# SCREEN CONTENT IMAGE QUALITY ASSESSMENT AND PROCESSING

**GUO HENG**

**School of Electrical & Electronic Engineering**

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

**2021**

# Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

17-03-2021
.....
Date


.....
Guo Heng

# Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

17-03-2021

.....

Date

KKM

.....

Prof. Ma Kai-Kuang

# Authorship Attribution Statement

This thesis contains material from 1 paper accepted at conferences in which I am listed as an author.

Chapter 2 is published as H. Guo, K. Ma and H. Zeng, "A Log-Gabor Feature-Based Quality Assessment Model for Screen Content Images," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 4499-4503, doi:10.1109/ICIP.2019.8803491.

The contributions of the co-authors are as follows:

- Prof. Ma provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by both Prof. Ma and Prof. Zeng.
- I co-designed this work with Prof. Ma and performed all the experimental verification at the School of Electrical and Electronic Engineering in NTU.
- The dataset used in experimental verification was provided by Prof. Zeng.

17-03-2021 .....

Date

.....

Guo Heng

# Acknowledgements

I wish to express my greatest gratitude to my advisor my supervisor, Professor Kai-Kuang Ma from the School of Electrical and Electronic Engineering. He has spent munificent efforts to supervise me on this project. With his encouragement, I had the opportunity to explore various directions and approaches on solving technical problems, which is truly beneficial to build up a collective understanding of problems from a broad perspective and from the most fundamental point of view. Without his supervision, this thesis would not be possible.

I would also like to express my gratitude to our collaborator Professor Zeng Huanqiang from Huaqiao University, and my colleague Mr Ni Zhangkai from NTU, for the helpful and fruitful discussions. They have given tremendous support and valuable scientific insights to my research. Their wealth of knowledge contributed greatly to this thesis.

Special thanks to my wife and my family, who deeply inspired me in every aspect. Their love and support are always the greatest inspiration. Their care, understanding and encouragement helped me keep calm and carry on during any tough days.

I would like to reserve a special praise to EEE and NTU for the funding and platform they provided to enable truly integrated and meaningful research. I would like to thank Professor Chen Shoushun, with whom I started my PhD life six years ago. And I would like to express my deepest appreciation to Professor Tang Yi, and Mrs Chua-Goh Wei Jiuan, who provided wholehearted help throughout my PhD life.

# Contents

# Summary

With rapid development of information technology, screen content images (SCIs), which contain a mixture of natural images together with texts, charts, symbols, and/or computer generated graphics, are often encountered in various multimedia applications. SCIs usually contain a large amount of sharp edges, hence the perception of SCIs are quite different from that of natural images. However, how to display the SCIs with a high quality remains an unmet need in image processing, and even the state-of-the-art technology is not able to handle SCIs well considering the extensive amount of sharp edges in these images. To address this problem, three fundamental image processing techniques for screen content color images are investigated in this thesis; i.e., image quality assessment, image blurriness estimation, and image lightness estimation. Among these techniques, image quality assessment is a fundamental tool to objectively evaluate the quality of the screen content color image from the observer's point of view. Image blurriness estimation distinguishes the blurred regions and sharp regions of an image, providing a fundamental guidance to image enhancement algorithms to improve the perceptual quality of images. Finally, the image lightness estimation method is a generic tool for extracting the lightness component from an image, which is able to help facilitate a variety of image processing tasks that manipulate the lightness of images. Our contributions in these areas are respectively summarized as follows.

For image quality assessment, a log-Gabor feature-based model (LGFM) is proposed to conduct objective evaluations of screen content images (SCIs). The key success of our method lies in the fact that the response of log-Gabor filter is more consistent with the perception of visual cortex in our human visual system (HVS). Two remarkable characteristics of the log-Gabor filter are (i) zero response at the

DC, and (ii) stronger response at high frequencies. In LGFM, the log-Gabor filter is used to extract salient features from the luminance of the reference SCIs and that of the distorted SCIs to measure their degree of similarity. Extensive simulation results have shown that our proposed LGFM gives the most consistent quality assessment score with the human perception, compared to other state-of-the-art IQA models.

For image blurriness estimation, we first proposed a Rapid transform that operates in the binary domain with Boolean operators and then it is combined with local binary pattern (LBP) feature descriptor, to form a cyclic rotational invariant local binary pattern, named as RT-LBP in short. The usefulness of the proposed RT-LBP is demonstrated with an LBP based image blurriness estimation method to generate an accurate pixel-wise blurriness map to segment the blurred and sharp regions in an image. With extensive experiments, the proposed RT-LBP is proved to deliver accurate image blurriness estimation for blurriness-guided image enhancement applications, with a speed that outperforms many existing methods.

For image lightness estimation, a color-adaptive lightness compensation method to the CIELAB color space is proposed. Based on the observations that color has a huge impact on human's perceived lightness, a color-dependent lightness compensation method is developed to generate an accurate lightness estimation of each image pixel. The amount of compensation is calculated based on a set of physiological visual response data collected from real human observers. Extensive experiments on various test images have demonstrated that the proposed method can serve as the fundamental to various image processing algorithms, including the previously mentioned two applications of image quality assessment and image quality enhancement in this thesis, to deliver superior perceptual quality. Just using the proposed lightness estimation method in this work to replace the conventional ones, and keeping other parts of the algorithms unchanged, significant improvements in the final results can be seen.

# List of Figures

# List of Tables

# Chapter 1

# Introduction

With rapid development and popularity of computers, computer-generated imagery and data are often encountered in our daily lives. Consequently, the contents of digital images circulating around the Internet are no longer just limited to conventional natural scene images, such as human portrait, landscape and cityscape, close-up shootings on animals or still objects, to name a few, as illustrated in Fig. 1.1 (a). In fact, the contents of digital images nowadays have a mixture of different sources. Besides the above-mentioned natural imageries, the contents could include those generated by computers, such as graphics, texts, charts, hand-writings and hand-drawings, and even some special symbols or patterns, including logos, bar codes, QR codes, and more. The image containing such kind of mixtures of image contents is denoted as the *screen content images* (SCIs), which is the focused image content to be investigated in this thesis.

## 1.1 Motivation

In recent years, the SCIs have attracted more attention, since they are widely deployed in various multimedia applications and services, such as remote screen


FIGURE 1.1: Examples of natural scene images and screen content images.

sharing, cloud gaming, online conferences and education, electronic brochures, Internet live video broadcasting, and so on. It has been observed that the image content characteristics of SCIs is quite different from that of the natural scene images. Refer to Fig. 1.1 for some SCI examples, the image contents of SCIs tend to have sharp edges, high-contrast and vivid colors in certain regions on the image field, whereas the rest of large areas are filled with a smooth texture as background. All these make SCIs containing a lot of sharp edges, which is particularly challenging to be handled for image processing.

The sharp edges in SCIs bring up an unmet need in image processing. Even many of the state-of-the-art technologies are not able to handle SCIs well. Our

research is fully motivated by this. In this thesis, three fundamental topics dealing with screen content color images are covered: they are image quality assessment, image blurriness estimation, and image lightness calculation, which serve as the fundamentals in this research area. In the next section, a brief introduction of these research objectives in each topic is presented. The contributions of this thesis will be discussed in Section 1.3.

## 1.2 Objectives

**Image Quality Assessment for Screen Content Images**

Measuring the image quality in an objective and accurate way is of fundamental importance for image processing applications. In this work, the goal is to develop an image quality assessment model for screen content color images that can objectively evaluate the perceptual quality of the image and provide the evaluation results that is highly in accordance with the human visual system (HVS).

**Image Blurriness Estimation for Image Enhancement**

Image enhancement techniques are invented to improve the perceptual quality of images. In this work, we are going to propose an image local feature descriptor and an image blurriness estimation method. The extracted image blurriness information functions as a guidance to conventional image enhancement algorithms, to further increase the quality of the image enhancement operations.

**Image Lightness Estimation on Screen Content Images**

In many image processing algorithms, the operations such as feature extraction, feature enhancement, are all performed on the lightness component of the image. Hence an accurate estimation and calculation of an image's lightness component becomes the foundation to various image processing methods. In this work, we

will take the understanding of human's reaction to light and color into account, to propose an accurate image lightness estimation method with the aid of physiological experimental results involving human observers.

## 1.3 Major Contributions of the Thesis

In this thesis, three contributions have been made in the above-mentioned research areas, and can be summarized respectively as follows:

- *Image Quality Assessment for Screen Content Images*: An image quality assessment (IQA) model, called the log-Gabor feature-based model (LGFM), is proposed in this work to conduct objective evaluations of screen content images (SCIs). From the standpoint of perceptually-related signal representation, the log-Gabor filter is clearly more suitable than the classical Gabor filter, since the output of the log-Gabor filter is more consistent with the perception of visual cortex in our human visual system (HVS). This could be due to the following two remarkable characteristics of the log-Gabor filter: (i) zero response at the DC, and (ii) stronger response at high frequencies. Both are highly desirable for evaluating the distorted SCI with respect to its reference SCI, as the subtle differences between them can be well captured with the aid of these two properties. In our proposed LGFM, the log-Gabor filter is used to extract salient features from the luminance of the reference SCIs and that of the distorted SCIs to measure their degree of similarity. Extensive simulation results have shown that our proposed LGFM is highly consistent with the human perception, compared to other state-of-the-art IQA models.

- *Image Blurriness Estimation for Image Enhancement*: In this work, we first proposed a Rapid transform that operates in the binary domain with Boolean

operators. The proposed Rapid transform has the properties of cyclic rotation invariance, reflection invariance and feature dimension reduction. Combining the proposed Rapid transform with local binary pattern (LBP) feature descriptor, a cyclic rotational invariant local binary pattern is then proposed, named as RT-LBP in short. The usefulness of the proposed RT-LBP is demonstrated using an LBP based image blurriness estimation method to generate an accurate pixel-wise blurriness map to segment the blurred and sharp regions in an image. With extensive experiments, the proposed RT-LBP is proved to deliver accurate image blurriness estimation for blurriness-guided image enhancement applications, with a speed that outperforms many existing methods.

- *Color-Adaptive Image Lightness Estimation*: In this work, an image lightness estimation method based on human perception is proposed. The proposed method incorporates the physiological visual response from real human observers. Based on the observations that color has a huge impact on human's perceived lightness, a color-dependent lightness compensation method is developed to generate an accurate lightness estimation of each image pixel. Extensive experiments on various test images have demonstrated that the proposed method can serve as the fundamental to various image processing algorithms, including the previously mentioned two applications of image quality assessment and image quality enhancement in this thesis, to deliver superior perceptual quality. Using the proposed lightness estimation method in this work to replace the conventional and other existing color spaces, and keeping other parts of the algorithms in the IQA and enhancement unchanged, significant improvements in the final results can be seen.

## 1.4 Organization of the Thesis

The remaining of this thesis is organized as follows.

Chapter 2 provides a literature review over the related SCI quality assessment and processing topics.

In Chapter 3, a detailed introduction of the image quality assessment on screen content images as well as the proposed log-Gabor feature-based image quality assessment model is presented. Experimental results from extensive simulations are provided to show that the proposed log-Gabor model clearly outperforms existing state-of-the-art IQA methods and delivers the most consistent assessment score with human eye's perception.

