60/4.48 | -0.32/4.18 | -0.01/5.42 | 0.50/5.21 | -0.78/5.44 | -0.20/4.96 |
| ***IFC*** | 0.01/3.3 | 0.40/3.5 | 0.05/3.6 | 0.63/4.3 | 0.01/3.7 | 0.12/3.6 |
| ***VIF*** | -0.37/4.02 | 0.16/4.3 | -0.43/4.55 | 0.94/5.7 | -0.22/4.73 | -0.04/4.47 |
| ***RFSIM*** | -0.50/3.37 | -0.03/3.53 | -0.04/4.66 | 0.11/4.15 | -0.35/3.75 | -0.02/3.92 |
| ***FSIMc*** | -0.33/3.65 | -0.13/4.10 | 0.07/5.31 | 0.17/4.78 | -0.30/4.11 | -0.09/4.33 |

Table 3.6: The $F$-test results for all datasets

| | ***PSNR*** | ***IWPSNR*** | ***S-CIELAB*** | ***VSNR*** | ***UQI*** | ***SSIM*** | ***MSSIM*** | ***IWSSIM*** | ***IFC*** | ***VIF*** | ***RFSIM*** | ***FSIMc*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***PSNR*** | - - - - - - | -1- - -1 | - - -0-0 | -111-1 | 111111 | 111111 | 111111 | 111111 | 111111 | -1- - - - | -1- - - - | 1111-1 |
| ***IWPSNR*** | -0- - -0 | - - - - - - | -000-0 | - -11- - | 111111 | 1-1111 | 1-1111 | 111111 | 111111 | - - - - -0 | -0- - -0 | 1-1111 |
| ***S-CIELAB*** | - - -1-1 | -111-1 | - - - - - - | 1111-1 | 111111 | 111111 | 111111 | 111111 | 111111 | 1111-1 | -111-1 | 111111 |
| ***VSNR*** | -000-0 | - -00- - | 0000-0 | - - - - - - | 111111 | 1-1111 | - -1- - - | 111111 | 111111 | - - -0-0 | -0- - -0 | - -1- - - |
| ***UQI*** | 000000 | 000000 | 000000 | 000000 | - - - - - - | 000000 | 000000 | 000000 | 0-0000 | 000000 | 000000 | 000000 |
| ***SSIM*** | 000000 | 0-0000 | 000000 | 0-0000 | 111111 | - - - - - - | - - - - - - | -11- -1 | 111111 | 0-0000 | 000000 | - - -000 |
| ***MSSIM*** | 000000 | 0-0000 | 000000 | - -0- - - | 111111 | - - - - - - | - - - - - - | 111-11 | 111111 | - -0000 | 000000 | - - -0- - |
| ***IWSSIM*** | 000000 | 000000 | 000000 | 000000 | 111111 | -00- -0 | 000-00 | - - - - - - | 111111 | 000000 | 000000 | 000000 |
| ***IFC*** | 000000 | 000000 | 000000 | 000000 | 1-1111 | 000000 | 000000 | 000000 | - - - - - - | 000000 | 000000 | 000000 |
| ***VIF*** | -0- - - - | - - - - -1 | 0000-0 | - - -1-1 | 111111 | 1-1111 | - -1111 | 111111 | 111111 | - - - - - - | -0- - - - | - -1- -1 |
| ***RFSIM*** | -0- - - - | -1- - -1 | -000-0 | -1- - -1 | 111111 | 111111 | 111111 | 111111 | 111111 | -1- - - - | - - - - - - | 111-11 |
| ***FSIMc*** | 0000-0 | 0-0000 | 000000 | - -0- - - | 111111 | - - -111 | - - -1- - | 111111 | 111111 | - -0- -0 | 000-00 | - - - - - - |

# CHAPTER 4

# Objective Image Quality Assessment

## 4.1 Introduction

Although the subjective quality assessment methods are accurate in estimating the perceived quality of images, but such routines are time-consuming, cumbersome, expensive to conduct, and the methodology is difficult to embed into real-world applications. What we really need are automatic algorithms for objective quality assessment that could analyze images and report their quality without human involvement. Such objective image quality assessment methods are computer-based methods that can predict the subjective image quality automatically without carrying out subjective experiments. However, automatic objective image quality assessment methods do not necessarily correlate well with perceived quality [90]. The objective image quality assessment methods can be employed for the following functions [57]:

1. It can be used to *monitor* image quality for quality control systems. For example, an image quality metric can be used to monitor an image acquisition system and automatically adjust its related parameters to obtain the best quality image data. A network video server can use an image quality metric to verify the quality of the digital video transmitted on the network to control and allocate streaming resources.

2. It can be used to *benchmark* image-processing systems and algorithms. For instance, if multiple of image denoising and restoration algorithms are available to enhance the quality of images captured using digital cameras, then a quality metric can be deployed to determine which of them provides the best quality results.

3. It also can be *embedded* into image-processing and transmission systems to optimize the systems and allow the systems to adjust its parameters autonomously. For example, in a visual communication system, an image quality measure can assist in the optimal design of the pre-filtering and bit assignment algorithms at the encoder and of optimal reconstruction, error concealment, and post-filtering algorithms at the decoder.

In this chapter, I propose to approach the problem of objective quality assessment of color quantized images. I present two quality assessment methods: the first is a general-purpose method for the quality assessment of color quantization distortions as well as other types of image distortion while the second method is an application-specific for the quality assessment of color quantization distortion only.

## 4.2 Structural Similarity Index for Color Images

It is known that the number of colors that can be distinguished by human eye is much more than that of the grayscale levels and more information is contained in color than what is contained in a grayscale image. Regardless of the advantages of the existing well-known objective image quality metrics, one of the common and major limitations of these metrics is that the assessment of a color image is always performed by assessing its luminance component. As a result, the color information in the image is largely ignored and precision of assessment result is influenced accordingly.

Wang et al. [58] introduced a general purpose grayscale image quality index. The quality index is computationally simple and correlates with the subjective evaluations of human observers for a wide variety of distortions. In this section, we show that this grayscale image metric can be extended to color version, by adding a color comparison to the criteria of the grayscale image quality index.

### 4.2.1 The Grayscale Image Quality Metric

Wang et al. [57] proposed the structural similarity index (SSIM) based on the assumption that the HVS is highly adapted to extract structural information from visual scenes. The structural similarity index consists of three local comparison functions namely luminance comparison, contrast comparison, and structure comparison. For a given original and test image signals $x$ and $y$ respectively, the luminance, contrast, and structure comparison functions are defined as follows:

$$l\left(x,y\right)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1} \tag{4.1}$$

$$C(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2} \tag{4.2}$$

$$S(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \tag{4.3}$$

where $\mu_x$ and $\mu_y$ are the sample means of $x$ and $y$ respectively, $\sigma_x$ and $\sigma_y$ are the sample standard deviations of $x$ and $y$ respectively, and $\sigma_{xy}$ is the sample correlation coefficient between $x$ and $y$. The constants $C_1$, $C_2$, and $C_3$ are used to stabilize the algorithm when the denominators approach zero. The structural similarity index is designed to satisfy the following reasonable conditions:

1. *Symmetry*: $SSIM(x,y) = SSIM(y,x)$. When quantifying the similarity between two signals, exchanging the order of the input signals should not affect the resulting measurement.

2. *Boundedness*: $SSIM(x,y) \leq 1$. An upper bound can serve as an indication of how close the two signals are to being perfectly identical.

3. *Unique maximum*: $SSIM(x,y) = 1 \iff x = y$. The perfect score is achieved only when the signals being compared are identical.

The general form of the SSIM index is given by combining the three comparison functions:

$$SSIM = l(x,y)^\alpha \cdot C(x,y)^\beta \cdot S(x,y)^\gamma \tag{4.4}$$

where $\alpha$, $\beta$, and $\gamma$ are parameters which define the relative importance of the three components. Usually, $\alpha = \beta = \gamma = 1$ and $C_1 = C_2$ yielding a specific form of the SSIM index:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \tag{4.5}$$

The local statistics $\mu_x$, $\mu_y$ and $\sigma_{xy}$ are computed within a local window which moves pixel-by-pixel over the entire image. At each step, the local statistics and SSIM index are calculated within the local window. Then the mean of SSIM values is taken to evaluate the overall image quality.

The single-scale SSIM was extended into a multiscale version (MSSIM) [58] which provides more flexibility than single-scale approach in incorporating the variations of image resolution and viewing conditions. In MSSIM, the contrast and the structure comparisons are calculated at each scale, and the luminance comparison is computed only at highest scale $M = 5$.

$$MSSIM = [l_M(x,y)]^{\alpha_M} \prod_{i=1}^{M} [C_i(x,y)]^{\beta_i} \cdot [S_i(x,y)]^{\gamma_i} \tag{4.6}$$

the parameters $\alpha_i$, $\beta_i$, and $\gamma_i$ are selected such that $\alpha_i = \beta_i = \gamma_i$ and $\sum_{i=1}^{M} \gamma_i = 1$ .

### 4.2.2 The Color Image Quality Metric

In this section, we extend the grayscale image quality metric MSSIM [58] to include color by adding a color comparison to the criteria of the grayscale image quality index MSSIM. The rationale for this approach is the fact that the color images reveal more meaningful information to the human observers rather than grayscale ones and that the number of colors that can be distinguished by human eye is

much more than that of the grayscale levels

The CIELAB color space is designed so that Euclidean distances between different colors in the CIELAB correspond approximately to perceived color differences. The three coordinates of the CIELAB color space, $L^*$, $a^*$, and $b^*$, represent the color luminance, the position between red and green, and the position between yellow and blue respectively. The XYZ coordinate system is used to transform RGB coordinates to $L^*a^*b^*$ coordinates using following equations:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{4.7}$$

$$L^* = \begin{cases} 116\left(Y/Y_n\right)^{1/3} - 16, & \text{if } Y/Y_n > 0.008856 \\ 903.3\left(Y/Y_n\right), & \text{if } Y/Y_n \leq 0.008856 \end{cases} \tag{4.8}$$

$$a^* = 500\left(f\left(X/X_n\right) - f\left(Y/Y_n\right)\right) \tag{4.9}$$

$$b^* = 200\left(f\left(Y/Y_n\right) - f\left(Z/Z_n\right)\right) \tag{4.10}$$

where

$$f\left(t\right) = \begin{cases} t^{1/3}, & \text{for } t > 0.008856 \\ 7.787 * t + 16/116, & \text{for } t \leq 0.008856 \end{cases} \tag{4.11}$$

Here $Y_n = 1.0$ is the luminance, and $X_n = 0.950455$, $Z_n = 1.088753$ are the chrominances for the $D_{65}$ white point.

A useful rule of thumb in CIELAB color space is that any two colors can be perceptually distinguishable from each other only if the Euclidean distance between these two colors is greater than a threshold value of 3 [91]. This threshold is known as the Just Noticeable Color Difference (JNCD) threshold. Therefore all the colors within a sphere of radius equal to the JNCD threshold are perceptually indistinguishable from each other.

The Color comparison can be obtained by the following steps:

- *Step 1*: The reference and test images are preprocessed by spatial filtering to simulate the spatial blurring by the human visual system in a way that the filtering operation to the image affects only the fine patterned colors [47].
- *Step 2*: The reference and test images are transformed to the CIELAB color space.
- *Step 3*: The color comparison is computed by averaging the number of colors in the reference image that are indistinguishable from their corresponding colors in the test image based on the CIELAB color space JNCD threshold value 3.

The general form of the proposed image quality measure is given as:

$$CMSSIM = [Clr(x,y)]^{\delta}[l_M(x,y)]^{\alpha_M}\prod_{i=1}^{M}[C_i(x,y)]^{\beta_i}\cdot[S_i(x,y)]^{\gamma_i} \tag{4.12}$$

where the contrast comparison $C$ and the structure comparison $S$ are calculated at each scale, the luminance comparison $l$ is computed only at highest scale $M = 5$, and the color comparison $Clr$ is calculated at the lowest scale only. The parameters $\alpha_i$, $\beta_i$, and $\gamma_i$ are selected as in [58], while the optimal value of $\delta$ is 0.7.

### 4.2.3 Results and Discussions

In this section, I present results on validation of the proposed CMSSIM quality metric on the database presented in Section 3.3. The database contains 875 test images obtained from 25 reference images using five different color quantization algorithms. Each image was evaluated by 22 human subjects and more than 19200 individual human quality judgments were carried out to give the final Mean Opinion Scores (MOS).

The performance of the proposed image quality metric in terms of the ability of predicting the subjective ratings is analyzed and compared with the performance of other quality assessment metrics. Specifically, I will compare the performance of the proposed metric against peak signal to noise ratio (PSNR), information content weighted PSNR (IW-PSNR) [56], visual signal to noise ratio (VSNR) [55], visual information fidelity (VIF) [73], structural similarity index (SSIM) [57], multiscale structural similarity index (MSSIM) [58], information content weighted SSIM (IW-SSIM) [56], Riesz transform based feature similarity (RFSIM) index [74], and feature similarity (FSIMc) [75].

A procedure similar to the one described in Section 3.4 was used to transfer scores given by an objective image quality metric into a predicted MOS. The purpose of this procedure is to map the scores of the objective image quality metric into the range of the subjective MOS and to remove any nonlinearity between them using non-linear regression [84].

Table 4.1 shows the Pearson's linear correlation coefficients (PLCC), the Spearman's rank order correlation coefficient (SROCC), and the Root Mean Square Error (RMSE) of the scores given by different quality assessment metrics against

Table 4.1: Validation scores for different quality assessment methods. The methods were tested against subjective MOS from the subjective study after a non-linear mapping. The validation criteria are: Pearson's linear correlation coefficients (PLCC), the Spearman's rank order correlation coefficient (SROCC), and the Root Mean Square Error (RMSE).

| | ***PLCC*** | ***SROCC*** | ***RMSE*** |
|---|---|---|---|
| ***PSNR*** | 0.945 | 0.939 | 9.774 |
| ***IWPSNR*** | 0.934 | 0.931 | 10.715 |
| ***VSNR*** | 0.926 | 0.923 | 11.289 |
| ***VIF*** | 0.942 | 0.938 | 10.100 |
| ***SSIM*** | 0.913 | 0.911 | 12.209 |
| ***MSSIM*** | 0.917 | 0.918 | 11.933 |
| ***IWSSIM*** | 0.888 | 0.890 | 13.771 |
| ***RFSIM*** | 0.946 | 0.941 | 9.751 |
| ***FSIMc*** | 0.924 | 0.926 | 11.474 |
| ***CMSSIM*** | **0.960** | **0.954** | **8.421** |

MOS from the subjective study after a non-linear mapping. It is clear from that table that the proposed CMSSIM metric outperforms state-of-the-art image quality metrics in assessing the quality of color quantized images in terms of all validation criteria. We can also see how the proposed CMSSIM metric demonstrably improved the performance of the structural similarity-based metrics as it includes color comparison.

We also used the popular Tampere Image Database (TID2008) [92] to further test the performance of the proposed CMSSIM metric for different distortion types. This database is the most recent and largest database, in terms of number of images and distortion types, so far available for verification of full reference quality metrics. The TID2008 database contains 1700 distorted images (25 reference images $\times$ 17 types of distortions $\times$ 4 levels of distortions). Mean Opinion Scores for this database have been obtained as a result of 838 subjective experiments.