In Chapter 4, the proposed binary domain Rapid transform algorithm is elaborated in detail first, followed by the proposed cyclic rotational invariant local binary pattern feature descriptor, RT-LBP. The effectiveness of the proposed RT-LBP descriptor is demonstrated with an LBP-based image blurriness estimation, which clearly demonstrate that the proposed fast RT-LBP algorithm is able to deliver superior image blurriness estimation performance while having significant advantage in terms of algorithmic simplicity and computational efficiency.

In Chapter 5, a novel color space for calculating the image lightness is proposed, which takes into account the understanding of human's reaction to light and color. The proposed method can benefit applications such as image quality assessment, and being adopted in image enhancement applications as well. Performance evaluations of the proposed lightness method is demonstrated in terms of both subjective comparison on the lightness estimation result and its contribution to image quality and image enhancement algorithms. The proposed lightness estimation method is able to deliver state-of-the-art image lightness estimation accuracy and serves as the fundamental to various higher level image processing tasks.

Finally, Chapter 6 concludes this thesis and provides a brief discussion on several promising future research directions.

# Chapter 2

# Literature Review

In this chapter, existing works related to the research topics covered in this thesis have been reviewed. In Section 2.1, various image quality assessment models, specifically those with an emphasis on exploiting the properties of the human visual system, have been reviewed. In Section 2.2, image blurriness estimation methods and local binary pattern are discussed. In Section 2.3, fundamentals on color science and color spaces are presented. Section 2.4 summarizes the findings of the literature review. An overview of the related works and their relationship to our research is depicted in Fig. 2.1.

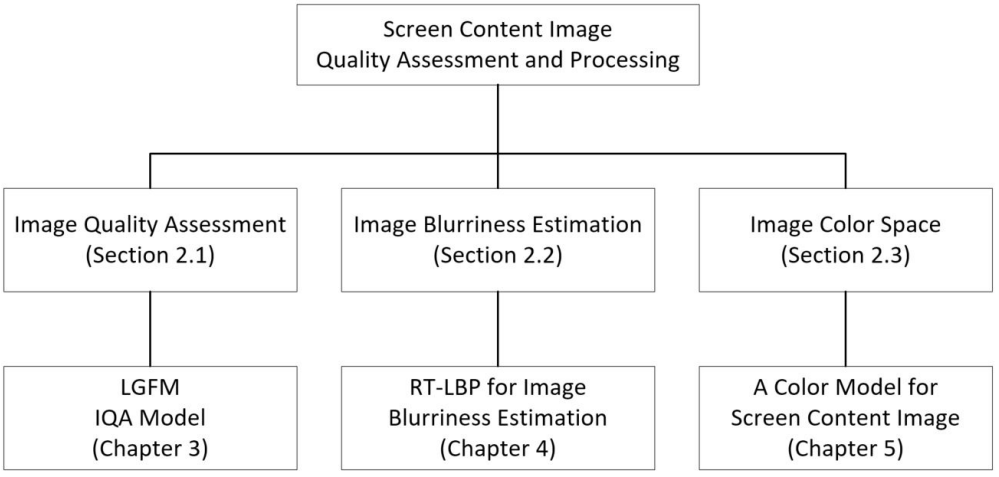



FIGURE 2.1: An overview of the related work and its relationship to our research.

## 2.1 Literature Review on Screen Content Image Quality Assessment

The quantitative assessment of image quality is an important topic in the pipeline of image acquisition, processing, transmission and display. For that, an objective evaluation metric is essential to conduct accurate image quality assessment (IQA). With the advent of multimedia technology, *screen content images* (SCIs) have been encountered in various information processing applications. A typical SCI could further include a mixture of texts, symbols, graphs, charts, and animation images, besides the natural images. Consequently, SCIs tend to contain sharp edges, and the evaluation of a SCI's image quality by the human visual system (HVS) [1, 2] becomes very different from the way conducted on the *natural* images. Therefore, it is essential to develop an accurate perceptual-based IQA model for the evaluation of SCIs.

Conventionally, the *mean square error* (MSE) and *peak signal-to-noise ratio* (PSNR) have been exploited to measure the visual quality of natural images. However, they are often inconsistent with the judgments made by the human visual system (HVS) [1, 2] since they are initially designed for analysis of electrical signal quality. As the human eyes are the final receiver of the images, it is important to development an IQA model in accordance with the human visual system (HVS).

A lot of physiological and psychophysical observations on how human visual perceives light and color [3–6] have been made over the past century, and many experiments [7–9] have been designed to investigate and model how the human visual system processes information and why the visual system works in such way.

Based on findings that the HVS is highly sensitive to the image's structural information, the development of *structural similarity* (SSIM) [2] was invented as a milestone development of the IQA models and further led to various improved models,

such as *information weighted SSIM* (IW-SSIM) [10], followed by *visual information fidelity* (VIF) [11], *feature similarity* (FSIM) [12], *gradient similarity* (GSIM) [13], *gradient magnitude similarity deviation* (GMSD) [14], *visual saliency-based index* (VSI) [15], *structural contrast-quality index* (SC-QI) [16], *perceptual similarity* (P-SIM) [17], *structural variation analysis* (SVQI) [18], and *multi-scale difference of gaussian* (MDOGS) [19]. All of these models exploit various properties of the HVS (e.g., edges [14], salient regions [15], micro- and macro-structures [17–19], and so on).

Since the inception of convolutional neural network (CNN), machine learning based approach has become a popular way to tackle various problems in computer vision and image processing, such as image classification and object detection, to name a few. In general, machine learning based approach involves two stages: (1) extracting salient features and (2) designing learning algorithms for performing regression or classification. Such machine learning based methods have also been developed to solve the IQA problem as well. Zuo *et al.* [20] directly fed the sequentially cropped image patches into CNNs for feature extraction, followed by computing quality score regression. Other deep learning-based models, which jointly learn feature representation and quality prediction, also showed some promising results in IQA [21–24].

However, Bosse *et al.* [25] has pointed out that as the number of parameters to be trained in deep networks is usually very large, but current publicly available quality-annotated image databases are rather small. This makes training of a deep network a challenging task and the machine learning based models would not perform well if it is trained with one dataset but being tested with another dataset. There is room for optimization in terms of feature dimensionality and balancing the ratio between network parameters. This remains a fundamental problem for all machine learning based IQA models. Furthermore, the computational complexity of machine learning based approaches are much higher compared to conventional

approaches, which makes it difficult to be a fundamental and generalized tool that can be easily exploited by other image-based applications. Hence machine-learning-based approaches have less attraction to us.

In summary, to develop an IQA model that is highly in accordance with the HVS still remains a challenging yet important task.

## 2.2 Literature Review on Image Blurriness Estimation and Local Binary Pattern

### 2.2.1 Blurriness Estimation

The purpose of blurriness estimation it to separate the blurred and non-blurred regions in an image. It is worth mentioning that blurriness estimation and salient object detection are not the same. Salient object detection aims to find out the most visually distinctive objects and the most important or noticeable objects in an image [26]. This is subject to humans understanding on the images content. Besides blurriness and sharpness of the object in images, color, brightness, contrast, etc., and even the meaning of the objects, all contribute to salient object detection. Blurriness estimation is an objective description on image patches to tell whether it is blurred or sharp, and how blur or how sharp it is. Some images from DUTS dataset [27] for salient object detection tasks and Fixations in Webpage Images (FiWI) dataset [28] for predicting webpage saliency are selected in Fig.2.2 to explain the similarity and difference between image blurriness estimation and salient object detection.

In some cases, the sharpest region is indeed the salient object in the image. For example, the butterfly shown in Fig.2.2 (a). However, this is not always the case that the sharpest objects in image are all the salient objects. As illustrated in

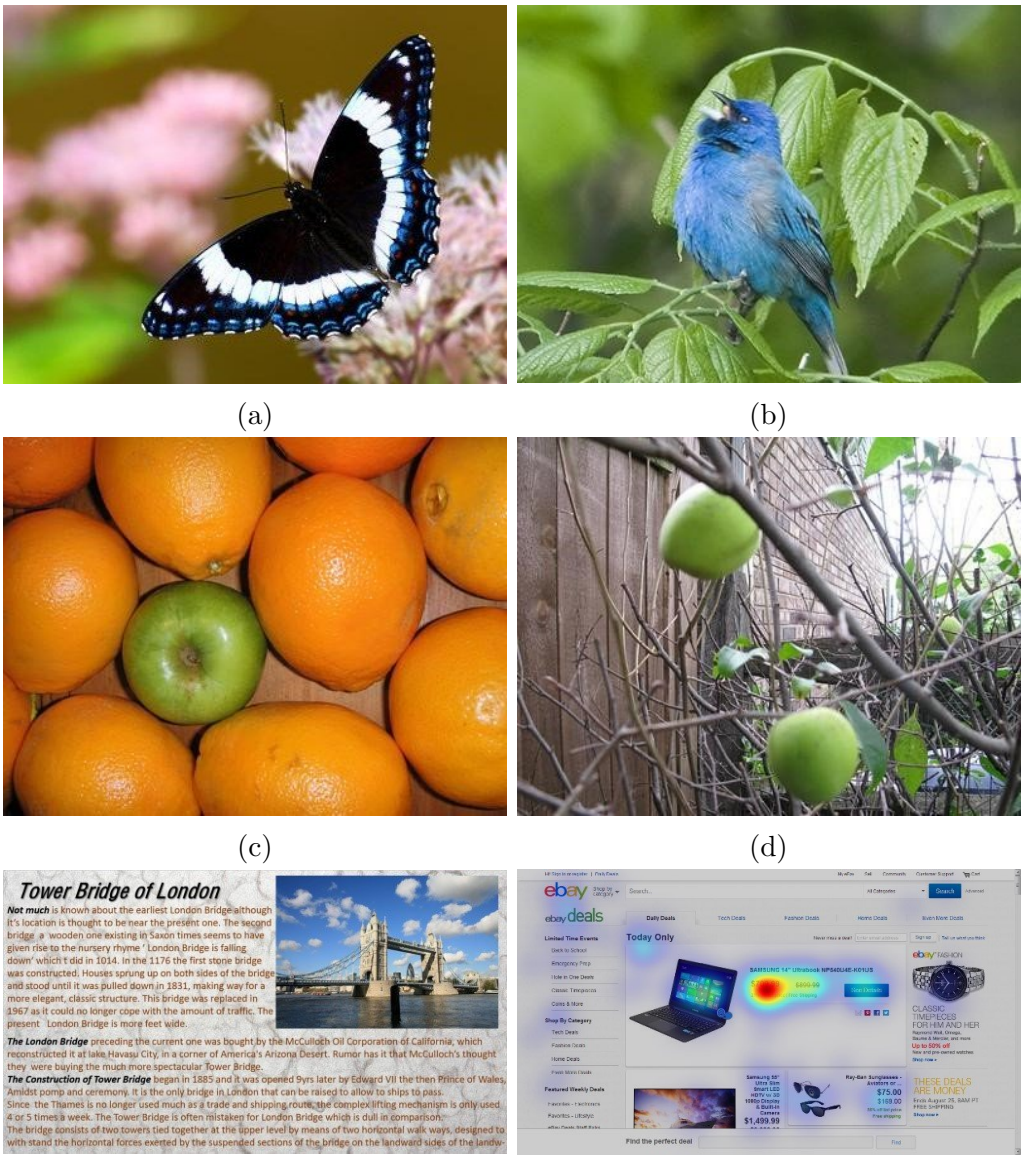



FIGURE 2.2: Examples on the similarity and difference between image blurriness estimation and salient object detection.

Fig.2.2 (b), only the blue bird is considered as the salient object even though the green leaves are also clear and sharp. As for the case in Fig.2.2 (c), the green apple is the only salient object among the oranges, despite all the fruits in this image are of similar blurriness level. For the case in Fig.2.2 (d), the green blurred fruits on the tree are salient objects, although the focal point of this image falls on the background brick wall and the fruits are blurred due to out of focus.

When the salient object detection goes to SCIs which are usually rich in text,

pictures, charts, etc., from the image blurriness point of view, a SCI is full of sharp salient stimuli everywhere. Conventional blurriness based salient object detection methods would fail. As illustrated in Fig.2.2 (e), although all the text in this SCI Tower Bridge is sharp, the most eye-catching objects in this image is the photo of Tower Bridge on the upper right corner. As for the case in Fig.2.2 (f), an SCI of a shopping webpage (overlaid with saliency heat map from user survey) from FiWI dataset, the price of the goods is the most salient object in this image. Shen *et al.* [28] thus proposed a new way for salient object detection for webpages considering the different characteristics of SCIs.

Blurriness may be helpful in salient objection object detection, but not always. Blurriness estimation is one of the fundamental image processing techniques that can serve as a part in salient object detection and many other image processing applications.

Existing blurriness estimation methods have adopted a wide variety of approaches. The vast majority of these methods are performed in the intensity and gradient domain of the images. The assumption is that the appearance of edges is affected by blur, and accordingly those methods estimate blurriness by using various edge-related models. For example, Marziliano *et al.* [29] measure the blurriness according to edge widths calculated in $8 \times 8$ blocks. Liu *et al.* [30] used Sobel edge detector to find the edge features, and then combine these features via a circular back propagation neural network system to estimate the blurriness. Li *et al.* [31] compare the kurtosis of blocks of dominant edge pixels, which captures the "peakedness" of the edge pixels distribution to tell how blur or how sharp the edge appears.

Other methods have used spectral information to estimate blurriness in the frequency domain. For example, Shaked *et al.* [32] use the discrete Fourier transform (DFT) to estimate blurriness based on the ratio of high-pass to low-pass energy of the the spatial derivative of each row and column. Vu *et al.* [33] measure the

slope of the magnitude spectrum and the total spatial variation of image blocks to give a local blurriness estimation. In some other works, blurriness has also been estimated based on the kurtosis of the energy spectrum [34], and the entropy of the energy spectrum [35].

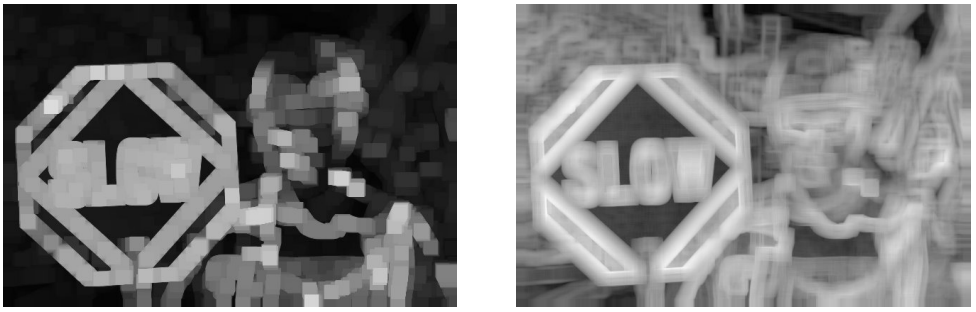
More recently, just noticeable blur estimation (JNBE) [36] is developed based on a sparse representation that exploits the sparseness measured from an image patch's sparse representation to infer its blurriness and it is believed to deliver superior performance.

Given those blurriness measurement metrics, Yi *et al.* [37] noticed that most of the existing methods fail to estimate the blurriness of "almost-smooth" textures accurately. The "almost-smooth" textures are texture of low contrast in general but has more detail or pattern than pure smooth regions, for example the lemon in Fig.4.8 (a) and the yellow road sign region in Fig.4.9 (a). This has greatly drawn our attention since such "almost-smooth" texture largely exists in the screen content images that we are particularly interested in in this thesis. Extensive experiments found that none of the above-mentioned metrics are well suited for measuring low contrast sharp regions, such as the almost smooth region in the example. This is because the low contrast region has very small intensity variance which leads to low gradient and low frequency response. This drawback of existing blurness estimation methods is further shown in Fig. 2.3. The low contrast yellow region of the road sign does not have a correct response for all measures as illustrated in the figure except for the Local Binary Pattern (LBP) based method proposed by Yi *et al.* [37].

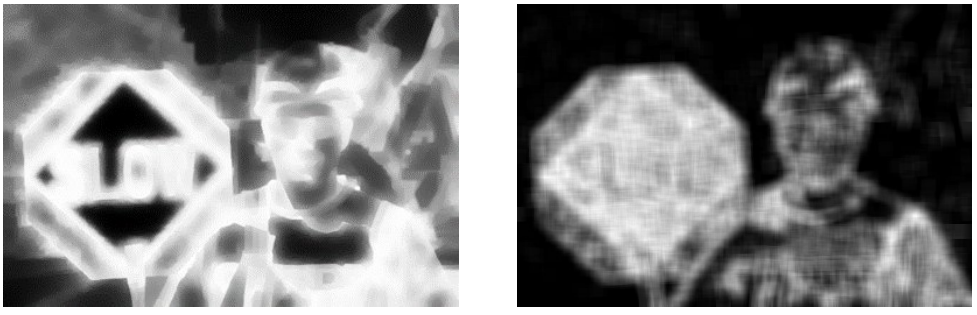
Attracted by the low computational complexity inherited from LBP and the relatively good blurriness estimation performance, we are interested to further explore in the LBP-based image blurriness estimation approaches.


(a) Original image (b) Ground truth

(c) Marziliano *et al.* [29] (d) Li *et al.* [31]




(e) Vu *et al.* [33] (f) Zhang *et al.* [34]



(g) JNBE [36] (h) Yi *et al.* [37]

FIGURE 2.3: Comparison of some existing blurriness estimation methods. All of them fail to measure the blurriness of the yellow region of the road sign correctly, except for (h).

### 2.2.2 Local Binary Pattern

Local binary pattern (LBP) has the attractive properties such as low computational complexity for real-time applications, less sensitive to illumination change [38], and low algorithmic complexity for ease of VLSI implementation [39] and in general good descriptive power. Due to these benefits, LBP and its modern variants have been continuously developed and exploited in many image processing works in the past few years, including texture classification [40–44], dynamic texture recognition [45, 46], face detection and recognition [47, 48], fingerprint and palmprint recognition [49, 50], object detection [51–53], and image quality assessment [54–56].

LBP is computed from the analysis of a 3×3 patch of the local neighbourhood around a central pixel. The LBP binary code is generated by thresholding the neighbourhood pixel values with the pixel value of the centre. If the tested pixel value is below that of the central pixel, then it is labelled as 0, otherwise it is assigned the value 1. With this technique, 256 ($2^8$) possible binary patterns could be obtained.

The obtained binary pattern is then multiplied by weights given to the corresponding pixel location. Summing the obtained results gives the measure of the LBP value. The process to convert LBP binary pattern to LBP value of a pixel is equivalent to converting a binary number to decimal. The corresponding LBP feature map is then constructed by assigning the LBP value ($X$) to each of the central pixel in every patch. Equation (2.1) shows the logic on how to assign binary values to the neighbouring pixels and the LBP value of the local neighbourhood is calculated as Equation (2.2). The process is further illustrated in Fig. 2.4. An example of an input image and its corresponding LBP feature map is shown in Fig. 2.5.

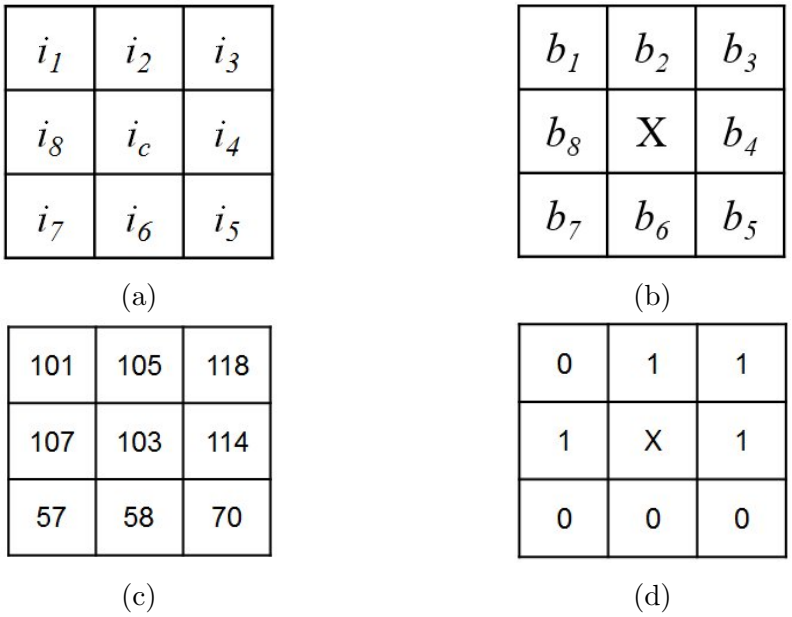


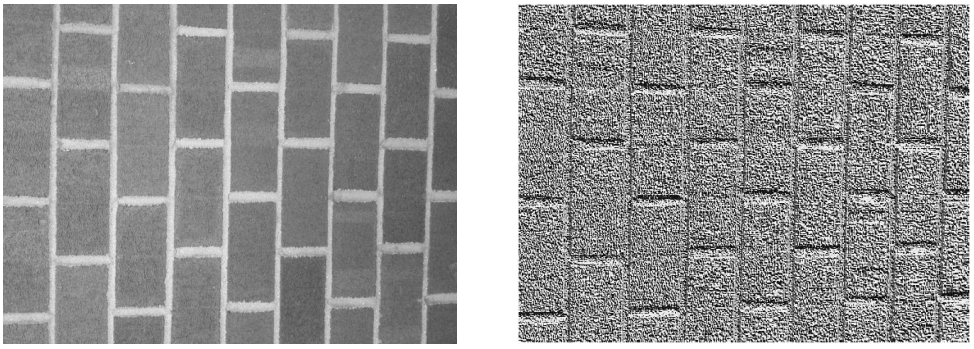

FIGURE 2.4: Illustration of generating LBP: (a) a 3×3 block of an input image. $i_x$ is intensity value. (b) obtained binary pattern of the patch, thresholded by the central pixel's intensity $i_c$. (c) and (d) give an numerical example: in the 3×3 block illustrated in (c), the threshold is 103, LBP pattern = 01110001, LBP value $(X) = 113$.

$$b_i = \begin{cases} 0, \text{ if } i_x < i_c; \\ 1, \text{ otherwise.} \end{cases} \tag{2.1}$$

$$X = \sum_{x=1}^{8} b_x \times 2^{x-1} \tag{2.2}$$

where $b_x$ is the obtained binary code, $i_x$ is the pixel intensity at position $x$ and $i_c$ is the central pixel intensity.