Table 4.2 shows the Pearson's linear correlation coefficients (PLCC) and the

Table 4.2: The Pearson's linear correlation coefficients (PLCC) and the Spearman's rank order correlation coefficient (SROCC) for the TID2008 image database

| | ***PLCC*** | | | ***SROCC*** | | |
|---|---|---|---|---|---|---|
| | ***MSSIM*** | ***CMSSIM*** | ***Δ(%)*** | ***MSSIM*** | ***CMSSIM*** | ***Δ(%)*** |
| ***Additive Gaussian noise*** | 0.748 | **0.919** | 22.872 | 0.81 | **0.917** | 13.223 |
| ***Additive noise in color components*** | 0.778 | **0.926** | 19.020 | 0.806 | **0.916** | 13.662 |
| ***Spatially correlated noise*** | 0.76 | **0.863** | 13.575 | 0.819 | **0.869** | 6.110 |
| ***Masked noise*** | **0.787** | 0.751 | -4.624 | **0.815** | 0.779 | -4.472 |
| ***High frequency noise*** | 0.822 | **0.957** | 16.439 | 0.868 | **0.939** | 8.104 |
| ***Impulse noise*** | 0.625 | **0.874** | 39.846 | 0.687 | **0.902** | 31.366 |
| ***Quantization noise*** | 0.757 | **0.853** | 12.728 | 0.**854** | 0.849 | -0.613 |
| ***Gaussian blur*** | 0.877 | **0.902** | 2.743 | **0.961** | 0.913 | -4.928 |
| ***Image denoising*** | 0.915 | **0.958** | 4.714 | **0.957** | 0.950 | -0.737 |
| ***JPEG compression*** | 0.931 | **0.943** | 1.248 | 0.935 | **0.936** | 0.064 |
| ***JPEG2000 compression*** | 0.939 | **0.944** | 0.529 | **0.973** | 0.945 | -2.869 |
| ***JPEG transmission errors*** | 0.824 | **0.864** | 4.792 | **0.874** | 0.864 | -1.097 |
| ***JPEG2000 transmission errors*** | 0.788 | **0.830** | 5.436 | 0.853 | **0.859** | 0.756 |
| ***Non eccentricity pattern noise*** | 0.665 | **0.724** | 8.937 | 0.733 | **0.763** | 4.043 |
| ***Local block-wise distortions*** | 0.796 | **0.845** | 6.088 | 0.761 | **0.810** | 6.433 |

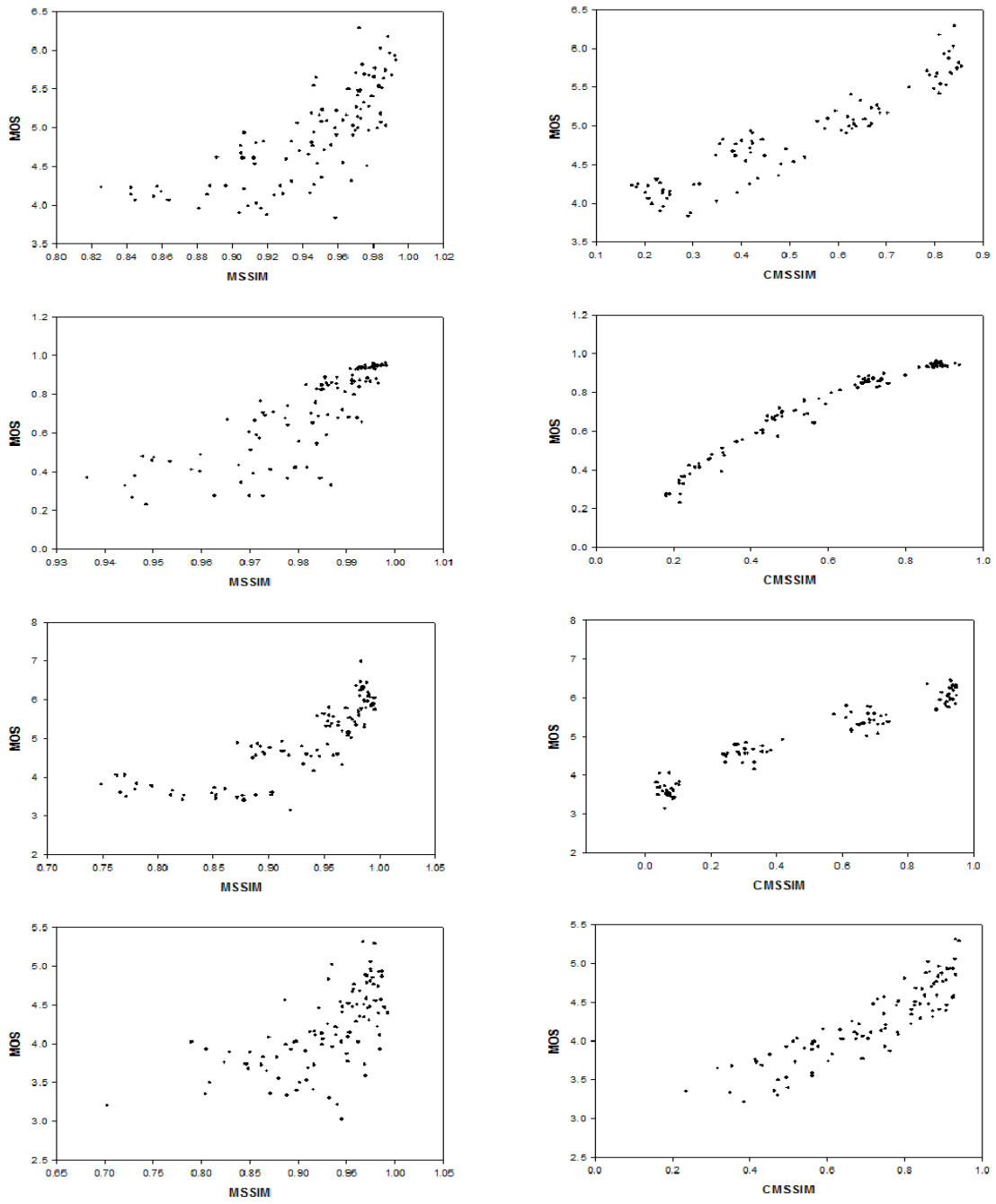



Figure 4.1: Scatter plots of MSSIM ($1^{st}$ column) and the proposed CMSSIM ($2^{nd}$ column) image quality metrics for a subset of distortion types from TID2008 image database. The distortion types are: Additive Gaussian noise ($1^{st}$ row), Additive noise in color components ($2^{nd}$ row), High frequency noise ($3^{rd}$ row), and Impulse noise ($4^{th}$ row).

Spearman's rank order correlation coefficients (SROCC) between the MOS form the TID2008 image database and the scores given by proposed CMSSIM as well as for the state-of-the-art MSSIM index. It is clear that the proposed CMSSIM metric greatly improves the performance of MSSIM for many distortion types. This improvement reaches around 40% in impulse noise and 23% in additive Gaussian noise.

It may be noted for the quantization noise that our extension has improved the performance of the original MSSIM metric about 12.7% for the Pearson's linear correlation coefficients criterion, but the two metrics have a competitive performance for the Spearman's rank order correlation coefficient. This is because there are only one hundred quantized images produced by single quantization algorithm in the TID2008 database while in our subjective study there are 875 quantized images produced by five color quantization algorithms. This number of quantized images is large enough to distinguish between the performance of the two metrics. Fig. 4.1 depicts scatter plots of the scores given by the proposed CMSSIM as well as MSSIM for some distortion types of TID2008 database. It also shows the improvement in the performance of CMSSIM metric over the MSSIM.

## 4.3 An Objective Metric for Quality Assessment of Color Quantized Images

In color quantization, a true color image is transformed into a color-mapped image consisting of a small number carefully selected representative colors. The ultimate goal of color quantization is to select the representative colors to ensure minimization of the perceived difference between the original and the quantized images. The performance of color quantization techniques greatly depends on how these representative colors are chosen and how close are the representative colors to original colors.

Certain types of degradation take place in the quantized image due to the use of limited number of representative colors. The appearance of spurious edges is one of the most annoying of these degradations. Another degradation is the loss of structure in the quantized image. Although color quantization is not in the first place a structural distortion, reducing the number of colors in an image may result in disturbing of the structure of the quantized image. For instance, two adjacent regions having close colors may be merged into one region of single color.

In this section, a new application-specific image quality metric is proposed for the quality assessment of color quantized images. The proposed metric is defined as a combination of three factors: how close is the set of representative colors to the original set of colors, the presence of spurious edges, and the loss of image structure. Therefore the color quantization quality metric models any color image quantization distortion as a combination of three similarities: color similarity, edge similarity, and structure similarity as shown in Fig. 4.2. For two images $x$ and $y$,

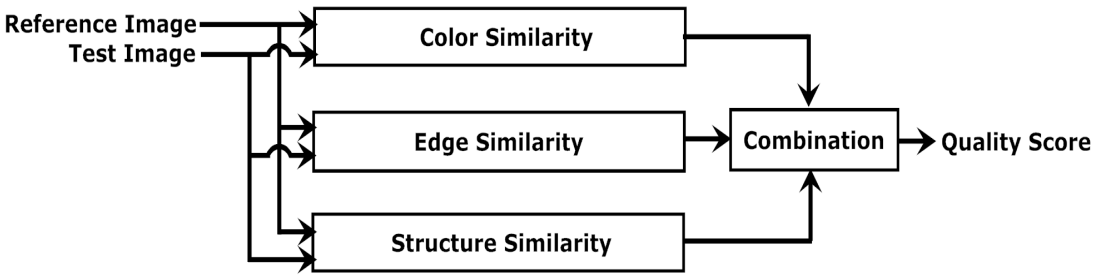



Figure 4.2: System diagram of the proposed color quantization quality metric.

the proposed quality metric is defined as

$$CQM(x,y) = C^{\alpha}(x,y) \cdot E^{\beta}(x,y) \cdot S^{\gamma}(x,y) \tag{4.13}$$

where $C$ is the color similarity, $E$ is the edge similarity, and $S$ is the structure similarity. $\alpha$, $\beta$ and $\gamma$ are parameters that define the relative importance of the three similarities. The optimal values of the parameters are $\alpha = 0.5$, $\beta = 0.7$ and $\gamma = 0.7$. The three statistical features are measured locally by moving a 3×3 pixels window through all the image pixels starting from the top left pixel to the bottom right one.

### 4.3.1 Color Similarity

The color similarity here is similar to the one described in Section 4.2.2 which is based on the Just Noticeable Color Difference (JNCD) threshold in the CIELAB color space. The Color similarity can be obtained by the following steps:

- *Step 1*: The reference and test images are preprocessed by spatial filtering to simulate the spatial blurring by the human visual system in a way that the filtering operation to the image affects only the fine patterned colors [47].

- *Step 2*: The reference and test images are transformed to the CIELAB color space.

- *Step 3*: The color similarity map is created by averaging the number of colors in the reference image that are indistinguishable from their corresponding colors in the test image based on the CIELAB color space JNCD threshold value 3.

Fig. 4.3 shows the color similarity maps of a full color image against some of its color quantized versions. The white regions contain perceptually indistinguishable colors.

### 4.3.2 Edge Similarity

We used the difference vector (DV) edge detection operator [93], which is color based edge detector, to create the edge maps of the reference and quantized images. In DV operator, each pixel represents a vector in the RGB color space and the maximum gradient is calculated across the central pixel in each of the four possible directions: 0, 45, 90, and 135 *deg.* The gradients are defined as:

$$|\nabla f|_{0\,\mathrm{deg}} = \|U_{0\,\mathrm{deg}} - V_{0\,\mathrm{deg}}\| \tag{4.14}$$

$$|\nabla f|_{90\,\mathrm{deg}} = \|U_{90\,\mathrm{deg}} - V_{90\,\mathrm{deg}}\| \tag{4.15}$$

$$|\nabla f|_{45\,\mathrm{deg}} = \|U_{45\,\mathrm{deg}} - V_{45\,\mathrm{deg}}\| \tag{4.16}$$

$$|\nabla f|_{135\,\mathrm{deg}} = \|U_{135\,\mathrm{deg}} - V_{135\,\mathrm{deg}}\| \tag{4.17}$$

where $\|\cdot\|$ denotes the $L_2$ norm, $U$ and $V$ are RGB color vectors in the neighborhood of the central pixel. Then the difference vector is given by

$$DV = \max(|\nabla f|_{0\,\mathrm{deg}}, |\nabla f|_{45\,\mathrm{deg}}, |\nabla f|_{90\,\mathrm{deg}}, |\nabla f|_{135\,\mathrm{deg}})$$

A threshold can be applied to the maximum gradient vector to locate edges. The edge similarity of the reference and test images $x$ and $y$ is then given as:

$$E(x,y) = \frac{2E_xE_y + C_1}{E_x^2 + E_y^2 + C_1} \tag{4.18}$$

where $E_x$ and $E_y$ are the edge responses of applying the DV edge detector to the reference and test images $x$ and $y$ respectively. $C_1$ is a stabilizing constant to avoid division by zero. Fig. 4.4 shows the edge similarity maps of a full color image against some of its color quantized versions.

### 4.3.3 Structure Similarity

The structural information in an image was defined as those attributes that represent the structure of objects in the scene, independent of the average luminance and contrast. Therefore, the structure of an image is determined after removing the luminance and contrast normalization [57].

Suppose that $x$ and $y$ are two nonnegative image signals with one of the signals having perfect quality. First, the luminance of each signal is estimated as the mean intensity:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4.19}$$

Second, the mean intensity is removed from the signals. The standard deviation (the square root of variance) is used as an estimate of the signal contrast. An unbiased estimate in discrete form is given by:

$$\sigma_x = \left( \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)^2 \right)^{1/2} \tag{4.20}$$

Third, the signal is normalized (divided) by its own standard deviation; so that the two signals being compared have unit standard deviation. The structure similarity is computed on these normalized signals $(x - \mu_x)/\sigma_x$ and $(y - \mu_y)/\sigma_y$. The correlation (inner product) between these is a simple and effective measure to quantify the structural similarity. Thus, the structure similarity is defined as follows [57]:

$$S(x, y) = \frac{\sigma_{xy} + C_2}{\sigma_x \sigma_y + C_2} \tag{4.21}$$

where $\sigma_x$ and $\sigma_y$ are the sample standard deviations of $x$ and $y$ respectively, $\sigma_{xy}$ is the sample correlation coefficient between $x$ and $y$, and $C_2$ is a stabilizing constant. Fig. 4.5 shows the structure similarity maps of a full color image against some of its color quantized versions.