The description above then brings about an important question that is not addressed in the initial invention of LBP – which location of the surrounding pixels should be chosen as the starting point of the LBP binary pattern? In other words, which location should be given the highest weight when calculating the LBP value of the neighbourhood? Further more, should the LBP binary code be formed by taking the patterns in a clockwise order or counter-clockwise order?



(a) An input image "brick" (b) LBP feature map of (a)

FIGURE 2.5: Illustration of LBP feature map.

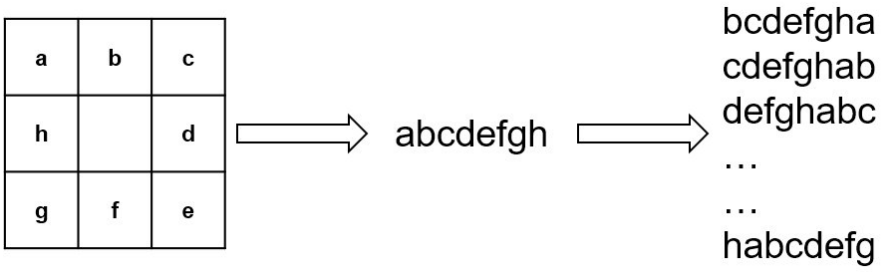



FIGURE 2.6: Illustration of the procedure to generate rotation invariant LBP pattern.

Later on, a rotation invariant version of LBP, always denoted as $LBP^{ri}$ in short [57], was proposed that partially answered the above question. In his approach, the rotation invariant version of LBP was obtained by performing the circular bitwise right shift of each LBP code until it reaches the minim the value. For instance, the 8-bit LBP codes 10000011, 00111000, and 00000111 all belong to one rotation invariant LBP category, and will be all mapped to the minimum code 00000111, as this LBP code has the minimum LBP value $X = 7$ in decimal. The process is demonstrated in Fig. 2.6. In this way, number of unique patterns is reduced to 36.

Ojala *et al.* [57] further found that not all rotation invariant patterns sustain rotation equally well, and so they proposed to use uniform patterns only, which are a subset of the rotation invariant patterns and named them as $LBP^{riu}$. A pattern is defined to be uniform if the circular sequence of bits contains no more than two

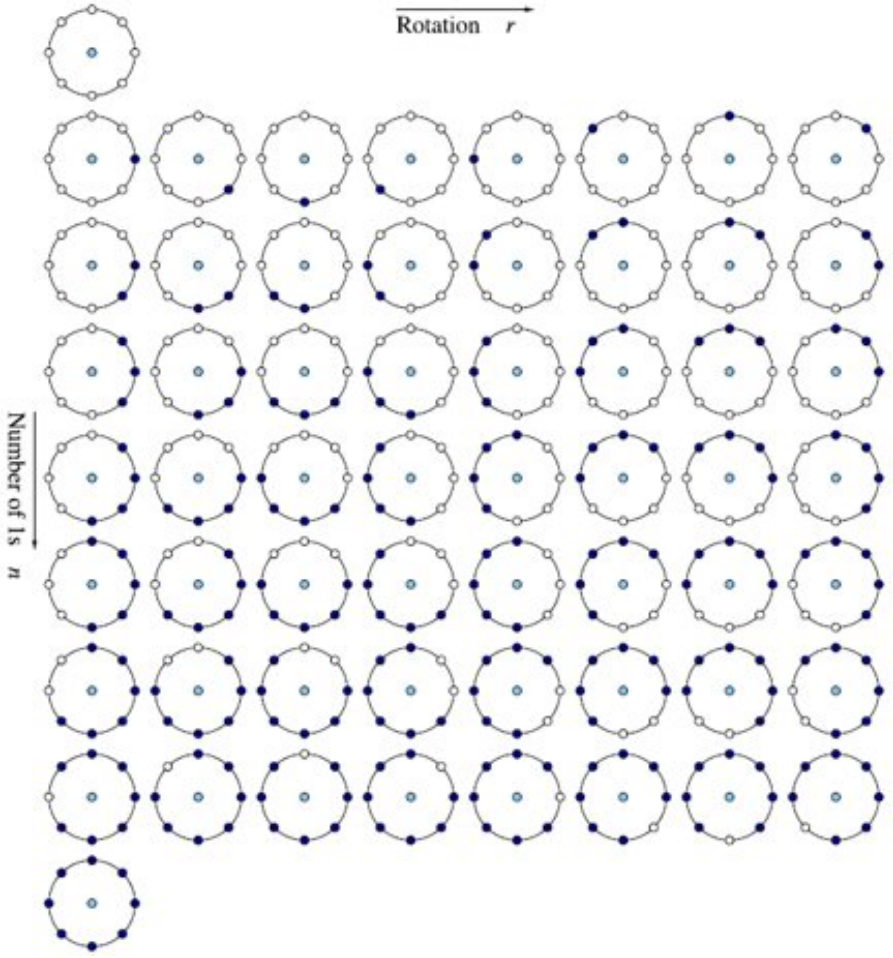



FIGURE 2.7: Illustration of the uniform LBP patterns [57].

transitions from one to zero, or zero to one. The remaining patterns are all considered as non-uniform and grouped into one single pattern category. This further reduces the number of unique patterns to 10 (for 8-bit LBP), that is, 9 uniform patterns, plus 1 category consisting of all non-uniform patterns. The 9 uniform patterns are shown in Fig. 2.7. In this figure, the white hollow circle represents LBP code 0 and the black solid dot represents LBP code 1. Each row in this figure represents one category of the uniform pattern, and shows all LBP patterns contained in each category. All other LBP patterns that are not shown in this figure are non-uniform patterns by definition, and they are grouped together as the $10^{th}$ category of LBP pattern.

Although both of the rotation invariant LBP, named $LBP^{ri}$, and the rotation

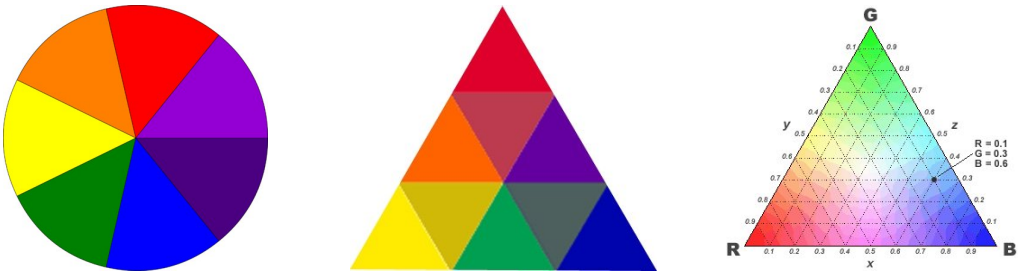
invariant uniform LBP, named $LBP^{riu}$, has been proved to perform well in many applications, the definition of both descriptors is still in an ad-hoc way and does not answer our above-mentioned questions on LBP very well.

This gives us a motivation to further develop a local descriptor that has a more solid scientific support. A *Cyclic Rotation Invariant Local Binary Pattern* for Blurriness Estimation will hence be proposed and introduced in Chapter 4 in this thesis.

## 2.3 Fundamentals of Color Spaces

In modern computer science and display technologies, a color is represented as tuples of numbers in order to transmit and reproduce such color, in both analog and digital domains. A color model is a mathematical model describing how colors can be represented. For example, a color can be represented as triples in the RGB color model, which is the most commonly used color model in computer graphics. The same color can be represented as quadruples in the CMYK color model, the most popular color model used in printing industries. When a model is associated with a precise description of how the components are to be defined and calculated, a "color space" is created. This chapter describes the color spaces in which human color vision can be modeled.

The research on finding a perfect color space to describe what a color is and how exactly it can be represented and reproduced has been going on for hundreds of years. Numerous models and systems have been developed for this purpose. In 1660, Sir Issac Newton observed that white light could be broken down into the colors of the rainbow using a prism. Based on his observation from the experiment, the Newton's color wheel (see Fig. 2.8 (a)) was created to describe the components of colors and the order of primary colors. However, Newton did not describe how a specific color could be generated by mixture of different primary colors.




(a) Newton's color wheel (b) Goethe's color triangle (c) Maxwell's color triangle

FIGURE 2.8: Newton's color wheel[1], Goethe's color triangle[2]and Maxwell's color triangle[3].

Johann Wolfgang von Goethe made a big leap in the color theory in the early 1800's and proposed his Goethe's color triangle in [58], as illustrated in Fig. 2.8 (b). This model opened up the possibilities for investigating color mixing. In Goethe's color triangle, red, yellow, and blue are chosen as three primary colors. They are arranged at the vertices of a triangle. The other sub-divisions of the triangle are grouped into secondary and tertiary triangles, where the secondary triangle colors represent the mix of the two primary triangles to either side of it, and the tertiary triangle colors represent the mix of the primary triangle adjacent to it and the secondary triangle directly across from it.

Sir James Clerk Maxwell, whose most notable achievement was to formulate the classical theory of electromagnetic waves, had also developed a color triangle from his understandings of the electromagnetic theory of light in 1861 [59]. Maxwell chose red, green, and blue as primary colors, the same primary colors as in the modern RGB color model. He presented the world's first demonstration of colour photography by projecting three black-and-white photographs of a scene taken through red, green and blue filters respectively and superimposing the three projection onto a screen using three projectors equipped with the same color filters. Maxwell has built the foundation for modern RGB color model.

[1]https://commons.wikimedia.org/wiki/File:Disque_newton.png
[2]https://i.imgur.com/lPchM.jpg
[3]http://www.appstate.edu/ steelekm/classes/psy3215/Color/maxwelltriangle.gif

There was not any standard for measurement and representation of color until 1931. The International Lighting Commission (usually abbreviated as CIE for its French name, Commission Internationale de l'Eclairage) made the first attempt to establish a world standard to quantitatively model the links between physical colors in electromagnetic spectrum and physiologically perceived colors in human color vision. Since then, the CIE 1931 RGB color space and CIE 1931 XYZ color space were created based on a series of experiments involving a total of 17 human observers.