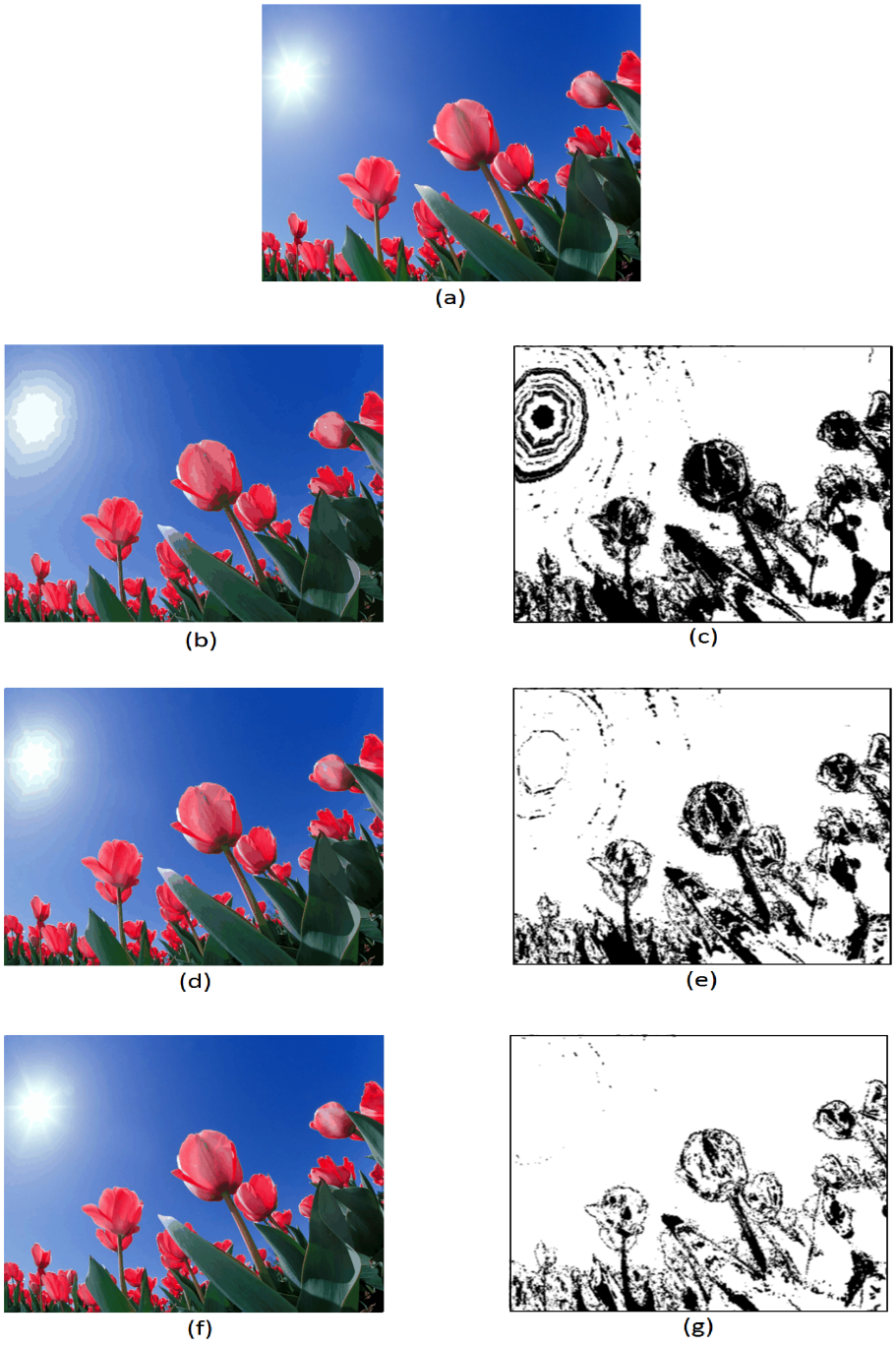



Figure 4.3: Color similarity maps. (a) Full color image, (b) 64 colors quantization of (a), (c) Color similarity map of (a) and (b), (d) 128 colors quantization of (a), (e) Color similarity map of (a) and (d), (f) 256 colors quantization of (a), (g) Color similarity map of (a) and (f).

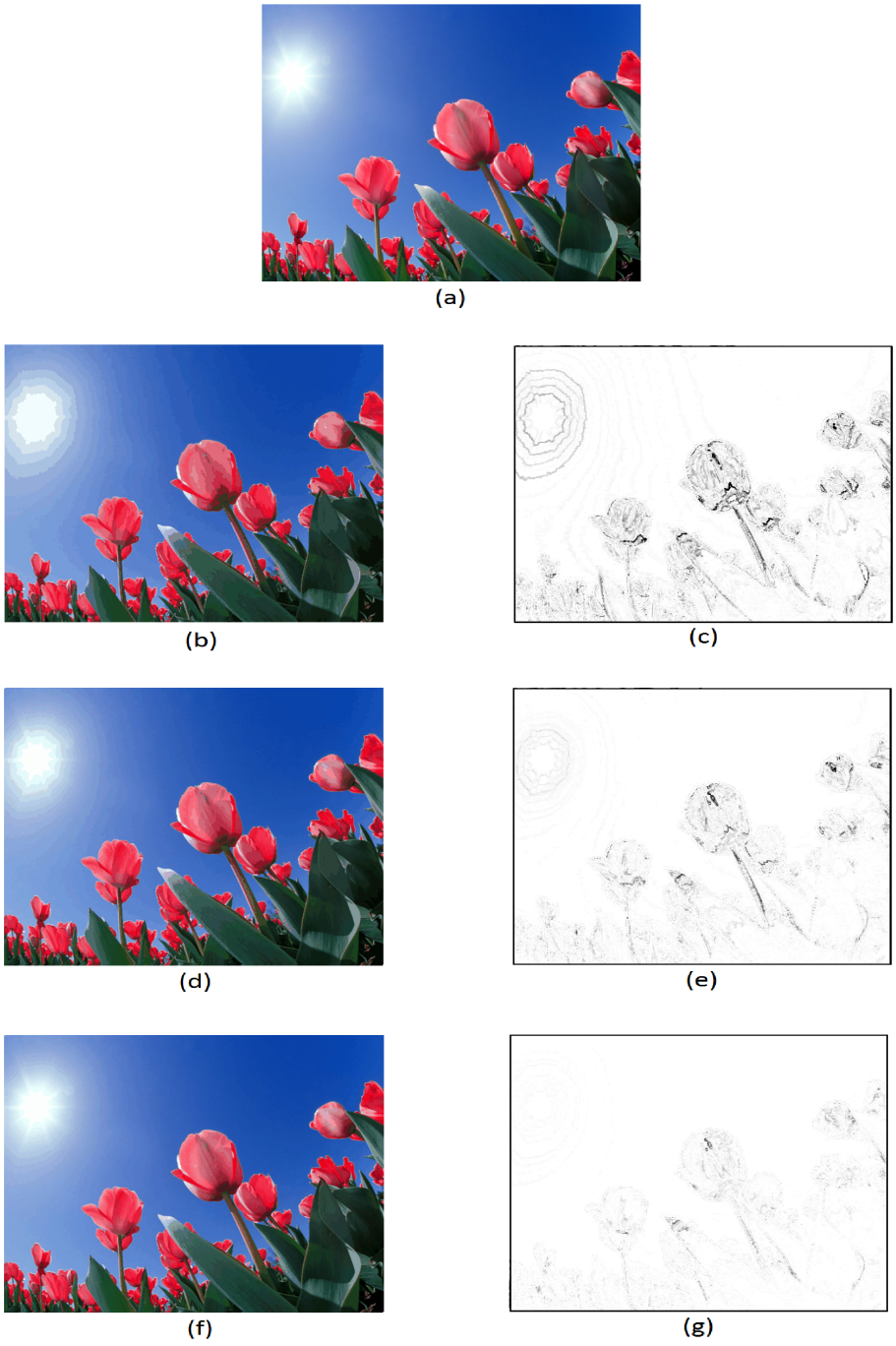



Figure 4.4: Edge similarity maps. (a) Full color image, (b) 64 colors quantization of (a), (c) Edge similarity map of (a) and (b), (d) 128 colors quantization of (a), (e) Edge similarity map of (a) and (d), (f) 256 colors quantization of (a), (g) Edge similarity map of (a) and (f).

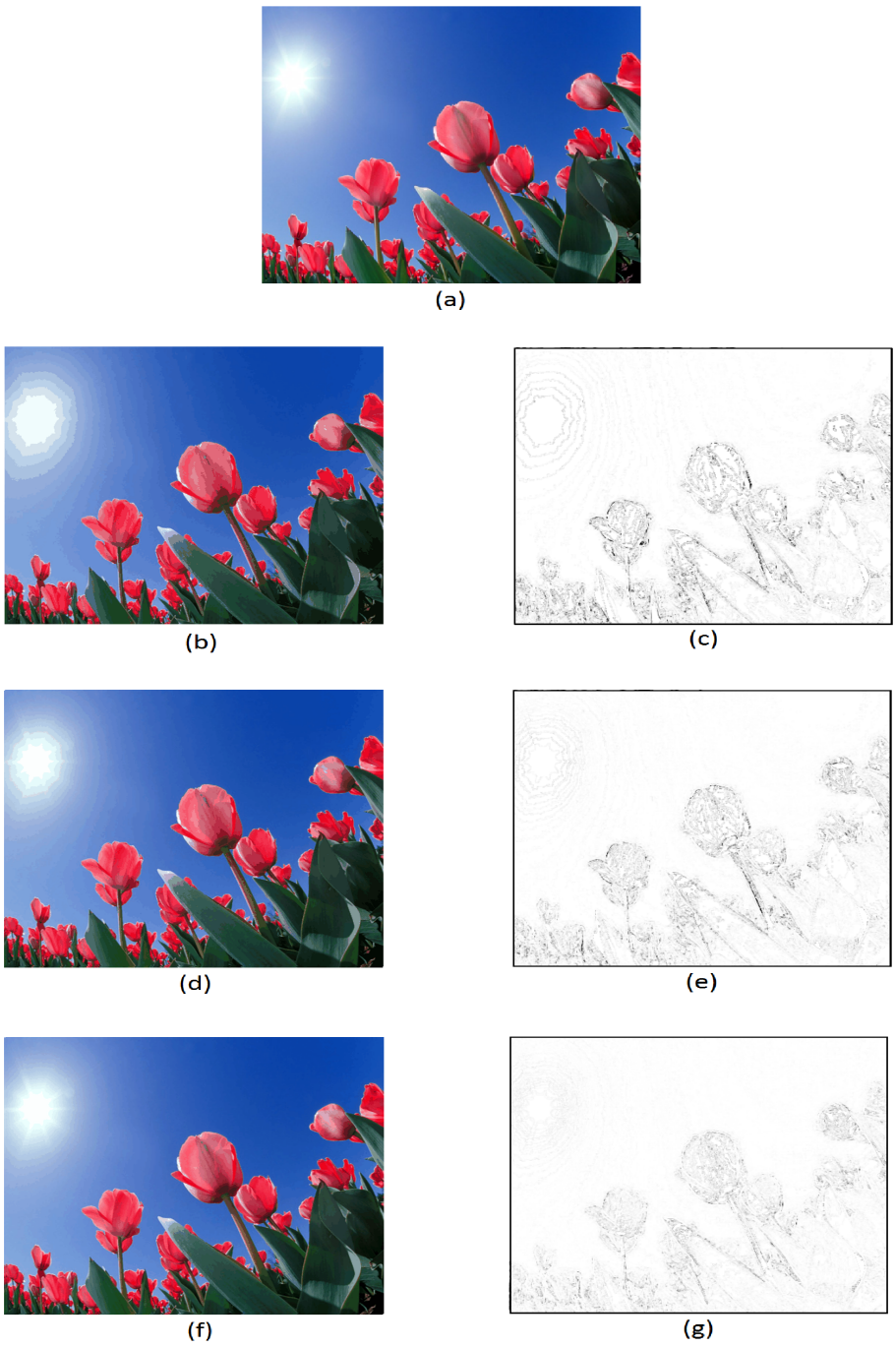



Figure 4.5: Structure similarity maps. (a) Full color image, (b) 64 colors quantization of (a), (c) Structure similarity map of (a) and (b), (d) 128 colors quantization of (a), (e) Structure similarity map of (a) and (d), (f) 256 colors quantization of (a), (g) Structure similarity map of (a) and (f).

### 4.3.4 Results and Discussions

In this section, the performance of the proposed image quality metric in terms of the ability of predicting the subjective ratings is analyzed. The proposed quality metric was applied to the set of images from the database presented in Section 3.3, and the results were compared to the subjective MOS. For comparison, the same set of images were presented to twelve well-known objective image quality metrics that are commonly used and their implementations are publicly available on the Internet. These metrics are selected to represent different categories. The error sensitivity based metrics: peak signal to noise ratio (PSNR), information content weighted PSNR (IW-PSNR) [56], visual signal to noise ratio (VSNR) [55], and S-CIELAB [47]. The structural similarity based metrics: universal image quality index (UQI) [53], structural similarity index (SSIM) [57], multiscale structural similarity index (MSSIM) [58], and information content weighted SSIM (IW-SSIM) [56]. The information fidelity based metrics: information fidelity criterion (IFC) [72], and visual information fidelity (VIF) [73]. The feature similarity based metrics: Riesz transform based feature similarity (RFSIM) index [74], and feature similarity (FSIMc) [75].

A mapping procedure similar to the one described in Section 3.4 was used to transfer scores given by an objective image quality metric into the range of the subjective MOS and to remove any nonlinearity between them.

Tables 4.3-4.6 show Pearson's correlation coefficient, Spearman's rank order correlation coefficient, root mean square error, and outliers ratio of the scores given by the different image quality assessment metrics included in the comparison after nonlinear mapping. It is clear from those tables that proposed metric CQM

outperforms all other metrics by a sizeable margin using any of the validation criteria. However, error sensitivity based metrics still outperform other metrics and among them the S-CIELAB metric is the best as it is color reproduction metric.

Table 4.3: Pearson's correlation coefficient of the scores given by different quality assessment methods against MOS from the subjective study after a non-linear mapping.

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.956 | 0.965 | 0.960 | 0.970 | 0.957 | 0.945 |
| ***IWPSNR*** | 0.959 | 0.935 | 0.957 | 0.969 | 0.963 | 0.934 |
| ***S-CIELAB*** | 0.963 | 0.966 | 0.969 | 0.977 | 0.961 | 0.961 |
| ***VSNR*** | 0.949 | 0.929 | 0.943 | 0.955 | 0.953 | 0.926 |
| ***UQI*** | 0.732 | 0.772 | 0.662 | 0.804 | 0.720 | 0.728 |
| ***SSIM*** | 0.929 | 0.940 | 0.911 | 0.935 | 0.930 | 0.913 |
| ***MSSIM*** | 0.935 | 0.934 | 0.910 | 0.944 | 0.940 | 0.917 |
| ***IWSSIM*** | 0.910 | 0.897 | 0.868 | 0.930 | 0.913 | 0.888 |
| ***IFC*** | 0.806 | 0.783 | 0.791 | 0.869 | 0.812 | 0.805 |
| ***VIF*** | 0.950 | 0.938 | 0.951 | 0.967 | 0.957 | 0.942 |
| ***RFSIM*** | 0.957 | 0.955 | 0.954 | 0.964 | 0.963 | 0.946 |
| ***FSIMc*** | 0.943 | 0.928 | 0.924 | 0.958 | 0.949 | 0.924 |
| ***CQM*** | **0.967** | **0.969** | **0.971** | **0.981** | **0.964** | **0.966** |

Table 4.4: Spearman's Rank Order Correlation Coefficient of the scores given by different quality assessment methods against MOS from the subjective study after a non-linear mapping.