With the advent of the computer age, the RGB color space becomes the most popular candidate to describe colors. Most televisions, computer displays, and projectors produce colors by combining primary colors in varying intensities. Such mixtures in RGB color space can reproduce a wide variety of colors. However, the relationship between the constituent amounts of red, green, and blue light and the resulting color is unintuitive, especially for inexperienced users. For example, if we are given an orange color as illustrated in Fig 2.9 (a), whose RGB values are R = 237, G = 125, B = 49, and we want to make it less *saturated*, or in other words less *colorful* as Fig 2.9 (b), what we need to do is to increase R, G and B by 7, 52, and 82, respectively. The intention of making it less colorful and the action of increasing all color components sounds contrary to each other; in fact, it is quite confusing, unintuitive, and unsynchronized. The relationship among how much R, G and B changes respectively is also unclear and hard to tell by intuition. Thus, attempts have been made to create some other color space models that can define color relationships the same way as the human eye perceives, leading to the creation of HSL color space, HSV color space, CIELAB color space, and so on. Lightness has become one of the most important definitions of color-making attributes since then [60].


(a) (b)

FIGURE 2.9: Example on making a color less colorful in sRGB color space. In this example, (b) is less orange than (a). The actual R, G, and B values that made up the respective orangeness are shown inside the image.

### 2.3.1 CIE 1931 RGB and XYZ Color Spaces

The CIE 1931 RGB color space and CIE 1931 XYZ color space can be considered as the world's first standard to quantitatively model the links between physical colors in electromagnetic spectrum and physiologically perceived colors in human color vision. Created by the International Commission on Illumination (CIE) in 1931 [61], these two models resulted from a series of independent experiments done in the late 1920s by William David Wright using 10 observers [62] and John Guild using 7 observers [63].

In human visual system, the cone cells of human retina are responsible in the sensation of color. There are three different types of cone cells in our human visual system, namely long-wavelength-sensitive (L), middle-wavelength-sensitive (M), and short-wavelength-sensitive (S) cones [64]. As their names suggest, these three kinds of cone cells differ in their sensitivity to different wavelengths of light. Based on experiments involving human observers, CIE proposed the standard photonic luminous efficiency function CIE 1924 $V(\lambda)$ to quantify the relationship between the wavelength of the light and the luminance perceived by our human visual system. The experimental results were combined into the specification of the CIE RGB color space, from which the CIE XYZ color space was derived.

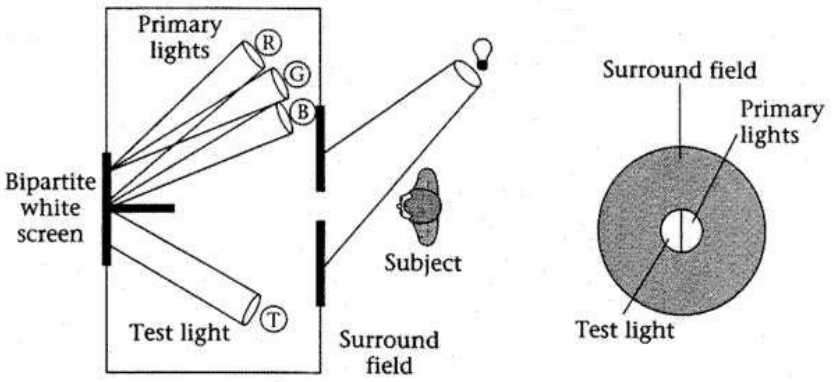



FIGURE 2.10: The color matching experiment setup [63].

### CIE 1931 RGB Color Space

The experiment setup to obtain the CIE 1931 RGB color space is illustrated in Fig. 2.10. In this experiment, a reference test color is projected onto one side of the bipartite screen, while an adjustable color of a mixture of three primary colors, red, green, and blue, is projected onto the other side of the screen. The three primary colors are with fixed chromaticity, but the observer can adjust the three primary color's brightness separately. The observer views the bipartite field and adjusts the intensities of the three primary lights until the mixture test color matches the appearance of the test light. Then the amount needed for each primary color to match a certain reference test color is recorded. Based on this experiment, CIE proposed the CIE 1931 RGB color matching functions as shown in Fig 2.11, which describes the amounts of each primary colors needed to match the monochromatic test color at the wavelength shown on the horizontal axis.

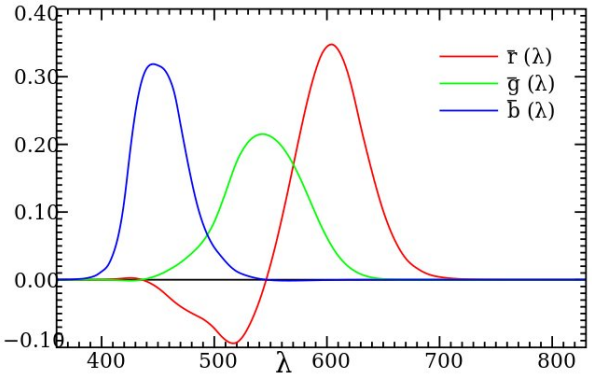



FIGURE 2.11: The CIE 1931 RGB color matching functions [63].

Hereinto, the RGB tri-stimulus values for a color in the CIE 1931 RGB color space can be then calculated as:

$$\begin{aligned} R &= \int_0^\infty S(\lambda)\bar{r}(\lambda)d\lambda, \\ G &= \int_0^\infty S(\lambda)\bar{g}(\lambda)d\lambda, \\ B &= \int_0^\infty S(\lambda)\bar{b}(\lambda)d\lambda. \end{aligned} \tag{2.3}$$

where $S(\lambda)$ is the *spectral power distribution* (SPD) that describes the power per unit area per unit wavelength of an illumination and detailed numbers can be found at [65, Appendix 5]. $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ are the color matching functions as described in Fig. 2.11.

### CIE 1931 XYZ Color Space

Having developed the RGB model for the human visual system using the CIE RGB matching functions, the members of the CIE wished to develop another color space that would relate to the CIE RGB color space but with simplification in computation, since in the age of 1930's computations were done by hand. Hence, CIE XYZ color space is introduced as a linear transformation from the CIE RGB color space and the standardized transformation settled by the CIE was as follows [65]:

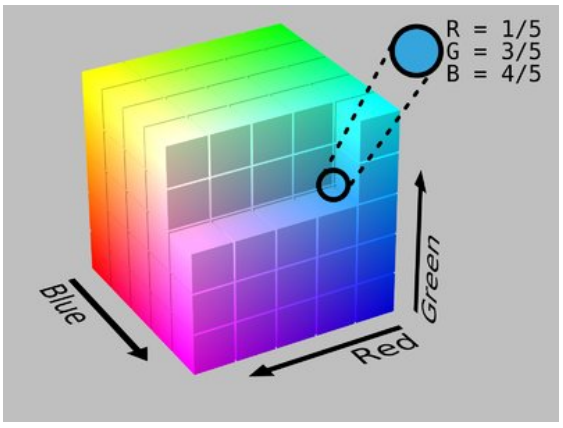



FIGURE 2.12: Visual representation of the sRGB color space[4].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49000 & 0.31000 & 0.20000 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00000 & 0.01000 & 0.99000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{2.4}$$

The CIE 1931 color spaces opened the era to quantitatively model the links between physical colors in electromagnetic spectrum and physiologically perceived colors in human visual system. Both of the CIE 1931 color spaces are still widely used today.

### 2.3.2 sRGB Color Space

sRGB is one of the RGB color spaces that was initially proposed by Hewlett Packard Enterprise Company (in short, HP) and Microsoft [66] for their products as well as for the Internet. Later on, it was standardized by the International Electrotechnical Commission [67] and is often exploited as the "default" color space, especially for the 24-bit color depth true-color images whose pixels are stored in 8-bit integers per color channel for R, G, and B, respectively. The 24-bit true-color is used in virtually every computer and hand-phone display as the dominant image storage format.

[4]https://commons.wikimedia.org/wiki/User:Datumizer#/media/File:RGB_Cube_Show_lowgamma_cutout_b.png

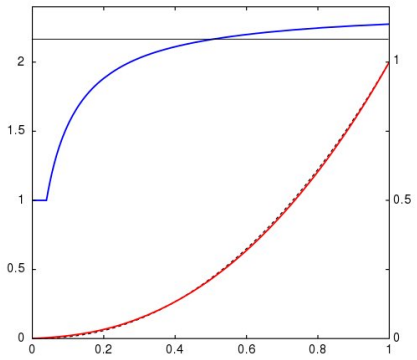



FIGURE 2.13: Plot of the actual sRGB intensities to be displayed versus the sRGB numerical values stored in image file (red line), and the effective gamma value (blue line) at each sRGB numerical value[5].

sRGB also defines a nonlinear transfer function, linking between the desired intensity of the R, G and B primary colors to be displayed and the actual number stored as the image file, named as *gamma correction*. The function is defined in terms of the power-law expression; that is,

$$f(u) = u^{\gamma}, \tag{2.5}$$

where a non-negative real input value $u$ is raised to the power of $\gamma$ to get the desired output value. The gamma correction function for the color space sRGB is defined as

$$f(u) = \begin{cases} \frac{u}{12.92}, & u \leqslant 0.04045, \\ (\frac{u+0.055}{1.055})^{2.4}, & \text{otherwise.} \end{cases} \tag{2.6}$$

where $u$ denotes the sRGB component values in the range of [0, 1]. For images stored in 8-bit per color channel, such values $u$ can be obtained by simply dividing each component value by 255.

The gamma correction function is visualized in Fig. 2.13. The red curve is the plot of the sRGB intensities ($y$-axis) versus sRGB numerical values ($x$-axis), while the

[5]https://upload.wikimedia.org/wikipedia/commons/e/ef/SRGB_gamma.svg

blue curve represents the effective gamma value $\gamma$ ($y$-axis) in (2.5) at each point of sRGB numerical value.

Images can be transformed from sRGB color space to CIE XYZ color space using the following linear transformation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950277 \end{bmatrix} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix}. \tag{2.7}$$

where $R_{linear}$, $G_{linear}$ and $B_{linear}$ are the sRGB values after gamma correction. Colors are represented in a cubic geometry as illustrated in Fig. 2.12.

In sRGB color space, the image lightness is the attribute of a visual sensation according to which an area appears to emit more or less light as defined by Fairchild in [60]. It is usually calculated as the numerical average of the three RGB components and denoted as intensity $I$, i.e.:

$$I = avg(R, G, B) = \frac{1}{3}(R + G + B). \tag{2.8}$$

### 2.3.3 YIQ and YUV Color Spaces

YIQ and YUV color spaces are initially designed for color TV broadcasting applications. Both have one *luminance* component (i.e., $Y$) and two *chrominance* components each (i.e., $I$ and $Q$ in the former or $U$ and $V$ in the latter), respectively [68]. In both color spaces, their luminance component $Y$ has exactly the same definition, which is calculated as the weighted average of red, green, and blue primary colors. For the chrominance components, the I-Q plane and the U-V plane have different coordinate systems. To be more specific, the I-Q plane is a 33°of

rotation relative to the U-V plane. Color space conversion from RGB to YIQ or to YUV are given by the following linear transformations; that is,

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{2.9}$$

and

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & -0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{2.10}$$

The luminance component $Y$ is very important in the old days when it was designed because it enabled then existing black-and-white TV receivers to receive the TV broadcasted programs with color pictures on top of those black-and-white TV programs in high majority. Component $Y$ represents the brightness of the image while components $UV$ or $IQ$ components are the add-ons to the black-and-white brightness signal to make the TV pictures being perceived with colors. Nowadays, these two color spaces still beneficial to image processing and computer vision algorithms such as color-to-grayscale image transformation.