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.950 | 0.961 | 0.952 | 0.965 | 0.953 | 0.939 |
| ***IWPSNR*** | 0.954 | 0.938 | 0.950 | 0.961 | 0.960 | 0.931 |
| ***S-CIELAB*** | 0.954 | 0.958 | 0.960 | 0.972 | 0.957 | 0.956 |
| ***VSNR*** | 0.944 | 0.925 | 0.938 | 0.954 | 0.952 | 0.923 |
| ***UQI*** | 0.742 | 0.774 | 0.679 | 0.802 | 0.728 | 0.735 |
| ***SSIM*** | 0.921 | 0.938 | 0.909 | 0.934 | 0.929 | 0.911 |
| ***MSSIM*** | 0.934 | 0.931 | 0.912 | 0.950 | 0.939 | 0.918 |
| ***IWSSIM*** | 0.911 | 0.897 | 0.872 | 0.935 | 0.913 | 0.890 |
| ***IFC*** | 0.812 | 0.790 | 0.798 | 0.873 | 0.816 | 0.810 |
| ***VIF*** | 0.945 | 0.936 | 0.946 | 0.962 | 0.955 | 0.938 |
| ***RFSIM*** | 0.950 | 0.950 | 0.948 | 0.960 | 0.959 | 0.941 |
| ***FSIMc*** | 0.942 | 0.930 | 0.928 | 0.959 | 0.949 | 0.926 |
| ***CQM*** | **0.962** | **0.962** | **0.961** | **0.980** | **0.960** | **0.960** |

Table 4.5: Root Mean Square Error

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 8.540 | 8.004 | 8.678 | 7.008 | 8.815 | 9.774 |
| ***IWPSNR*** | 8.281 | 10.860 | 9.029 | 7.216 | 8.234 | 10.715 |
| ***S-CIELAB*** | 7.885 | 7.920 | 7.679 | 6.158 | 8.400 | 8.312 |
| ***VSNR*** | 9.168 | 11.372 | 10.304 | 8.629 | 9.187 | 11.289 |
| ***UQI*** | 19.871 | 19.511 | 23.205 | 17.243 | 21.147 | 20.574 |
| ***SSIM*** | 10.831 | 10.490 | 12.787 | 10.302 | 11.176 | 12.209 |
| ***MSSIM*** | 10.334 | 10.970 | 12.862 | 9.532 | 10.390 | 11.933 |
| ***IWSSIM*** | 12.076 | 13.596 | 15.375 | 10.672 | 12.429 | 13.771 |
| ***IFC*** | 17.283 | 19.099 | 18.938 | 14.332 | 17.774 | 17.811 |
| ***VIF*** | 9.127 | 10.609 | 9.586 | 7.409 | 8.815 | 10.100 |
| ***RFSIM*** | 8.508 | 9.102 | 9.291 | 7.760 | 8.231 | 9.751 |
| ***FSIMc*** | 9.692 | 11.403 | 11.823 | 8.298 | 9.647 | 11.474 |
| ***CQM*** | **7.385** | **7.592** | **7.423** | **5.579** | **8.087** | **7.716** |

Table 4.6: Outliers Ratio (an outlier is a point that falls far twice or more the subjective MOS standard deviation from the subjective MOS [84])

| | ***SOM*** | ***Median*** | ***Kmeans*** | ***Octree*** | ***Wu*** | ***All Data*** |
|---|---|---|---|---|---|---|
| ***PSNR*** | 0.206 | **0.143** | 0.189 | 0.286 | 0.206 | 0.280 |
| ***IWPSNR*** | 0.183 | 0.314 | 0.257 | 0.280 | **0.171** | 0.317 |
| ***S-CIELAB*** | 0.183 | 0.160 | 0.194 | 0.251 | 0.234 | 0.240 |
| ***VSNR*** | 0.229 | 0.337 | 0.240 | 0.349 | 0.194 | 0.335 |
| ***UQI*** | 0.589 | 0.606 | 0.686 | 0.617 | 0.611 | 0.647 |
| ***SSIM*** | 0.274 | 0.251 | 0.349 | 0.417 | 0.269 | 0.376 |
| ***MSSIM*** | 0.257 | 0.297 | 0.320 | 0.349 | 0.240 | 0.329 |
| ***IWSSIM*** | 0.303 | 0.394 | 0.383 | 0.377 | 0.326 | 0.383 |
| ***IFC*** | 0.469 | 0.577 | 0.537 | 0.520 | 0.497 | 0.542 |
| ***VIF*** | 0.217 | 0.240 | 0.223 | 0.280 | 0.194 | 0.282 |
| ***RFSIM*** | 0.217 | 0.246 | 0.234 | 0.314 | 0.194 | 0.280 |
| ***FSIMc*** | 0.200 | 0.326 | 0.337 | 0.331 | 0.229 | 0.328 |
| ***CQM*** | **0.166** | 0.206 | **0.189** | **0.211** | 0.223 | **0.230** |

# CHAPTER 5

# Joint Color Quantization and Dithering Techniques

## 5.1 Introduction

Color quantization is an important problem for many applications in graphics and multimedia where only a limited number of colors can be displayed or printed simultaneously. Regardless of color quantization algorithm used, reconstruction of an image with a limited number of representative colors (color palette) will cause highly visible degradations in image quality. The most disturbing of these degradations is the appearance of false contours. False contours appear when batches of colors in a smooth gradient area are mapped into a constant palette colors forming a flat regions; boundaries between such regions may be visible as false contours as shown in Fig. 5.1.

A common way to overcome the problem of false contours is the use of dithering techniques. Dithering techniques make use of the averaging property of the human eye to the colors in a neighborhood of the point of interest and create the illusion of more colors. A dithering technique called error diffusion achieves this effect by distributing the error encountered in quantizing a pixel to the neighboring pixels. This results in an alternation of palette colors in the neighborhood and is perceived as a new color by the eye. Some well-known error diffusion filters are

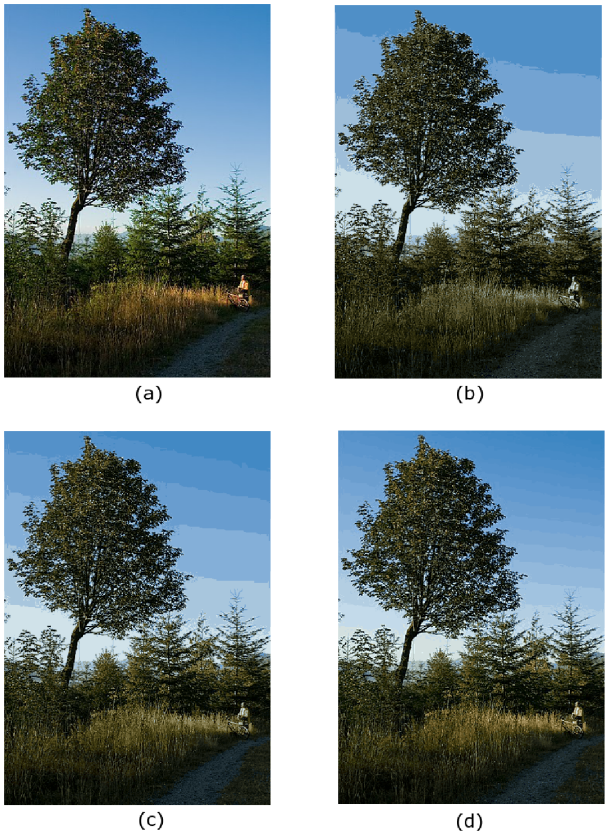



Figure 5.1: Appearance of false contours in color quantized images. (a) Full color image, (b) Color quantized image with 16 colors, (c) Color quantized image with 32 colors, (d) Color quantized image with 64 colors.

Floyd-Steinberg [94], Stucki [95], and Jarvice-Judice-Ninke [96].

Usually color quantization and dithering are performed sequentially; the quality of the dithering process depending on the selection of the color palette. Standard methods of color quantization do not take dithering into account. Therefore, the desirable qualities obtained by quantization are often disturbed by the dithering process. In this Chapter, two methods for color quantization are designed to obtain better results with minimum visual degradation after the error diffusion technique is applied.

## 5.2 Optimal Color Palette for Error Diffusion Techniques

In this section, I present a method for quantization of color images to be used with error diffusion techniques. This method is designed to obtain better results with minimum visual degradation after the error diffusion technique is applied. It consists of three steps: in the first step, the colors in original image are quantized into a number of colors that is much smaller than the desired number of colors. In the second step, the remaining palette colors are selected around each palette color generated in the first step based on the CIELAB color space Just Noticeable Color Difference (JNCD) threshold, and also on the size of the clusters. Finally error diffusion technique is incorporated within the process of pixel mapping.

### 5.2.1 Choice of Color Space

The CIE $L^*a^*b^*$ (CIELAB) and $L^*u^*v^*$ (CIELUV) color spaces [97] are considered perceptually uniform and referred to as uniform color spaces in which the Euclidean distances between different colors in the color space correspond approximately to perceived color differences. The application of uniform color spaces to the problem of color quantization has been studied by Gentile et al. [23] and to half-toning [98], the results indicate that such color spaces can significantly improve quality of the quantized image relative to the standard RGB color space. The CIE recommended using XYZ coordinate system to transform RGB to $L^*a^*b^*$. The following equations are used to transfer RGB to XYZ.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{5.1}$$

And then from XYZ to CIELAB color space

$$L^* = \begin{cases} 116\left(Y/Y_n\right)^{1/3} - 16, & \text{if } Y/Y_n > 0.008856 \\ 903.3\left(Y/Y_n\right), & \text{if } Y/Y_n \leq 0.008856 \end{cases} \tag{5.2}$$

$$a^* = 500\left(f\left(X/X_n\right) - f\left(Y/Y_n\right)\right) \tag{5.3}$$

$$b^* = 200\left(f\left(Y/Y_n\right) - f\left(Z/Z_n\right)\right) \tag{5.4}$$

where

$$f\left(t\right) = \begin{cases} t^{1/3}, & \text{for } t > 0.008856 \\ 7.787 * t + 16/116, & \text{for } t \leq 0.008856 \end{cases} \tag{5.5}$$

Here $Y_n = 1.0$ is the luminance, and $X_n = 0.950455$, $Z_n = 1.088753$ are the chrominances for the $D_{65}$ white point.

Two colors in CIELAB color space are perceptually distinguishable from each other if the Euclidean distance between these two colors is greater than a threshold value 3 [91]. This threshold is known as the Just Noticeable Color Difference (JNCD) threshold.

### 5.2.2 The Proposed Color Quantization Method

After transforming the input image into the CIELAB color space, the quantization process is composed of the following steps:

- ***Step 1:*** *Generation of Initial Palette Colors*

  In this step, the colors in the original image are quantized into a small set of $M$ colors that is much less than the desired $K$ palette colors ($M = 2^{\lfloor \frac{\log_2 K}{2} \rfloor}$) using k-means algorithm or any other quantization algorithm.

- ***Step 2:*** *Selection of the Remaining Palette Colors*

  After the generation of the initial palette colors, the remaining $(K - M)$ palette colors are selected around the initial palette color based on CIELAB JNCD threshold value 3. The use of CIELAB JNCD threshold is to ensure that the selected palette colors are far from each other by a distance that is greater than the CIELAB JNCD threshold to be distinguishable from each other. The number of selected palette colors around a given initial palette color is proportional to the size of cluster of image colors represented by that initial palette color.

  For an initial palette color $P_i$ that represents a cluster $\{C_1, C_2, ..., C_{R_i}\}$ of image colors, a set $S_i$ of $\frac{R_i}{N}(K - M)$ distinct image colors will be selected around $P_i$ from $\{C_1, C_2, ..., C_{R_i}\}$ and are assigned as new palette colors. $N$ is the total number of distinct colors in the original image. The selection of the $S_i$ palette colors around $P_i$ is done as follows:

$$P_{i_j} = \min_{P \in \{P_{i_1}, P_{i_2}, .., P_{i_{j-1}}\}} \{d(P, C) > 3 : C \in \{C_1, C_2, .., C_{R_i}\}\} \tag{5.6}$$

$$j = 2, \ldots, |S_i| + 1, (P_{i_1} = P_i)$$

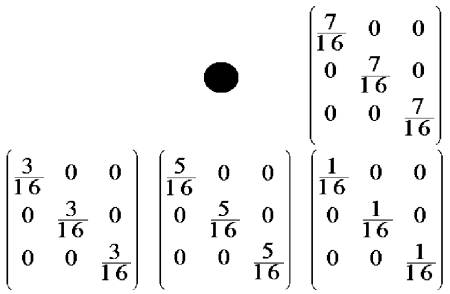


Figure 5.2: Floyd-Steinberg vector filter for a single pixel.

where $d$ is the Euclidian distance between two colors in the CIELAB color space.

- ***Step 3:*** *Pixel Mapping with Error Diffusion*

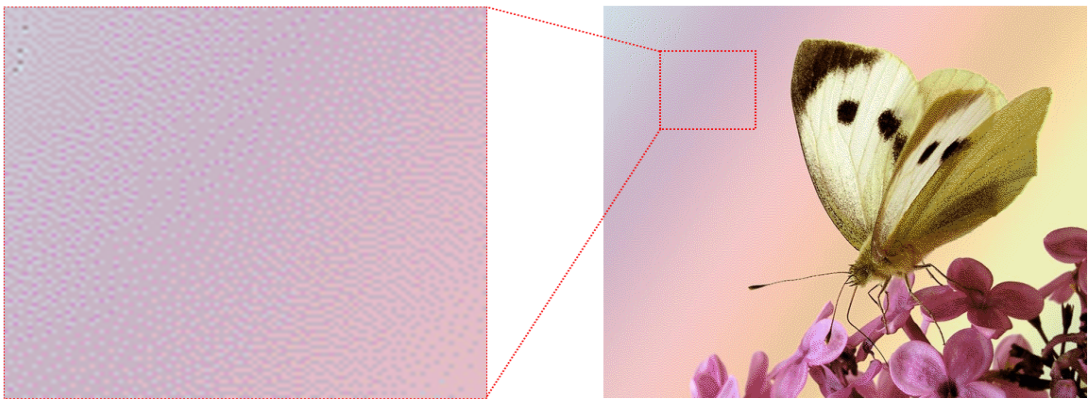
  Once the color palette has been designed, the quantized image is produced by mapping each pixel to the closest color from the color palette. To overcome the problem of false contours that are commonly sensed in color quantized images regardless of the color quantization algorithm used, we incorporate error diffusion technique within the pixel mapping process. The pixels in the original image are chosen in raster ordering from left to right and top to bottom. Each pixel is quantized and the quantization error is propagated forward to the neighboring pixels that have not yet been quantized. Fig. 5.2 shows Floyd-Steinberg [94] filter that we used to perform the error diffusion.

### 5.2.3 Results and Discussions

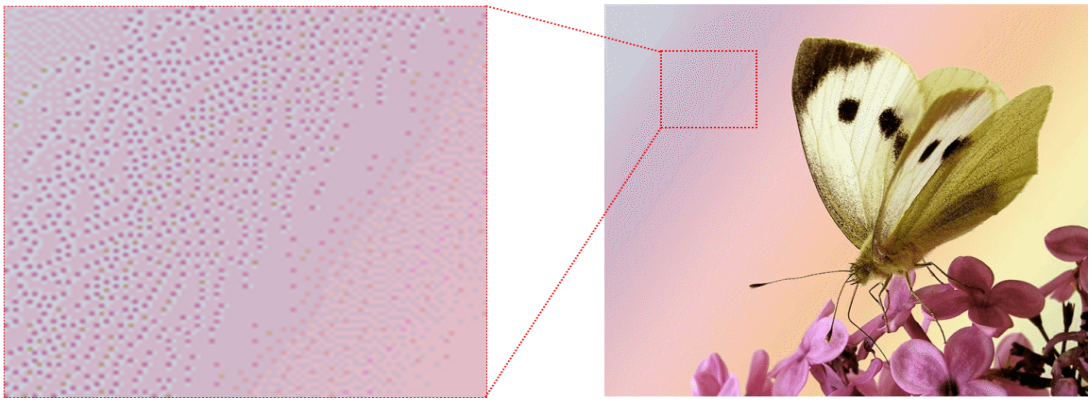
Standard color quantization process is generally considered as a clustering problem to find the $K$ palette colors that minimize some error criterion for all the colors in an image. However, two problems are associated with performing error diffusion techniques for color quantized images. The first problem is, if a given

color is close to the boundary of a color cluster, the quantization errors accumulate and eventually a palette color from a different cluster is produced. This exposes itself as a color impulse that is very disturbing. On the other hand, in smooth regions usually the colors are very close to a palette color with a small dithered error. Therefore those colors will be mapped to the same palette color forming flat regions. This results the existence of some false contours after dithering. These two problems disappeared in our proposed method because the palette colors are distributed around few initially selected palette colors. Therefore it needs a small diffused error to jump from one palette color into another close palette color in the same cluster. This results a high alternation of palette colors between neighboring pixels to give full illusion of smooth gradation and hence reduces visible false contours. The selection of palette colors from the same cluster will avoid the presence of color impulses.

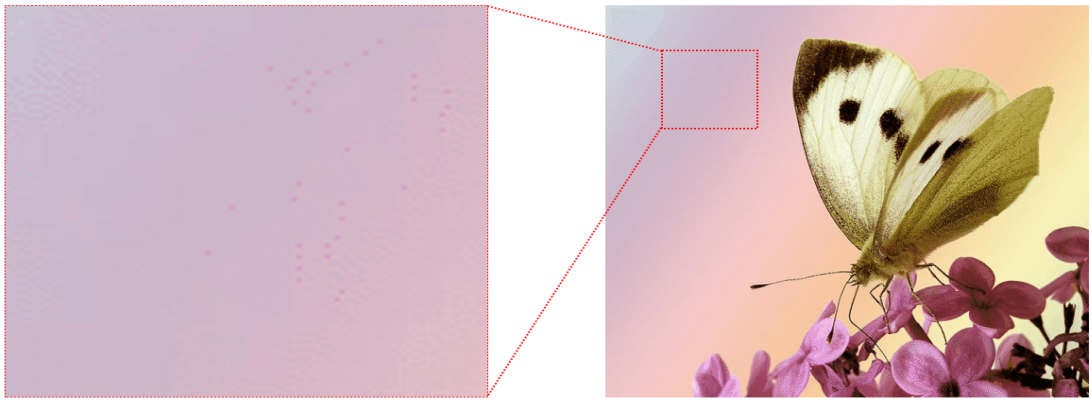
In Fig. 5.3 and Fig. 5.4 two full color images are quantized by the proposed method. The same two images are also presented to five popular quantization algorithms namely: k-means [25], median-cut [2], octree algorithm [16], SOM [31], Wu's quantizer [13] in which the Floyd-Steinberg error diffusion was incorporated within the pixel mapping process. It is observed that applying the Floyd-Steinberg dithering on the other quantization algorithms causes the creation of color impulses. The color impulses are in the form of isolated green and dark blue pixels on sky area in Fig. 5.4(b)-(f). It is also observed that some false contours remain visible in these images in the smooth gradient area in Fig. 5.3(b)-(f). Fig. 5.3(a) and Fig. 5.4(a) are the results of the proposed method that shows better performance when compared to the other algorithms. As expected, disturbing colored impulses and false contours are minimized compared with other algorithms.



(a)

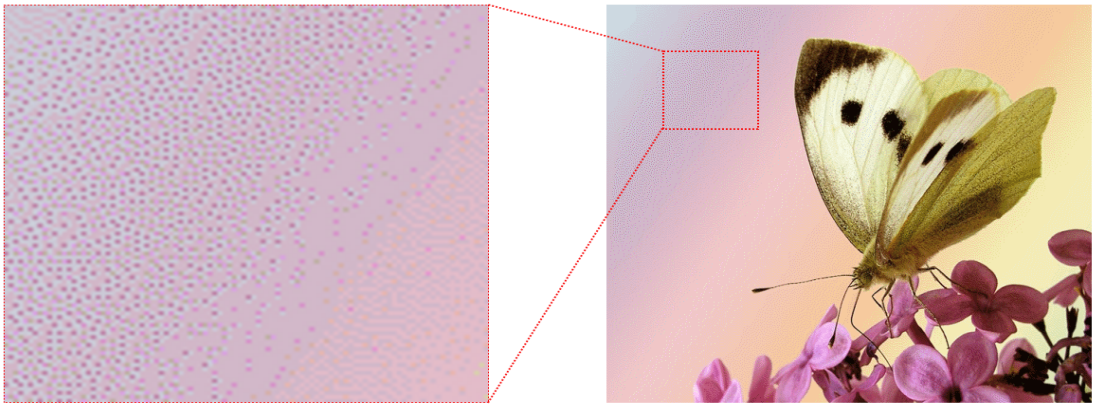


(b)

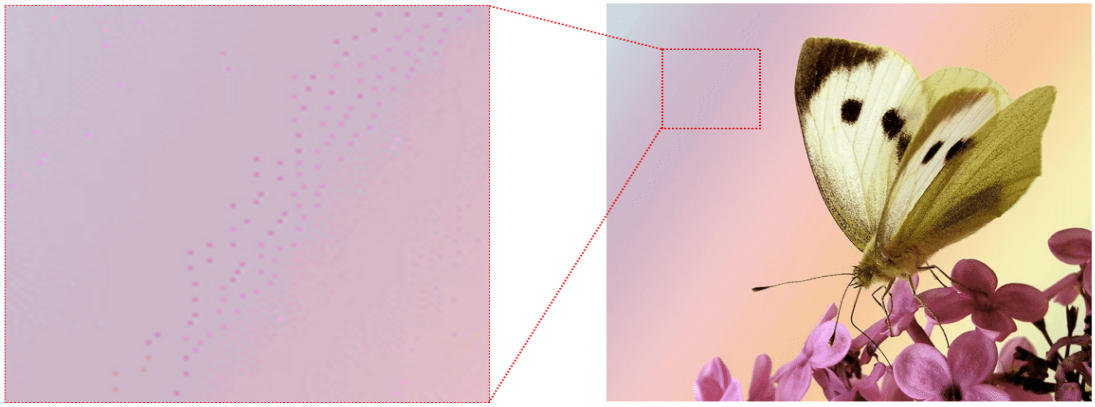


(c)

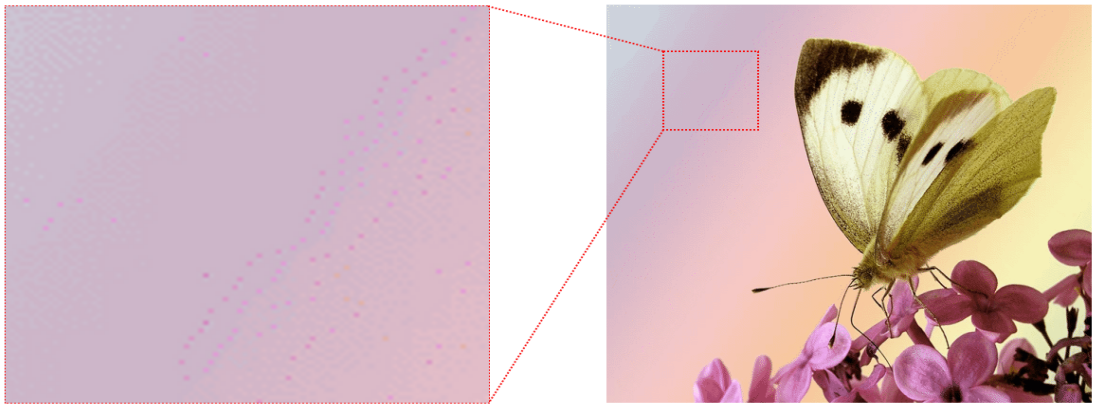
Figure 5.3. (cont'd)



(d)

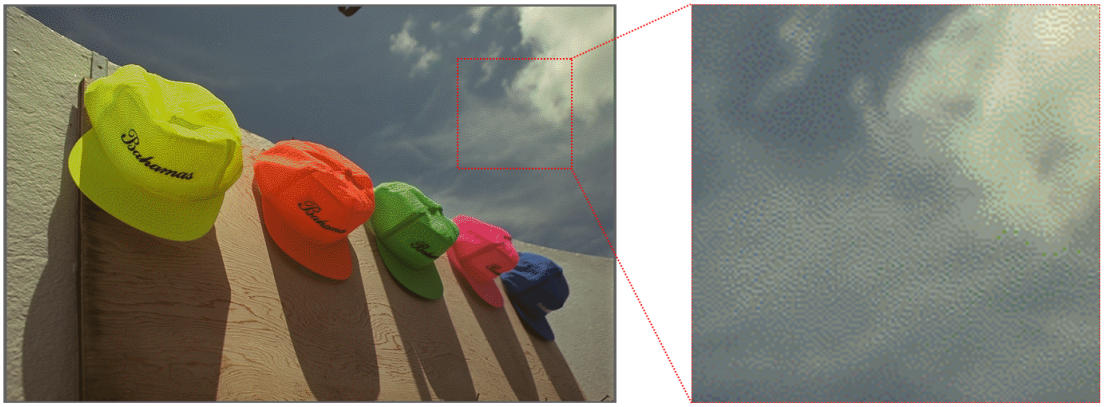


(e)

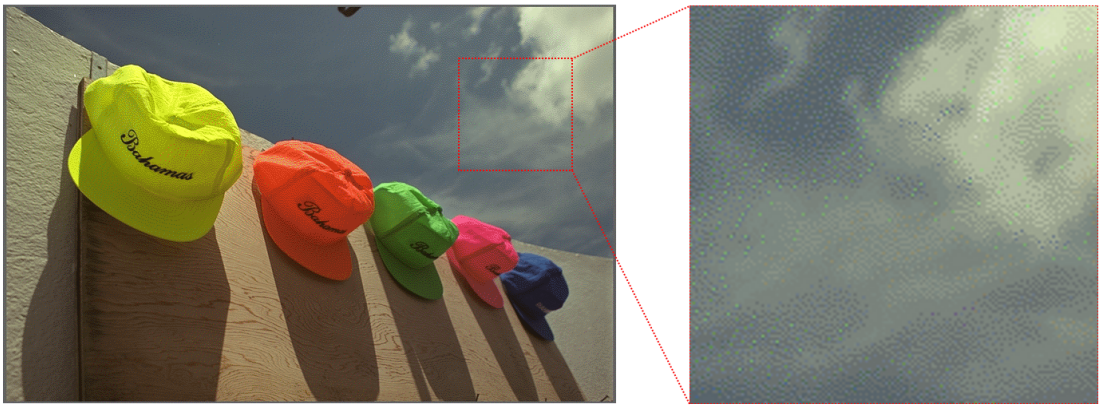


(f)

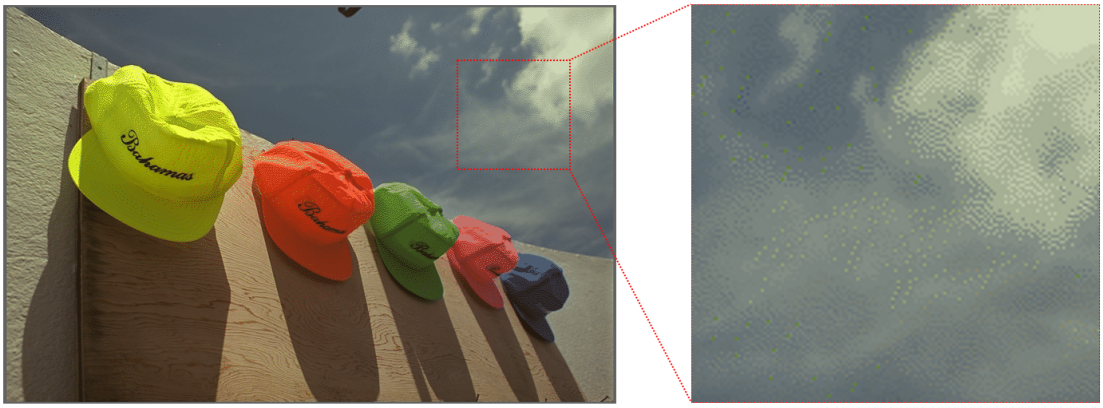
Figure 5.3: Color quantized images with 128 colors followed by Floyd-Steinberg error diffusion. (a) Proposed method, (b) K-means, (c) Median Cut, (d) Octree, (e) SOM, (f) Wu's quantizer.



(a)

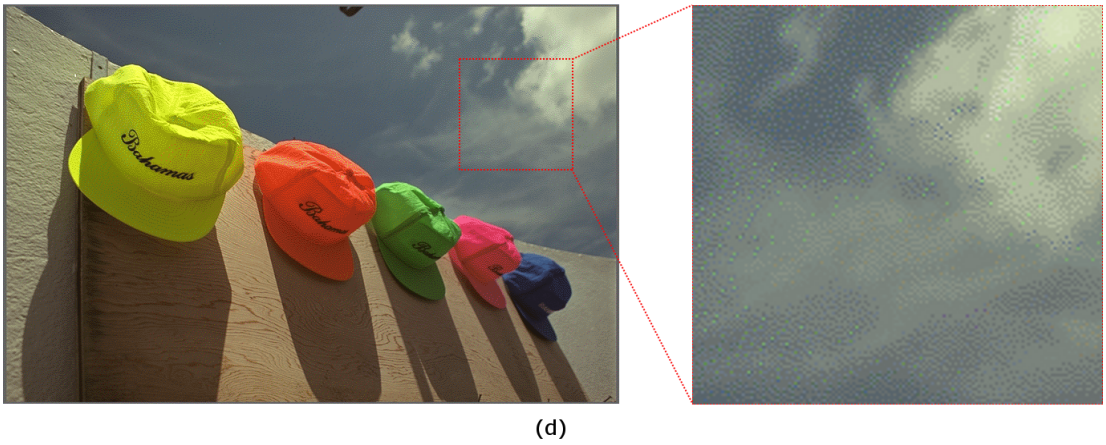


(b)

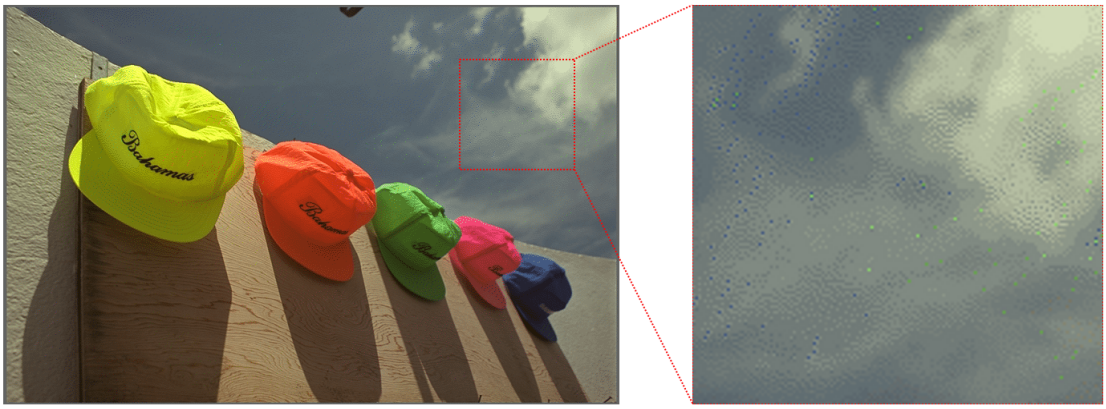


(c)

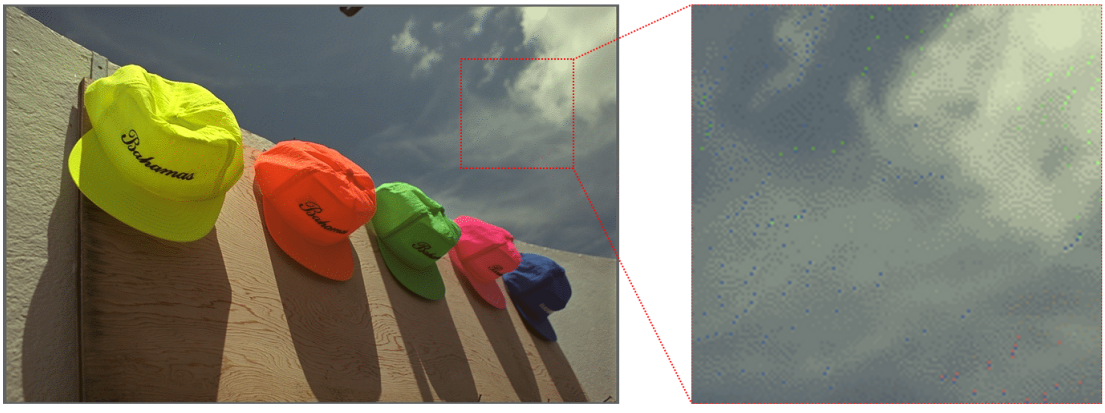
Figure 5.4. (cont'd)



(d)



(e)



(f)

Figure 5.4: Color quantized images with 128 colors followed by Floyd-Steinberg error diffusion. (a) Proposed method, (b) Kmeans, (c) Median Cut, (d) Octree, (e) SOM, (f) Wu's quantizer.