### 2.3.4 HSL and HSV Color Spaces

Designed in the 1970's by computer graphics researchers to more closely align with the way human vision perceives color and light, there are two alternative color spaces, called the HSL and HSV [69]. The HSL stands for *hue*, *saturation*, and *lightness*, while the HSV stands for *hue*, *saturation*, and *value*, respectively

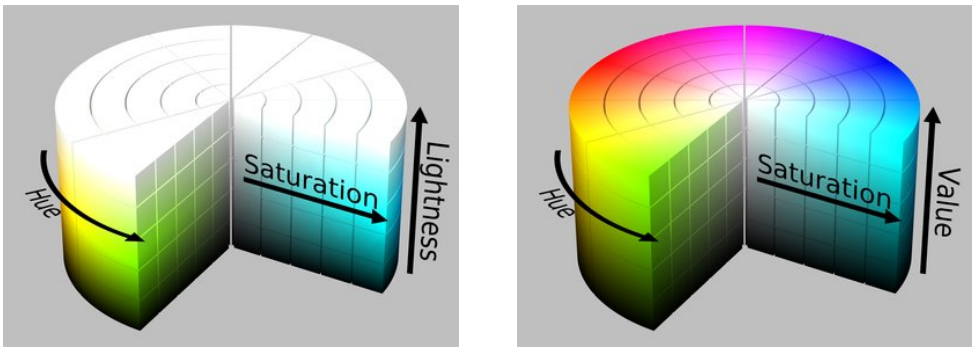

FIGURE 2.14: An illustration of the relationship between hue value and color appearance[6].

*Hue* is technically defined as "the *degree* to which a stimulus can be described as similar to or different from the stimuli that are described as red, green and blue [70]. From the visual perception point of view, the relationship between the *hue* and color is demonstrated in Fig. 2.14. Hue starts from the red primary at 0°, passing through the green primary at 120°, the blue primary at 240°, and finally to red at 360°, which is a wrap-back to the red primary at 0°as expected.

*Saturation* describes the *colorfulness* of a color. For colors with the same hue and brightness, the higher the saturation is, the deeper the color appears. For colors with *zero* saturation, they are named *neutral*, *achromatic*, or *gray* colors ranging from white at lightness 1 (i.e., value = 1) to black at lightness 0 (i.e., value = 0). Hue values for such neutral colors are not defined, and are often assigned a hue of 0°for convenience of color image representation.

Both HSL and HSV color spaces have a cylindrical geometry. In the HSL color space, each hue corresponds to one radial slice [see Fig. 2.15 (a)), which surrounds a central axis of neutral colors ranging from black at the bottom to white at the top. Fully saturated colors are placed around a circle in the middle of the cylinder at a lightness value of $V = 1/2$ by its definition, where a lightness value $V$ of 0 or 1 represents fully black or white, respectively. While in the HSV color space [see Fig. 2.15(b)], fully saturated colors are placed around a circle in the top of the cylinder and black is placed at the bottom of the cylinder. Thus, the fully saturated colors have lightness of 0.5 in HSL color space, while in HSV they have a value of 1. The HSL and HSV cylinders are illustrated in Fig. 2.15.

[6]https://i.stack.imgur.com/YOBFy.png




(a) (b)

FIGURE 2.15: Visual representation of HSL and HSV color space [69].

The relationship between HSL and RGB as well as that of HSV and RGB can be described as follows [71]. To make the definitions easier to write, the maximum value of the RGB components, the minimum value of the RGB components, and the chroma value (i.e., the range of the RGB components) are denoted as $M$, $m$, and $C$, respectively.

$$M = max(R, G, B), \tag{2.11}$$

$$m = min(R, G, B), \tag{2.12}$$

$$C = range(R, G, B) = M - m. \tag{2.13}$$

With the definitions as described above, we can define hue angle $H$ mathematically as:

$$H' = \begin{cases} Undefined, & \text{if} C = 0, \\ \frac{G-B}{C} mod 6, & \text{if} M = R, \\ \frac{B-R}{C} + 2, & \text{if} M = G, \\ \frac{R-G}{C} + 2, & \text{if} M = B. \end{cases} \tag{2.14}$$

$$H = 60^\circ \times H'. \tag{2.15}$$

As mentioned above, neutral colors (i.e., with $C = 0$) are often assigned a hue of $0°$ for convenience of representation.

The lightness or value in the HSL and HSV color spaces are defined in a simple way. In the HSL color space, lightness $L$ is defined as the average of the largest and smallest RGB color components, i.e. the mid-range of the RGB components. While in the HSV color space, value $V$ is defined as the largest component of a color.

$$L = mid(R, G, B) = \frac{1}{2}(M + m). \tag{2.16}$$

$$V = max(R, G, B) = M. \tag{2.17}$$

Saturation is defined as the colorfulness in proportion to its brightness. The saturation in HSL color space $S_L$ and the saturation in HSV color space $S_V$ can be simply calculated by the following equations respectively:

$$S_L = \begin{cases} 0, & L = 0\text{or}1, \\ \frac{C}{1-|2L-1|}, & \text{Otherwise.} \end{cases} \tag{2.18}$$

$$S_V = \begin{cases} 0, & V = 0, \\ \frac{C}{V}, & \text{Otherwise.} \end{cases} \tag{2.19}$$

### 2.3.5 CIELAB Color Space

CIELAB is an improved model based on CIE 1931 XYZ color space, which was designed to be perceptually uniform with respect to human color vision [72]. The goal is that the same amount of numerical changes in the CIELAB values corresponds to about the same amount of visually perceived changes. Hence, intuitively, CIELAB

would be more suitable for being exploited in human-eye perception-related image applications such as image quality assessment.

The CIELAB color space is a nonlinear transformation on the CIE XYZ space with the purpose of making the space more uniform with respect to human perceived color differences. Given an image recorded in the CIE XYZ color space, its values in the CLELAB components can be calculated based on the following equations:

$$\begin{aligned} L^* &= 116 f(Y/Y_n) - 16, \\ a^* &= 500 \left[ f(X/X_n) - f(Y/Y_n) \right], \\ b^* &= 200 \left[ f(Y/Y_n) - f(Z/Z_n) \right]. \end{aligned} \tag{2.20}$$

where $L^*$ is the lightness, and $a^*$ and $b^*$ are the indexes of redness-greenness and yellowness-blueness, respectively. The ($X_n$, $Y_n$, $Z_n$) is a tri-stimulus value of the reference white point under a reference illumination specified by the CIE, and $f$ is the pointwise nonlinearity function defined as:

$$f(x) = \begin{cases} x^{\frac{1}{3}}, & x > 0.008856, \\ 7.787x + 16/116, & x < 0.008856. \end{cases} \tag{2.21}$$

The reference white point is a set of tri-stimulus values or chromaticity coordinates that serve to define the color "white". For example, a photograph of a sheet of white paper taken under tungsten lights may appear relatively distorted and appeared as orange color when compared it to a photo taken under daylight. So it is important to choose a proper reference white point when transforming an image from RGB color space to CIELAB. For that, the CIE standard determines illuminant D50 as the default white point. It simulates warm daylight at sunrise or sunset with correlated color temperature of 5003 K, which is also known as *horizon light*.

Given an image in the CIELAB color space, its chroma value $C^*$ and hue angle $h$ can be calculated as

$$C^* = \sqrt{(a^*)^2 + (b^*)^2}, \tag{2.22}$$

$$h = \arctan\left(\frac{b^*}{a^*}\right). \tag{2.23}$$

Intuitively, CIELAB would be the most suitable color space to be used when estimating the lightness of an image, since it is designed to follow the way how human eyes perceive the images lightness.

### 2.3.6 The Gaussian Color Model

The Gaussian color model [73], or being denoted as the LMN color space in some literatures (e.g., [74]), has gained its popularity over the past few years, as this color model is derived from a physical reflectance model to describe how much light is reflected from an object in color images.

The conversion from the RGB color space to the LMN color space is calculated based on the following formula:

$$\begin{bmatrix} L \\ M \\ N \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{2.24}$$

where $L$ denotes the luminance and $M$, and $N$ are two chrominance components. The $L$ component in this LMN color space describes the intensity of the reflected light by the objects in an image.

In summary, a lot of attempts have been made to create color space models that can describe the color relationships in the same way as the human eye perceives.

Lightness has hence become one of the most important definitions of color-making attributes. To find an accurate lightness estimation method with human eye's physiological-perception taken into consideration is attracting our interest.

## 2.4 Summary of Literature Review

In summary, three fundamental issues in screen content color image have been reviewed in Chapter 2.

In Section 2.1, quantitative assessment of image quality is discussed. To develop an IQA model that is highly in accordance with the HVS yet simple and fast still remains a challenging yet important task.

In Section 2.2, various image blurriness estimation methods have been reviewed. LBP-based approach has particularly attracted our interest due to its computational simplicity and relatively good performance on the "almost-smooth" regions in the image. This gives us a motivation to further develop a local descriptor for image blurriness detection application.

In Section 2.3, fundamentals on color science has been presented and various popular color spaces have been reviewed. To create a color model for screen content images with an accurate lightness estimation in accordance with human eyes physiological-perception is attracting our interest.

# Chapter 3

# Image Quality Assessment Model for Screen Content Images

## 3.1 Introduction

An *image quality assessment* (IQA) model, called the *log-Gabor feature-based model* (LGFM), for conducting objective evaluations of *screen content images* (SCIs) is proposed in this chapter. From the standpoint of perceptual-based signal representation, it has been observed that the log-Gabor function is more suitable than the many classical image feature extraction functions, due to the fact that the output of the log-Gabor function is more consistent with the perception of visual cortex in our human visual system (HVS). Extensive experimtnral results have shown that our proposed LGFM outperforms other state-of-the-art IQA models and is highly consistent with the human perception on the IQA task.

The rest of this chapter is organized as follows. In Section 3.2 , a proposed LGFM IQA model for conducting objective evaluations of the SCIs is presented. In Section 3.3, extensive experiments based on the proposed model and several state-of-the-arts are conducted and compared. Section 3.4 concludes this chapter.

## 3.2 Proposed IQA Model Based on Log-Gabor Features

### 3.2.1 Motivation

Among the existing conventional/non-maching-learning-based IQA algorithms for the evaluation of SCIs, the *Gabor feature-based model* (GFM) proposed by Ni *et al.* [74] delivers state-of-the-art performance, which was motivated by the fact that the collection of the Gabor filter's output values are quite consistent with the judgemet responded by our HVS [75].

In [76], Field further suggested that the *log-Gabor* is a better signal representation scheme than the Gabor, since the image under the log-Gabor representation presented on the logarithmic frequency scale will yield a Gaussian distribution. Another support on the use of the log-Gabor function for shape recognition application can be found in Massot *et al.* [77]. Inspired by these, a new log-Gabor feature-based IQA model is proposed in this chapter, namely the *log-Gabor feature-based model* (LGFM), which exploits two remarkable characteristics of the log-Gabor filter as follows.