## 5.3 Local Fractal Dimension-Based Color Quantization for Error Diffusion Techniques

In this section, I propose a method for color quantization to be used with dithering techniques. The premise of the proposed method is to improve the perceived quality in low activity (smooth) regions where *false contours in an image are most likely to occur* while preserving structure that is commonly lost in high activity (busy) regions when applying dithering techniques. Improving the perceived quality in low activity regions is achieved by introducing more colors to regions with low activity in order to move quantization errors from low activity regions to high activity regions. The errors will be less visible to the human eye as the human visual system is less sensitive to quantization errors in high activity regions than to errors in low activity regions. Preserving structure in high activity regions will be achieved by applying dithering techniques only to low activity regions where false contours in an image are most likely to occur.

### 5.3.1 Local Fractal Dimension

Fractal dimension of imaged 3D surfaces can be used as a measure to perceptually distinguish between smooth and rough textured regions [99, 100]. It has been widely applied to many fields of digital image processing, such as texture classification and segmentation[101–104], image data compression [105], and computer graphics [106].

Sarkar and Chaudhuri [107, 108] proposed a method to find the fractal dimension of images known as differential box counting method. In their method, for

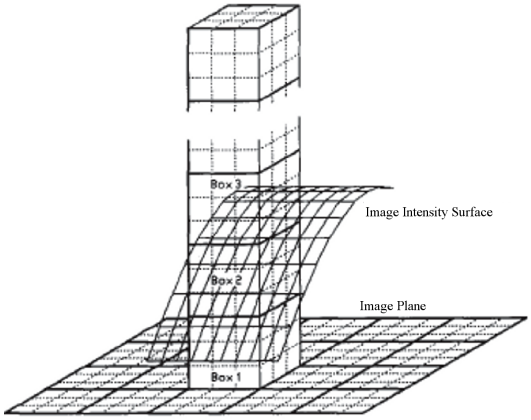



Figure 5.5: Estimation of fractal dimension using differential box counting method

an image of size $M \times M$ pixels scaled down to a size $s \times s$ where $M/2 \geq s > 1$ and $s$ is an integer. The entire image can be considered as a 3-D space with $(x, y)$ denoting 2-D position and the third coordinate $(z)$ denoting gray level. The $(x, y)$ space is partitioned into grids of size $s \times s$. On each grid there is a column of boxes of size $s \times s \times s'$ (see Fig. 5.5). If the total number of gray levels is $G$, then $\lfloor G/s' \rfloor = \lfloor M/s \rfloor$. The contribution of the $(i, j)^{th}$ grid is computed as

$$n_r = l - k + 1 \tag{5.7}$$

where the minimum and maximum gray level of the image in the $(i, j)^{th}$ grid fall in box number $k$ and $l$, respectively. For example in Fig. 5.5 where $s = s' = 3$, assign numbers 1, 2, . . to the boxes as shown. Let the minimum and maximum gray level of the image in the $(i, j)^{th}$ grid fall in box number 1 and 3, respectively. The contribution of the $(i, j)^{th}$ grid is given as $n_r(i, j) = 3 - 1 + 1$.

Taking the contribution of all grids, $N_r = \sum_{i,j} n_r$. $N_r$ is counted for different

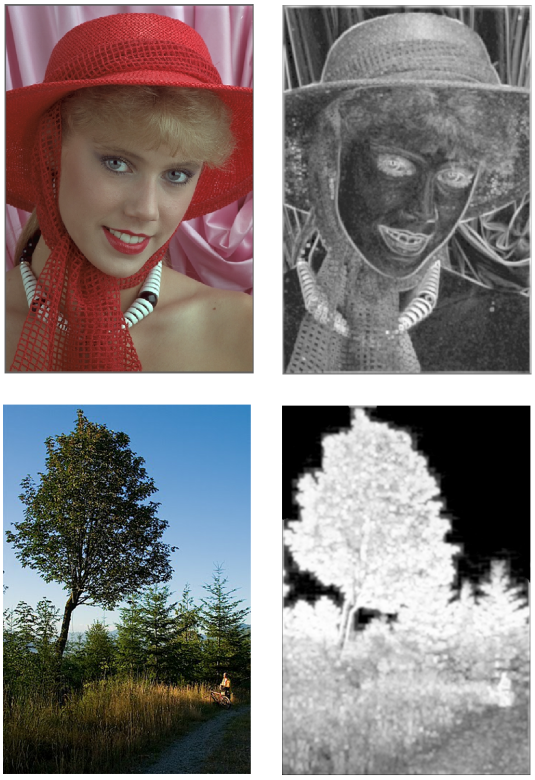


Figure 5.6: LFD map of an image: (left) the original images and (right) their LFD maps

values of $r$, i.e., different values of $s$. Using the following equation:

$$1 = N_r \times r^D \quad \text{or} \quad D = \log(N_r)/\log(1/r) \tag{5.8}$$

The fractal dimension $D$ can be estimated from the least square linear fit of $log(N_r)$ against $log(1/r)$.

Fig. 5.6 shows the original images (left) and their local fractal dimension (LFD) maps (right) where every position $(i, j)$ in the LFD map represents the LFD value estimated using a 3×3 local window centered at position $(i, j)$ in the original image. The LFD values are scaled into the rage 0-255 for display purpose. Notice how the low activity (smooth) regions are given small LFD values (darker) while the high

activity (busy) regions are given large LFD values (brighter). It would be easy to distinguish between such regions and weight the importance of a pixel based on the spatial activity of the area of the image where the pixel appears. The pixel's importance (activity weighting) is inversely proportional to the activity level of the region where it appears in the image.

### 5.3.2 Construction of Color Palette

In this section, we present the proposed algorithm for color image quantization using local fractal dimension (LFD) maps. To describe the proposed algorithm, we first explain the k-means [25] algorithm which is the basis of our proposed algorithm. Generally the k-means algorithm seeks to minimize an objective function which is defined as

$$j = \sum_{i=1}^{k} \sum_{C \in S_i} W(C) \left\| C - \bar{C}_{S_i} \right\|^2 \tag{5.9}$$

where $\bar{C}_{S_i}$ is the center of the cluster $S_i$, $k$ is the number of clusters, and $W(C)$ is the weight factor of the pixel $C$.

Initially the k-means algorithm starts with a set of $k$ centers that are randomly selected. At each iterative step, scan through all the pixels of the original image and assign each pixel to the nearest center in the sense of minimizing (5.9). Then the new centers are calculated as

$$\bar{C}_{S_i} = \frac{\sum_{C \in S_i} W(C) \times C}{\sum_{C \in S_i} W(C)} \qquad i = 1, 2, .., k \tag{5.10}$$

The steps are repeated until the algorithm converges or the number of iterations reaches a specified value. Usually in the conventional k-means algorithm

the weight factor $W(C) = 1$ for all pixels, therefore all pixels are treated equally regardless of their spatial distribution.

Our proposed color quantization algorithm follows the same steps as the k-means algorithm except that the weight factor $W(C)$ of pixel $C$ at the $(i, j)^{th}$ position in the original image is equal to the inverse of its corresponding LFD value at the $(i, j)^{th}$ position in the LFD map. In this way pixels are not equally treated where pixels in the low activity (smooth) regions are given larger weight (importance) than those pixels in the high activity (busy) regions. Accordingly, more levels of colors will be presented in low activity regions and this will not degrade the quality of the image as the human eyes are less sensitive to quantization errors in high activity regions than in low activity regions. As a result, the perceived quality in low activity regions will be improved after dithering techniques are applied as shown in the next section. Fig. 5.7 shows the color quantization outputs for the conventional k-means and the proposed algorithm for 16 and 32 colors. It is clear that the proposed algorithm allocates more levels of colors to the gradient sky region than the conventional k-means does to the same region for both 16 and 32 colors.

### 5.3.3 Pixel Mapping with Error Diffusion

To succeed in facing up to the problem of false contours in the color quantized images, we incorporate error diffusion technique within the pixel mapping process. In conventional systems for error diffusion, the pixels are chosen in raster ordering from left to right and top to bottom. Each pixel is quantized and the quantization error is propagated forward to the neighboring pixels that have not yet been

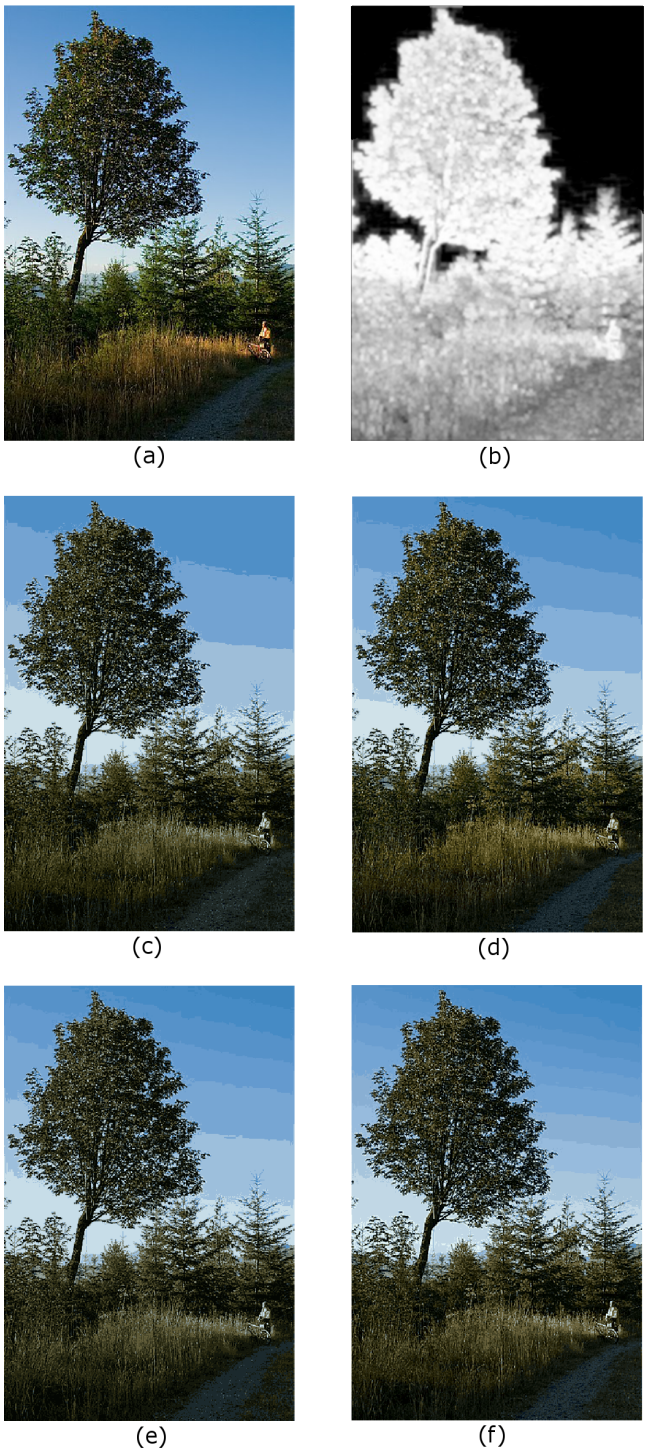



Figure 5.7: Color quantization results of 'Tree' image. (a) The original image, (b) LFD map of (a), (c) K-means 16 colors, (d) k-means 32 colors, (e) Proposed method 16 colors, (f) Proposed method 32 colors.

quantized with a causal filter. One problem is associated with this approach that the structure in high activity regions may be distorted due to the accumulated quantization error (see Fig. 5.9(a)). To avoid unnecessarily dithering of high activity regions, the pixels in high activity regions are masked using smooth-region map in a way that the Floyd-Steinberg error diffusion filter [94] is applied only to pixels in low activity regions where false contours in an image are most likely to occur. In this way false contours in low activity regions will be removed while preserving the structure in high activity regions that is commonly lost in high activity regions when applying dithering techniques.

The smooth-region map is created by thresholding the LFD map using the Otsu's global image threshold method. Fig. 5.9(b) and (d) show the LFD map of the 'Tree' image and its corresponding smooth-region map respectively. The fundamental steps involved in the local fractal dimension-based color quantization method are delineated in Fig. 5.8.

### 5.3.4 Results and Discussions

In Fig. 5.9(a) the full color image 'Tree' has been quantized to 16 colors using the conventional k-means where the Floyd-Steinberg error diffusion [94] has been incorporated within the pixel mapping process. It is observed that the application of the Floyd-Steinberg error diffusion causes the creation of color impulses and some false contours remain visible in the sky region as well as distorting structure in grass region. Those color impulses and false contours are the result of insufficient color levels presented in the gradient sky region, while the distorted structure is the result of unnecessarily dithering high activity regions where false contours in

**Input:** A full color image $I$, Desired number of palette colors $K$.
**Output:** A quantized image with $K$ colors .
**Begin**
%Color Palette Generation Step%
$LFD = Create\_LFD\_Map(I)$
*Randomly initialize* $K$ *centroids* $\{C_{S_1}, C_{S_2}, ..., C_{S_k}\}$
**while** *termination condition is not satisfied* **do**
*Assign image pixels to the closest centroid*
*Update cluster centroids as:* $C_{S_i} = \frac{\sum_{p \in S_i} LFD(p) \times p}{\sum_{p \in S_i} LFD(p)} \qquad i = 1, 2, .., k$
**end while**

%Pixel Mapping Step%
$Smooth\_Area\_Map = Otsu's\_Threshold(LFD)$
**for all** $p \in I$ **do**
**if** $Smooth\_Area\_Map(p) == 0$ **then**
$Quantize(p)$
$Error = Quantization\_Error(p)$
$Floyd\_Steinberg\_Filter(Error, p)$
**else**
$Quantize(p)$
**end if**
**end for**
**End**

Figure 5.8: Pseudocode for the local fractal dimension-based color quantization algorithm.

an image are not apparent to the human eye. On the other hand Fig. 5.9(c) shows the proposed color quantization outputs of 16 colors. It is clear that the false contours are less visible in the sky region compared to the conventional k-means and the result is free from color impulses. This is due to introducing more color levels in gradient sky region than the conventional k-means (see Fig. 5.7(c) and (e)). As a result, dithering in low activity regions makes a pixel jump to another close palette color for even small errors. This results a high alternation of palette colors between neighboring pixels to give full illusion of smooth gradation. At the same time, it may also be noted that structure in the grass region is preserved by

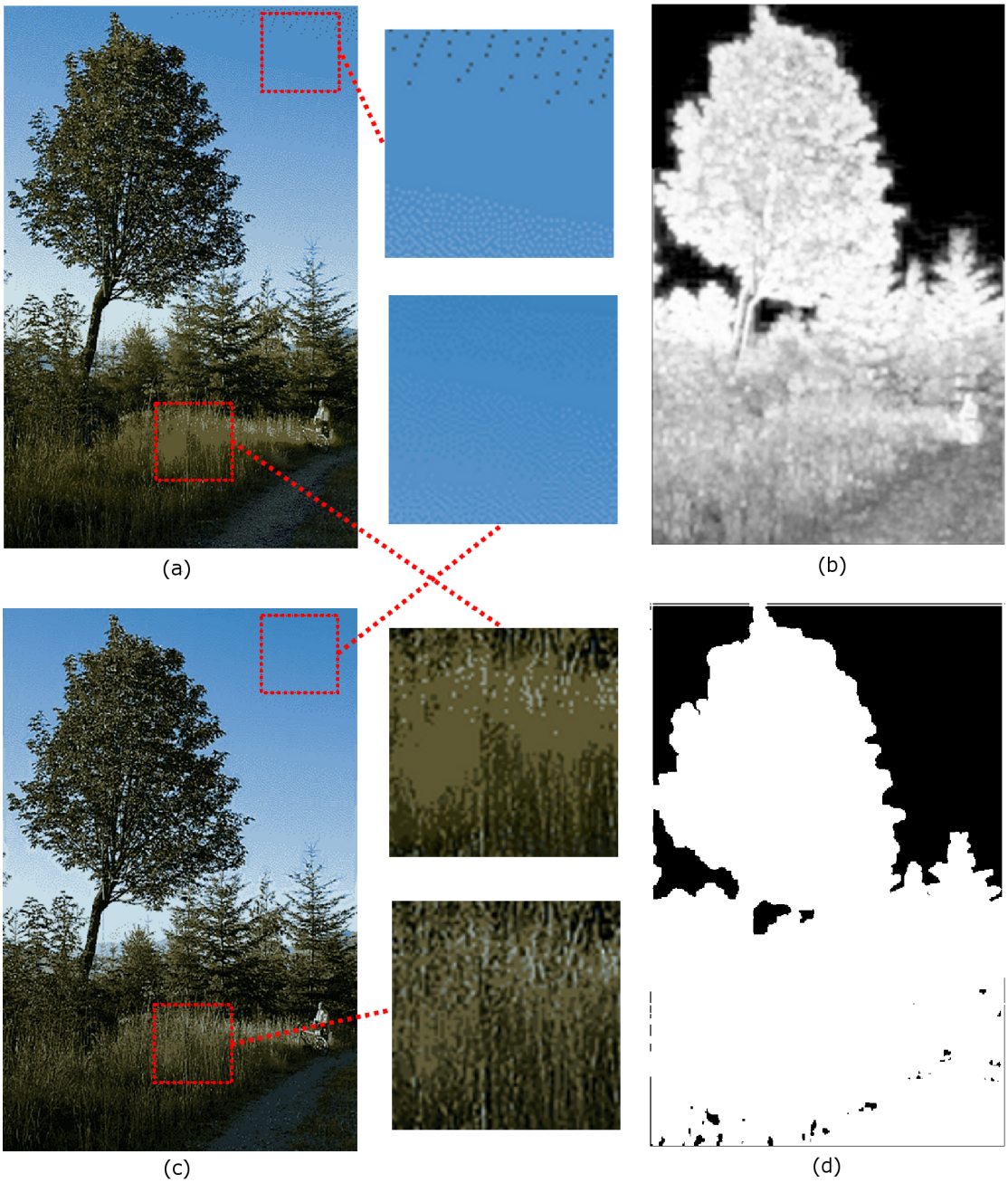



Figure 5.9: Results of joint color quantization and dithering of 'Tree' image. (a) k-means 16 colors, (b) LFD map of original 'Tree' image, (c) Proposed algorithm 16 colors, (d) smooth-region map of original 'Tree' image (high activity regions in white color and low activity regions in black color).

not performing dithering in the high activity regions.

For the objective evaluation, two image quality metrics have been used namely: the widely used peak signal-to-noise ratio metric (PSNR) where a higher PSNR value means that the distorted image is closer to the original, and the structural similarity measure (SSIM) [57] to measure how close is the structure of the resultant dithered quantized image to the structure of the original image. Table 5.1 shows the performance comparison of joint color quantization and dithering for the algorithms: median-cut [2], Orchard's method [18], k-means algorithm [25], and the proposed method in terms of PSNR, and SSIM metrics for the testing images in Fig. 5.10. The results show that the proposed method produces quantized images that are perceptually closer to the original images, and at the same time it preserves more structure of the original images when compared with the other methods.


Figure 5.10: Original images used for the different experiments. (a) Tree (59974 colors), (b) Woman (31744 colors), (c) Flowers (47946 colors) , (d) House (154605 colors).

| Image | Colors | Median Cut+ED | | Orchard's method | | K-means+ED | | Proposed Method | |
|---|---|---|---|---|---|---|---|---|---|
| | | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| **Tree** | 32 | 26.414 | 0.791 | 30.179 | 0.887 | 29.854 | 0.883 | **31.666** | **0.891** |
| | 64 | 28.468 | 0.834 | 32.305 | 0.914 | 32.014 | 0.924 | **35.214** | **0.934** |
| | 128 | 31.736 | 0.903 | 34.972 | 0.951 | 34.910 | 0.954 | **37.509** | **0.955** |
| | 256 | 34.287 | 0.934 | 37.335 | 0.969 | 36.924 | 0.968 | **39.129** | **0.975** |
| **Woman** | 32 | 27.165 | 0.679 | 30.143 | 0.776 | 29.039 | 0.778 | **30.724** | **0.808** |
| | 64 | 30.106 | 0.772 | 32.369 | 0.831 | 32.682 | 0.858 | **34.297** | **0.872** |
| | 128 | 31.912 | 0.839 | 35.002 | 0.892 | 34.986 | 0.905 | **36.682** | **0.921** |
| | 256 | 34.550 | 0.891 | 37.646 | 0.934 | 37.128 | 0.933 | **38.880** | **0.950** |
| **Flowers** | 32 | 26.403 | 0.745 | 28.833 | 0.750 | 28.756 | 0.776 | **30.017** | **0.794** |
| | 64 | 27.454 | 0.752 | 30.333 | 0.813 | 31.605 | 0.833 | **32.124** | **0.874** |
| | 128 | 29.963 | 0.809 | 33.039 | 0.853 | 33.527 | 0.902 | **35.405** | **0.911** |
| | 256 | 32.240 | 0.831 | 35.517 | 0.897 | 35.483 | 0.927 | **37.446** | **0.937** |
| **House** | 32 | 25.601 | 0.689 | 28.867 | 0.795 | 29.575 | 0.829 | **32.310** | **0.873** |
| | 64 | 27.927 | 0.787 | 31.159 | 0.873 | 31.753 | 0.881 | **34.518** | **0.916** |
| | 128 | 30.454 | 0.837 | 33.965 | 0.916 | 34.162 | 0.924 | **35.949** | **0.934** |
| | 256 | 31.952 | 0.839 | 36.043 | 0.938 | 36.287 | 0.947 | **38.059** | **0.957** |

Table 5.1: Objective validation of the LFD-based method. The validation criteria are: the widely used peak signal-to-noise ratio metric (PSNR) and the structural similarity measure (SSIM) [57].

# CHAPTER 6

# Conclusions and Future Work

Color quantization is still an important problem in the fields of image processing and computer graphics. In this dissertation, I addressed three problems related to color image quantization. The first problem is the creation of a ground truth database for human perception of the quality of color quantized images, the second is the objective quality assessment of color quantized images, and the third is the joint color quantization and dithering techniques. I summarize below the main contributions of this dissertation and present some possible future directions.

In Chapter 3, we investigated *the subjective quality assessment of color quantized images* in which we have conducted extensive psychometric experiments to collect subjective quality scores for a diverse set of images quantized by five popular quantization algorithms. The subjective mean opinion scores (MOS) from the subjective study were used for training and testing the algorithms presented in this dissertation and also for validating the correlation of the existing image quality metrics with human perception in evaluating the quality of color quantized images. To my knowledge, this study is the only extensive study conducted in the quality assessment literature for color quantization distortion, containing about 19,200 subjective quality evaluations by 22 human subjects. Moreover, we have made the image database publicly available for downloading for research in image quality assessment and other applications.

Chapter 4 addresses the problem of *objective quality assessment of color quan-*

*tized images*, where we proposed two image quality assessment methods. The first method was an extension to the well-known general-purpose multi-scale structural similarity (MSSIM) index in which a color comparison is added to the criteria of the gray-scale MSSIM index. The proposed method outperformed traditional image quality assessment methods in the quality assessment of color quantization distortion. The performance of the extended MSSIM has been analyzed also on the Tampere Image Database (TID2008), which is largest database so far available with 17 types of distortions, to further test the extended metric for different distortion types. The results show that the proposed extension greatly improves the performance of the MSSIM for the assessment of other distortion types. This improvement reaches around 40% in impulse noise and 23% in additive Gaussian noise.

The second method was an application-specific method for the quality assessment of color quantized images that models any distortion as a combination of three similarities: color similarity, edge similarity, and structure similarity. I demonstrated that the proposed method outperforms traditional image quality assessment methods in the quality assessment of color quantization distortion in our simulations by a sizable margin.

In Chapter 5, we present the problem of *joint color quantization and dithering techniques*. Usually color quantization and dithering are performed sequentially; the quality of the dithering process depending on the selection of the color palette. Standard methods of color quantization do not take dithering into account. Therefore, the desirable qualities obtained by quantization are often disturbed by the dithering process. In this chapter we proposed two color quantization methods that are designed to obtain better results with minimum visual degradation after

the error diffusion technique is applied. The first color quantization method consists of three steps: in the first step, the colors in original image are quantized into a number of colors that is much smaller than the desired number of quantization levels. These initial colors represent the dominant colors in the original image. In the second step, the remaining palette colors are selected around each palette color generated in the first step based on the CIELAB color space Just Noticeable Color Difference (JNCD) threshold. The use of JNCD threshold is to ensure that the selected palette colors will be far from each other by a distance that is enough for the selected palette colors to be distinguishable from each other. Finally error diffusion technique is incorporated within the pixel mapping processing. The results show that our method completely removes false contours as well as prevents color impulses that are common results of applying dithering techniques to color quantized images.

The second color quantization method is based on local fractal dimension. The premise of the proposed method is to improve the perceived quality in low activity (smooth) regions where false contours in an image are most likely to occur while preserving structure that is commonly lost in high activity (busy) regions when applying dithering techniques. The local fractal dimension is used to assign different weights to image pixels according to their spatial distribution. This results in introducing more colors to regions with low activity and thus moving quantization errors from low activity regions to high activity regions. The errors will be less visible to the human eye as the human visual system is less sensitive to quantization errors in high activity regions than to errors in low activity regions. Structure in high activity regions is preserved by applying dithering techniques only to low activity regions where false contours in an image are most likely to occur. Simula-

tion results show that the proposed method gives a remarkable improvement in the quality of joint color quantization and dithering both subjectively and objectively.

There are a number of possible future directions in which my dissertation research could be extended:

First, during my research I found that some of the color quantization algorithms well-correlate with human perception than others in most cases. The scores from the subjective study can be used to conclude the best color quantization algorithm, in terms of correlation with human perception, out of the five popular algorithms that are used in the subjective study.

Second, the color quantization quality assessment metric proposed in this dissertation can be embedded in a color quantization algorithm to guide the process of quantization in order to improve the perceptual quality of the delivered color quantized images.

Third, in spite of error diffusion is the most commonly dithering technique used on display device to reduce the appearance of false contours that result from the quantization process. Other dithering techniques such as ordered dithering and dot diffusion can also be studied to reduce the appearance of false contours in color quantized images.

Fourth, during the study of joint color quantization and dithering techniques, I dealt only with display devices. It is well-known that the appearance of an image on a display device differs from the appearance on a hardcopy device. Therefore, further study is required to validate the proposed color quantization algorithms on hardcopy devices.

# REFERENCES


[1] S. Cheng and C. Yang. A fast and novel technique for color quantization using reduction of color space dimensionality. *Pattern Recognition Letters*, 22(8):845–856, 2001.

[2] P. Heckbert. Color image quantization for frame buffer display. In *Proceedings of SIGGRAPH*, volume 16, pages 297–307, 1982.

[3] P. Cosman, K. Oehler, E. Riskin, and R. Gray. Using vector quantization for image processing. *Proceedings of the IEEE*, 81(9):1326–1341, 1993.

[4] J. Braquelaire and L. Brun. Comparison and optimization of methods of color image quantization. *IEEE Transactions on Image Processing*, 6(7):1048–1052, 1997.

[5] M. Swain and D. Ballard. Color indexing. *International journal of computer vision*, 7(1):11–32, 1991.

[6] C. Yang and W. Tsai. Color image compression using quantization, thresholding, and edge detection techniques all based on the moment-preserving principle. *Pattern Recognition Letters*, 19:205–215, 1998.

[7] E. van den Broek, T. Kok, T. Schouten, and L. Vuurpijl. Human-centered content-based image retrieval. In *Proceedings of SPIE*, volume 6806, page 54, 2008.

[8] A. Mojsilovic, J. Kovacevic, J. Hu, R. Safranek, and S. Ganapathy. Matching and retrieval based on the vocabulary and grammar of color patterns. *IEEE Transactions on Image Processing*, 9(1):38–54, 2000.

[9] G. Joy and Z. Xiang. Center-cut for color-image quantization. *The Visual Computer*, 10(1):62–66, 1993.

[10] P. Heckbert. Color Image Quantization for Frame Buffer Display. Bachelor's thesis, Massachusetts Institute of Technology, 1980.

[11] S. Wan, S. Wong, and P. Prusinkiewicz. An algorithm for multidimensional data clustering. *ACM Transactions on Mathematical Software*, 14(2):153–162, 1988.

[12] S. Wan, P. Prusinkiewicz, and S. Wong. Variance-based color image quantization for frame buffer display. *Color Research and Application*, 15(1):52–58, 1990.