There are multiple remarkable characteristics possessed by the log-Gabor-based signal representation. First, the log-Gabor function removes DC; hence, the filtered results has *no* DC component. This is especially desirable in IQA, since DC does not offer any minute differences for conducting IQA but only captures the global information of the image such as image's overall luminance or brightness. For example, the shape of a bridge should be perceived the same by our HVS regardless it is viewed at night or in the daytime. Second, compared with the Gabor function whose bandwidth is constantly limited to one octave, the bandwidth of the log-Gabor function is adjustable. Third, the frequency range yielded by the log-Gabor function is wider than that of the Gabor function. All the above-mentioned are

advantageous from the standpoint of signal representation, especially considering that the image content of the distorted SCIs and their corresponding reference SCIs are quite similar. Consequently, their subtle differences can only be reflected through its higher frequency range.

### 3.2.2 Model Overview

The block diagram of the proposed log-Gabor feature-based IQA model for conducting the evaluation of a distorted SCI's quality compared to its reference SCI is shown in Fig. 3.1. In the first stage, both the reference SCI and the distorted SCI (respectively denoted by $r$ and $d$ in their subscripts) are converted from the originally recorded RGB color space to another new color space, the LMN, according to the following formula [73]:

$$\begin{bmatrix} L_\zeta(x,y) \\ M_\zeta(x,y) \\ N_\zeta(x,y) \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R_\zeta(x,y) \\ G_\zeta(x,y) \\ B_\zeta(x,y) \end{bmatrix}, \tag{3.1}$$

where $\zeta$ denotes either the reference SCI $r$ or the distorted SCI $d$. Components $L_\zeta(x,y)$, $M_\zeta(x,y)$, and $N_\zeta(x,y)$ denote the luminance and two chrominance components at each pixel location $(x,y)$, respectively.

In the second stage, the log-Gabor filter is used to generate the log-Gabor feature maps on the luminance component of the reference SCI and that of the distorted SCI individually. In the third stage, the degree of similarity between these two sets of log-Gabor feature maps will be measured. The same filtering and similarity measurement will be applied to each chrominance component likewise. In the last stage, these measurements will be fused to arrive at the final IQA evaluation score. The details of each stage will be further described in the following sub-sections, respectively.

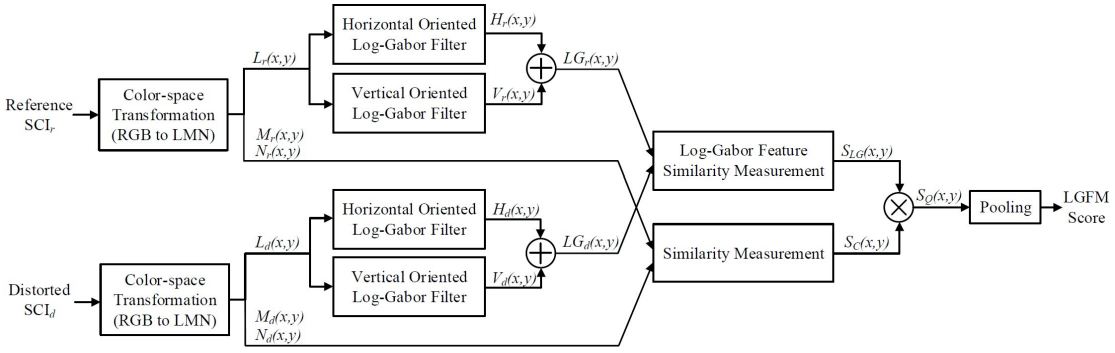



FIGURE 3.1: The block diagram of the proposed *log-Gabor feature-based* image quality assessment (IQA) model for the evaluation of the quality of distorted screen content images (SCIs).

### 3.2.3 Why the Log-Gabor Filter?

The Gabor filter has been one of the most popular choices in many signal processing algorithms. In image processing, its filtered result are fairly consistent with our HVS perception. The Gabor filter $G(\omega)$ is defined as [78]

$$G(\omega) = n_c \cdot \exp\left[-\frac{\sigma^2}{2}(\omega - \omega_0)^2\right], \tag{3.2}$$

where $\omega_0$ is the center frequency of the filter and $n_c$ is a normalization constant. The bandwidth $\beta$ of $G(\omega)$ is defined as the frequency difference between frequencies $\omega_1$ and $\omega_2$, such that the amplitude of the response is half of the peak. It is expressed in octaves as [78]

$$\beta = \log_2\left(\frac{\omega_2}{\omega_1}\right), \tag{3.3}$$

where $\omega_1$ and $\omega_2$ are the solutions of $G(\omega_1) = G(\omega_2) = 0.5 \cdot n_c$. The bandwidth $\beta$ can be also expressed in terms of $k_\beta$ as [78]

$$\beta = \frac{\ln\left(\frac{k_\beta + \sqrt{2\ln(2)}}{k_\beta - \sqrt{2\ln(2)}}\right)}{\ln(2)}, \tag{3.4}$$

where $k_\beta = \sigma \cdot \omega_0$, or alternatively,

$$k_\beta = \sqrt{2\ \ln 2}\left(\frac{2^\beta + 1}{2^\beta - 1}\right). \tag{3.5}$$

It is important to note that the Gabor filter has a *non-zero* DC component $G(0)$, which can be obtained as

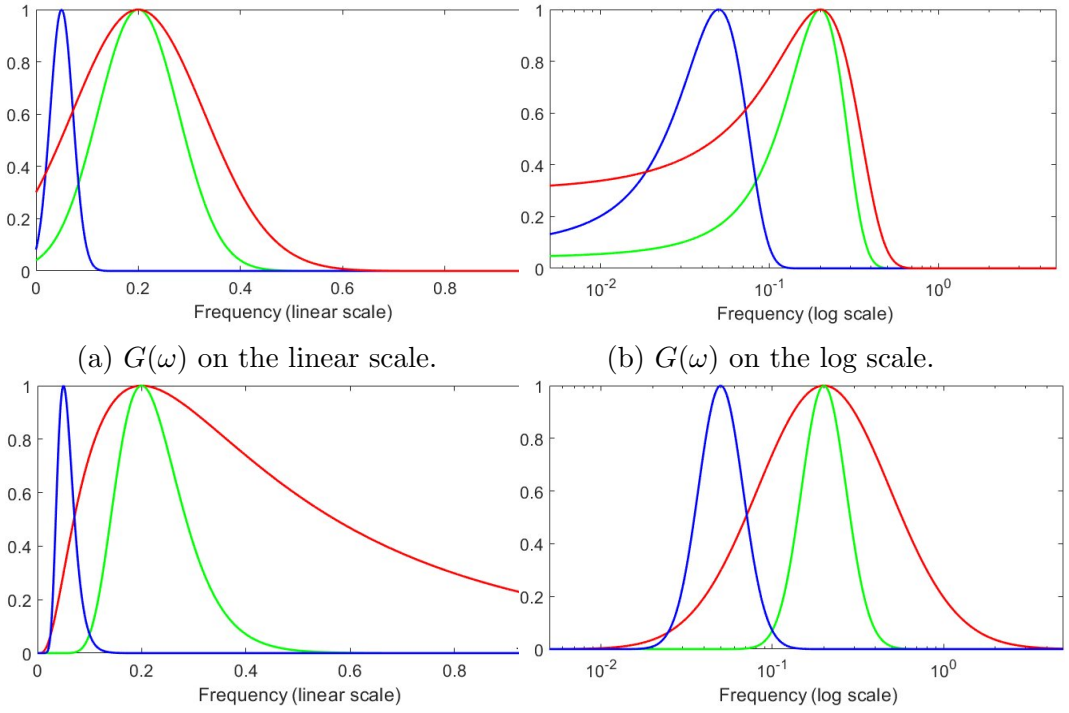
$$G(0) = \exp\left(-\frac{\sigma^2\omega_0^2}{2}\right) = \exp\left(-\frac{k_\beta^2}{2}\right). \tag{3.6}$$

From (3.5) and (3.6), it can be observed that the $G(0)$ is a function of $\beta$. Three examples are demonstrated in Fig. 3.2 (a) and (b), which are presented in the linear scale and the logarithmic scale, respectively. From which, one can appreciate how the value of $\beta$ affects $G(0)$: (Blue and Green) $\beta = 1$; (Red) $\beta = 3$. Note that the Gabor filter's DC response becomes non-negligible when the bandwidth is set to 3 octaves or above, as illustrated in Fig. 3.2 (b), which has a set of zoom-in plots of Fig. 3.2 (a) in the neighborhood of the DC. Ronse [79] and Heitger *et al.* [80] also pointed out this issue on the non-negligible DC response. As highlighted previously, the DC component of an image represents a global information (such as illumination), which is considered irrelevant to the information on edge and image structure from the perceptual viewpoint. For those applications that exploit Gabor filter, the bandwidth has been suggested to set at 0.7 octave, in order to maintain negligible DC response [78].

To overcome the above-mentioned dilemma on the Gabor filters DC response versus its bandwidth, Field [76] proposed the *log-Gabor* filter $G_l(\omega)$, which is defined as

$$G_l(\omega) = n_c \cdot \exp\left[-\frac{ln^2(\omega/\omega_0)}{2\ ln^2(\kappa_\beta)}\right], \tag{3.7}$$

where $\omega_0$ is the filter's center frequency, and parameter $\kappa_\beta$ (where $0 < \kappa_\beta < 1$) controls the shape of the filter, which is related to the bandwidth $\beta$ of the filter via




(a) $G(\omega)$ on the linear scale.

(b) $G(\omega)$ on the log scale.

(c) $G_l(\omega)$ on the linear scale.

(d) $G_l(\omega)$ on the log scale.

FIGURE 3.2: Frequency response of the Gabor filter and the log-Gabor filter, presented in the *linear* and *log* scales, respectively. At the DC, the Gabor filter has a non-zero response, while the log-Gabor has a zero-response. A stronger response at higher frequency range is also shown in the log-Gabor filter.

the following [76]

$$\beta = -\frac{2\sqrt{2}}{\sqrt{ln2}} \ln(\kappa_\beta) \quad \text{or} \quad \kappa_\beta = \exp\left(-\frac{\beta}{4}\sqrt{2\ ln2}\right). \tag{3.8}$$

In the logarithmic-scale, the DC component corresponds to the limit of the log function yielded at zero frequency. Since $\lim_{x\to 0^+} \ln(x) = -\infty$, $\lim_{\omega\to 0^+} G_l(\omega) = 0$ according to (3.7).