[13] X. Wu. Efficient statistical computations for optimal color quantization. In *Graphics Gems II*, pages 126–133. New York: Academic, James Arvo edition, 1991.

[14] X. Wu. Color quantization by dynamic programming and principal analysis. *ACM Transactions on Graphics*, 11(4):348–372, 1992.

[15] T.-S. Liu and L.-W. Chang. Greedy tree growing for color image quantization. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 19–22, 1994.

[16] M. Gervautz and W. Purgathofer. A simple method for color quantization:

Octree quantization. In *Graphics Gems*, pages 287–293. Academic Press Professional, Inc., San Diego, CA, USA, 1990.

[17] Y. Sirisathitkul, S. Auwatanamongkol, and B. Uyyanonvara. Color image quantization using distances between adjacent colors along the color axis with highest color variance. *Pattern Recognition Letters*, 25(9):1025–1043, 2004.

[18] M. Orchard and C. Bouman. Color quantization of images. *IEEE Transactions on Signal Processing*, 39(12):2677–2690, 1991.

[19] R. Balasubramanian, J. Allebach, and C. Bouman. Color-image quantization with use of a fast binary splitting technique. *Journal of the Optical Society of America A*, 11(11):2777–2786, 1994.

[20] C. Yang and J. Lin. Color quantization by RWM-cut. In *Proceedings of the Third International Conference on Document Analysis and Recognition*, volume 2, pages 669–672, 1995.

[21] I.-S. Hsieh and K.-C. Fan. An adaptive clustering algorithm for color quantization. *Pattern Recognition Letters*, 21:337–346, 2000.

[22] G. Braudaway. A procedure for optimum choice of a small number of colors from a large color palette for color imaging. In *Proceedings of International Conference on Electronic Imaging Exposition*, pages 75–79, 1986.

[23] R. Gentile, J. Allebach, and E. Walowit. Quantization of color images based on uniform color spaces. *Journal of Imaging Technology*, 16:11–21, 1990.

[24] M. Larabi, N. Richard, and C. Fernandez. A fast color quantization using a

matrix of local pallets. In *Proceedings of the 29$^{th}$ Applied Imagery Pattern Recognition Workshop*, pages 136–140, 2000.

[25] Y. Linde, A. Buzo, and R. M. Gray. An algorithm for vector quantizer desing. *IEEE Transactions on Communications*, 28(1):84–95, 1980.

[26] J. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[27] M. Astrahan. Speech analysis by clustering or the hyperplane method. Stanford A.I. Project memo, Stanford University, California, 1970.

[28] J. Hartigan and M. Wong. A k-means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.

[29] M. Belal and A. Daoud. A new algorithm for cluster initialization. *Worls Ancademy of Science, Engineering and Technology*, 4:74–76, 2005.

[30] B. Mirkin. *Clustering for data mining: A data recovery approach*. London: Chapman and Hall, 2005.

[31] A. Dekker. Kohonen neural networks for optimal colour quantization. *Network Computation in Neural Systems*, 5(3):351–367, 1994.

[32] T. Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag, Berlin, third edition, 1989.

[33] N. Papamarkos. Color reduction using local features and a sofm neural network. *International Journal of Imaging Systems and Technology*, 10(5):404–409, 1999.

[34] N. Papamarkos, A. Atsalakis, and C. Strouthopoulos. Adaptive color reduction. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 32(1):44–56, 2002.

[35] J. Zhang and J. Hu. A hierarchical clustering method for color quantization. In *Proceedings of 20$^{th}$ International Conference on Pattern Recognition*, pages 23–26, 2010.

[36] M. Omran, A.Engelbrecht, and A. Salman. A color image quantization algorithm based on particle swarm optimization. *Informatica*, 29:261–269, 2005.

[37] X. Hu, T. Wang, and D. Li. A new approach of color quantization based on ant colony clustering algorithm. In *Proceedings of the International Conference on Information Technology: Coding and Computing*, volume 1, pages 102–108, 2005.

[38] Z. Wang and C. Bovik. Modern image quality assessment. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1):1–156, 2006.

[39] Z. Wang, H. Sheikh, and C. Bovik. Objective video quality assessment. In B. Furht and O. Marques, editors, *The Handbook of Video Databases: Design and Applications*, pages 1041–1078. CRC Press, 2003.

[40] S. Ouni, M. Chambah, M. Herbin, and E. Zagrouba. Are existing procedures enough? image and video quality assessment: Review of subjective and objective metrics. In *Electronic Imaging, Image Quality and System Performance*, San Jose, USA, 2008. SPIE.

[41] W. Lin and C.-C. Kuo. Perceptual visual quality metrics: A survey. *Journal of Visual Communication and Image Representation*, 22(4):297–312, 2011.

[42] M. Pedersen and J. Hardeberg. Survey of full-reference image quality metrics: Classification and evaluation. *Foundations and Trends in Computer Graphics and Vision*, 7(1):1–80, 2012.

[43] T. Mitsa and K. Varkur. Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms. In *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal processing*, pages 301–304, 1993.

[44] C. Chou and Y. Li. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *IEEE Tranactions on circuits and systems for video technology*, 5(6):467–476, 1995.

[45] C. Lambrecht and J. Farrell. Perceptual quality metric for digitally coded color images. In *Proceedings of the European Signal Processing Conference*, pages 10–13, 1996.

[46] A. Poirson and B. Wandell. Appearance of colored patterns: Pattern-color separability. *Journal of the Optical Society of America*, 10(12):2458–2470, 1993.

[47] X. Zhang and B. Wandell. A spatial extension of CIELAB for digital color image reproduction. In *Proceedings of SID International Symposium Digest of Technical Papers*, volume 27, pages 731–734, 1996.

[48] R. Iordache and A. Beghdadi. A wigner-ville distribution-based image dis-

similarity measure. In *Proceedings of the 6$^{th}$ International Symposium on Signal Processing and Its Applications*, volume 2, pages 430–433, 2001.

[49] A. Beghdadi and B. Pesquet-Popescu. A new image distortion measure based on wavelet decomposition. In *Proceedings of Seventh International Symposium on Signal Processing and Its Applications*, volume 1, pages 485–488, 2003.

[50] A. Samet, M. Ayed, N. Masmoudi, and L. Khriji. New perceptual image quality assessment metric. *Asian Journal of Information Technology*, 4(11):996–1000, 2005.

[51] J. Munkberg, P. Clarberg, J. Hasselgren, and T. Akenine. High dynamic range texture compression for graphics hardware. *ACM Transactions on Graphics*, 25(3):698–706, 2006.

[52] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli. New full-reference quality metrics based on HVS. In *Proceedings of the Second International Workshop on Video Processing and Quality Metrics*, 2006.

[53] Z. Wang and A. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9:81–84, 2002.

[54] N. Ponomarenko, F. Silvestri, K.Egiazarian, M. Carli, and V. Lukin. On between-coefficient contrast masking of dct basis functions. In *Proceedings of Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2007.

[55] D. Chandler and S. Hemami. VSNR: A wavelet base visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9):2284–2298, 2007.

[56] Z. Wang and Q. Li. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5):1185–1198, 2011.

[57] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[58] Z. Wang, E. Simoncelli, and A. Bovik. Multi-scale structural similarity for image quality assessment. In *Proceedings of the 37$^{th}$ IEEE Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1398–1402, 2003.

[59] Z. Wang and E. Simoncelli. Translation insensitive image similarity in complex wavelet domain. In *Proceedings of IEEE international conference on acoustics, speech and signal processing*, volume 2, pages 573–576, 2005.

[60] C.-L. Yang, W.-R. Gao, and L.-M. Po. Discrete wavelet transform-based structural similarity for image quality assessment. In *Proceedings of IEEE 15$^{th}$ International Conference on Image Processing*, pages 377–380, 2008.

[61] A. Brooks and X. Zhao. Structural similarity quality metrics in a coding context: Exploring the space of realistic distortions. *IEEE Transactions on Image Processing*, 17(8):1261–1273, 2008.

[62] C. Li and A. Bovik. Content-partitioned structural similarity index for image

quality assessment. *Signal Processing: Image Communication*, 25(7):517–526, 2010.

[63] G.H. Chen, C.L. Yang, L.M. Po, and S.L. Xie. Edge-based structural similarity for image quality assessment. In *Proceedings of IEEE International Conference in Acoustics, Speech and Signal Processing*, volume 2, pages 933–936, 2006.

[64] G.H. Chen, C.L. Yang, and S.L. Xie. Gradient-based structural similarity for image quality assessment. In *Proceedings of IEEE International Conference on Image Processing*, pages 2929–2932, 2006.

[65] X. Gao, T. Wang, and J. Li. A content-based image quality metric. In *Proceedings of the $10^{th}$ international conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, pages 231–240, 2005.

[66] W. Dong, Q. Yu, C. Zhang, and H. Li. Image quality assessment using rough fuzzy integrals. In *Proceedings of the $27^{th}$ International Conference on Distributed Computing Systems Workshops - Supplements*, pages 1–5, 2007.

[67] E. Lam and K. Loo. An image similarity measure using homogeneity regions and structure. In *Proceedings of the SPIE*, volume 6978, pages 680811–1–680811–9, 2008.

[68] A. Kolaman and O.Yadid-Pecht. Quaternion structural similarity: a new quality index for color images. *IEEE Transactions on Image Processing*, 21(4):1526–1536, 2012.

[69] C. Moxey, S. Sangwine, and T. Ell. Hypercomplex correlation techniques for

vector images. *IEEE Transactions on Signal Processing*, 51(7):1941–1953, 2003.

[70] A. Shnayderman, A. Gusev, and A. M. Eskicioglu. An SVD-based gray-scale image quality measure for local and global assessment. *IEEE Transactions on Image Processing*, 15(2):422–429, 2006.

[71] M. Narwaria and W. Lin. Objective image quality assessment with singular value decomposition. In *Proceedings of 5$^{th}$ International Workshop Video Processing and Quality Metrics for Consumer Electronics*, 2010.

[72] H. Sheikh, A. Bovik, and G. de Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, 14(12):2117–2128, 2005.

[73] H. Sheikh and A. Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2):430–444, 2006.

[74] L. Zhang, L. Zhang, and X. Mou. RFSIM: A feature based image quality assessment metric using riesz transforms. In *Proceedings of IEEE 17$^{th}$ International Conference on Image Processing*, pages 321–324, 2010.

[75] L. Zhang, L. Zhang, X. Mou, and D. Zhang. FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8):2378–2386, 2011.

[76] D. Kim and R. Park. New image quality metric using the Harris response. *IEEE Signal Processing Letters*, 16(7):616–619, 2009.

[77] D. Kim, H. Han, and R. Park. Gradient information-based image quality metric. *IEEE Transactions on Consumer Electronics*, 56(2):930–936, 2010.

[78] A. Agrawal, R. Raskar, and R. Chellappa. Edge suppression by gradient field transformation using cross-projection tensors. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 2301–2308, 2006.

[79] D. Kim and R. Park. Image quality measure using the phase quantization code. *IEEE Transactions on Consumer Electronics*, 56(2):937–945, 2010.

[80] D. Kim and R. Park. Image quality assessment using the amplitude/phase quantization code. *IEEE Transactions on Consumer Electronics*, 56(4):2756–2762, 2010.

[81] G. Zhai, W. Zhang, Y. Xu, and W. Lin. LGPS: Phase based image quality assessment metric. In *Proceedings of IEEE Workshop on Signal Processing Systems*, pages 605–609, 2007.

[82] Y. Wang, T. Jiang, S. Ma, and W. Gao. Image quality assessment based on local orientation distributions. In *Proceedings of Picture Coding Symposium*, pages 274–277, 2010.

[83] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893, 2005.

[84] ITU-R. Methodology for the subjective assessment of the quality for television pictures. ITU-R Recommendation BT.500-11, 2002.

[85] B. Rosner. Percentage points for a generalized ESD many-outlier procedure. *Technometrics*, 25(2):165–172, 1983.

[86] A. van Dijk, J. Martens, and A. Watson. Quality assessment of coded images using numerical category scaling. In *Proceedings of SPIE*, volume 2451, pages 90–101, 1995.

[87] A. Moorthy, K. Seshadrinathan, R. Soundararajan, and A. Bovik. Wireless video quality assessment: a study of subjective scores and objective algorithms. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(4):587–599, 2010.

[88] K. Soundararajan, R. Soundararajan, A. Bovik, and L. Cormack. Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, 19(6):1427–1441, 2010.

[89] H. Sheikh, M. Sabir, and A. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 15(11):3441–3452, 2006.

[90] M. Chambah, S. Ouni, M. Herbin, and E. Zagrouba. Toward an automatic subjective image quality assessment system. In *IS&T/SPIE Electronic Imaging*, pages 72420E–72420E. International Society for Optics and Photonics, 2009.

[91] M. Mahy, L. Van Eycken, and A. Oosterlinck. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Journal of Color Research and Application*, 19:105–121, 1994.

[92] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti. TID2008 - A database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10:30–45, 2009.

[93] Y. Yang. Colour edge detection and segmentation using vector analysis. Master's thesis, Electrical and Computer Engineering, University of Toronto, Toronto, Canada, 1995.

[94] R. Floyd and L. Steinberg. An adaptive algorithm for spatial grey scale. In *Proceedings of the Society of Information Display*, volume 17, pages 75–77, 1976.

[95] P. Stucki. MECCA - A multiple-error correcting computation algorithm for bilevel image hardcopy reproduction. Technical Report RZ1060, IBM Research Laboratory, Zurich, Switzerland, 1981.

[96] J. Jarvis, C. Judice, and W. Ninke. A survey of techniques for the display of continuous tone pictures on bilevel displays. *Computer Graphics and Image Processing*, 5(1):13–40, 1976.

[97] A. Robertson. The cie 1976 color-difference formulae. *Color Research and Application*, 2:7–11, 1977.

[98] R. Gentile, E. Walowit, and J. Allebach. Quantization and multilevel halftoning of color images for near-original image quality. *Journal of the Optical Society of America A*, 7(6):1019–1026, 1990.

[99] A. Pentland. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:661–674, 1984.

[100] A. Pentland. Shading into texture. *Artificial Intelligence*, 29:147–170, 1986.

[101] Y. Liu and Y. Li. Image feature extraction and segmentation using fractal dimension. In *Proceedings of International Conference on Information, Communications and Signal Processing*, pages 975–979, 1997.

[102] H. Potlapalli and R. Luo. Fractal-based classification of natural textures. *IEEE Transactions on Industrial Electronics*, 45(1):142–150, 1998.

[103] L. Kaplan. Extended fractal analysis for texture classification and segmentation. *IEEE Transactions on Image Process*, 8(11):1572–1585, 1999.

[104] S. Novianto, Y. Suzuki, and J. Maeda. Near optimum estimation of local fractal dimension for image segmentation. *Pattern Recognition Letters*, 24(1):365–374, 2003.

[105] H. Matsumoto and K. Sasazaki. Color image compression with vector quantization. In *Proceedings of IEEE Conference on Soft Computing in Industrial Applications*, pages 84–88, 2008.

[106] G. Zorpette. Fractals: not just another pretty picture. *IEEE Spectrum*, 25(10):29–31, 1988.

[107] N. Sarkar and B. Chaudhuri. An efficient approach to estimate fractal dimension of textural images. *Pattern Recognition*, 25(9):1035–1041, 1992.

[108] N. Sarkar and B. Chaudhuri. An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Transactions on Systems, Man and Cybernetics*, 24(1):115–120, 1994.