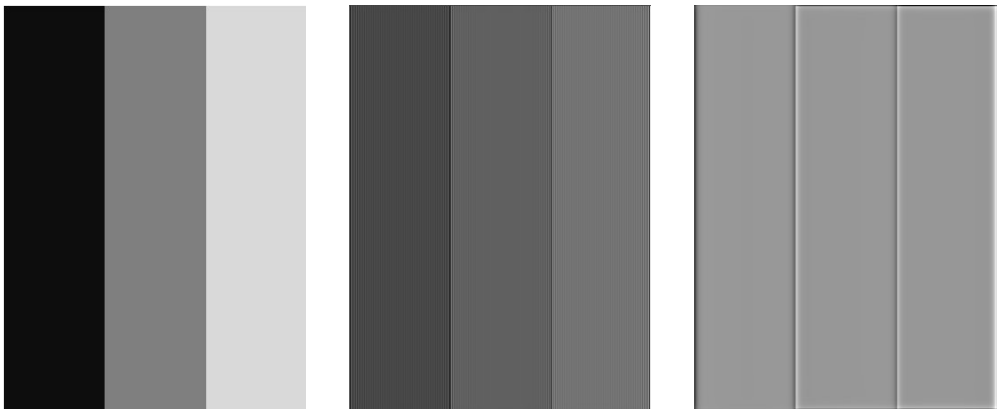
Referring to Fig. 3.2 for some comparisons of the Gabor and the log-Gabor filters in their frequency domains. Note that the *linear* scale is used in (a) and (c), while the *logarithmic* scale is used in (b) and (d). The color legends of these plots are: (Blue) center frequency at 0.05 and 1-octave bandwidth; (Green) center frequency at 0.2 and 1-octave bandwidth; (Red) center frequency at 0.2 and 3-octave bandwidth. For the Gabor filter, when its bandwidth is larger than 1 octave, its DC response

is non-negligible. On the other hand, the log-Gabor filter is able to have larger bandwidth while maintaining its zero-response at DC.

Compared with the Gabor filter, the log-Gabor filter produces two unique and highly desirable characteristics: (i) zero response at the DC, and (ii) wider bandwidth (i.e., yielding stronger response at higher frequencies). For (i), as mentioned previously, the DC component captures the global information, which is much less relevant to the evaluations of image quality. Refer to Fig. 3.3 for an illustration, different responses from the Gabor filter and the log-Gabor filter in the constant regions is seen. In the obtained feature maps, neutral grey means 0 response while darker color represents a lower response and a brighter color represents a higher response. For the log-Gabor case as illustrated in Fig. 3.3 (c), it can be seen that the response is the same in all the constant regions regardless of its original intensity. In this example, the three gray-scale intensity values of their corresponding strips shall be treated as irrelevant contributions from the viewpoint of IQA. This matches to the filtered result of the log-Gabor filter's as shown in Fig. 3.3 (c), where all three strips have 0 response. For (ii), image distortions tend to be reflected in the range of higher frequencies. These match the log-Gabor filter's characteristics on signal representation and therefore help to conduct image quality assessment more accurately.

For the log-Gabor filter, it has been shown in [81] that $\kappa_\beta = 0.74$ will result in a filter bandwidth of 1 octave, $\kappa_\beta = 0.55$ of 2 octaves, and $\kappa_\beta = 0.41$ of 3 octaves. Psychophysical research has shown that the bandwidth of the transfer functions of the visual cortex of mammal's cells is between 1 octave and 3 octaves [82]. Consider that the information acquired from the higher frequency range is more useful for the IQA perceived by the HVS, thus $\kappa_\beta = 0.41$ is exploited in this chapter.

The log-Gabor filter as shown in (3.7) is convolved with the *luminance* component of each SCI for extracting their features. Through an independent convolution



(a) Input image (b) Gabor convolved. (c) Log-Gabor convolved.

FIGURE 3.3: Illustrating the zero-DC response produced by the log-Gabor filter as shown in (c).

along each direction, the horizontal and vertical responses are generated as

$$\begin{aligned} H_\zeta(x,y) &= G_l^h(x,y) \otimes L_\zeta(x,y); \\ V_\zeta(x,y) &= G_l^v(x,y) \otimes L_\zeta(x,y), \end{aligned} \tag{3.9}$$

where the symbol $\otimes$ denotes the convolution operator. A summation of the obtained horizontal and vertical responses is defined as the *log-Gabor feature* of the luminance component in this chapter; that is,

$$G_{l\zeta}(x,y) = H_\zeta(x,y) + V_\zeta(x,y). \tag{3.10}$$

Fig. 3.4 and Fig. 3.5 give two examples of the extracted log-Gabor features of a reference SCI and its distorted version. In the example of the "Tower Bridge" SCI in Fig. 3.4 (a), the majority of people's attention falls on the texts, and some of the attention is on the photo of tower bridge. The background texture hardly contribute to the perceptual quality of such image. In the Gabor feature maps shown in Fig. 3.4 (b) and (e), a lot of unwanted response comes from the area other than text and that bridge, making the Gabor feature look messy. The response from such non-perceptual-quality-related areas is not helpful to IQA, and may even have negative effective on the accuracy of the IQA as additional error may be

brought in. In contrast, refer to the log-Gabor feature maps shown in Fig. 3.4 (c) and (f), they are highly consistent with our perceptual point view. All the edges of the texts and the bridge are accurately captured and have a high response to log-Gabor filter, while the response from other areas remains very low, as shown in the feature maps as very dark color. Fig. 3.5 gives another example of using the "Dashboard" SCI. Similar to the "Tower Bridge" example, the log-Gabor filter can capture the edges of all the elements in the image that people will watch on, while other areas are with low responses and are kept clean. Such property of log-Gabor filter lays the foundation for the success of our proposed LGFM IQA model.

### 3.2.4 Similarity Measurements

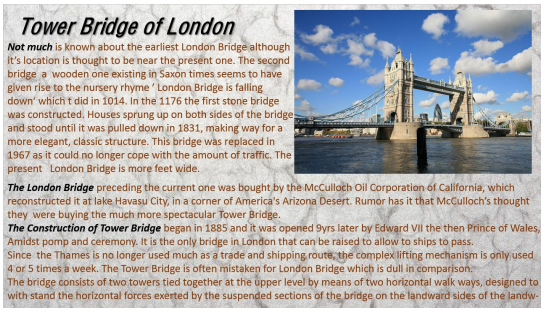
In compliance with many existing state-of-the-art works (e.g., [2, 74, 83]), the degree of similarity of the log-Gabor features, between the luminance components of the reference and distorted SCIs, $S_{G_l}$, can be calculated according to

$$S_{G_l}(x,y) = \frac{2G_{lr}(x,y)\cdot G_{ld}(x,y) + c_1}{G_{lr}^2(x,y) + G_{ld}^2(x,y) + c_1}, \tag{3.11}$$

where $G_{lr}(x,y)$ and $G_{ld}(x,y)$ are the log-Gabor feature maps of the SCIs $r$ and $d$, respectively, and $c_1$ is a constant inserted for avoiding numerical instability when the denominator becomes to zero.

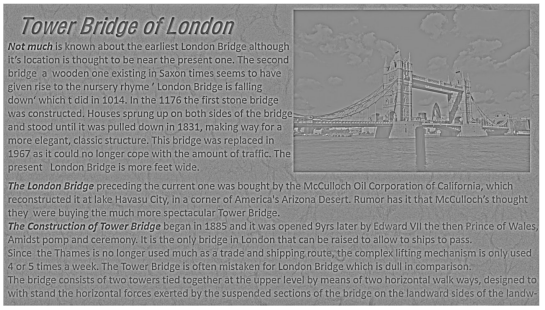
Besides the luminance distortion, the chrominance components (denoted as $M_\zeta(x,y)$ and $N_\zeta(x,y)$) also affect the perceptual assessment of the visual quality of color images to a large extent [15]. The chrominance similarity is used to reflect the quality degradation caused by color distortions, which is defined as

$$\begin{aligned} S_C(x,y) = &\frac{2M_r(x,y)\cdot M_d(x,y) + c_2}{M_r^2(x,y) + M_d^2(x,y) + c_2} \\ &\cdot \frac{2N_r(x,y)\cdot N_d(x,y) + c_2}{N_r^2(x,y) + N_d^2(x,y) + c_2}, \end{aligned} \tag{3.12}$$

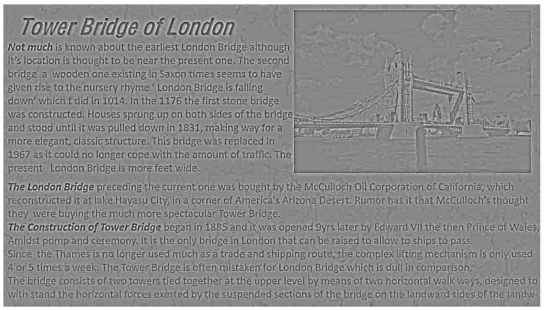



(a) Reference SCI "Tower Bridge"

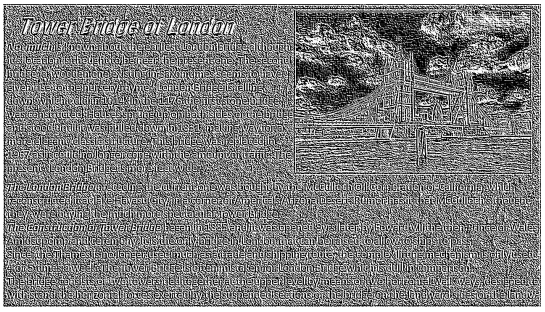

(e) Distorted SCI "Tower Bridge"



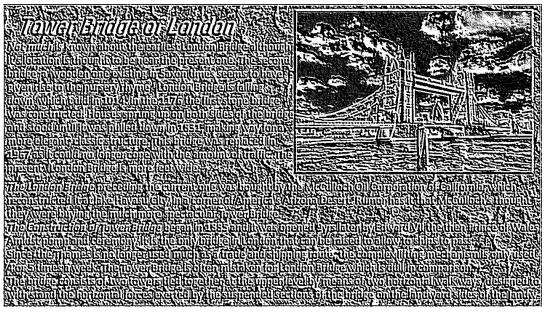

(b) Luminance component of (a)



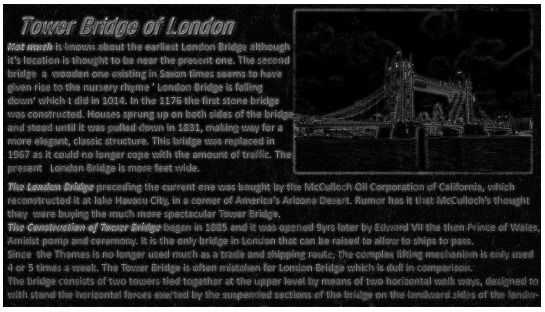

(f) Luminance component of (e)



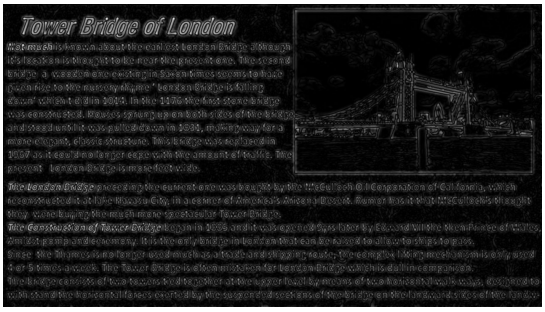

(c) Gabor feature map of (b)




(g) Gabor feature map of (f)




(d) Log-Gabor feature map of (b)




(h) Log-Gabor feature map of (f)

FIGURE 3.4: An illustration of the extracted log-Gabor features of SCI "Tower Bridge" and comparison to their Gabor features. Left column: reference image. Right column: distorted image.

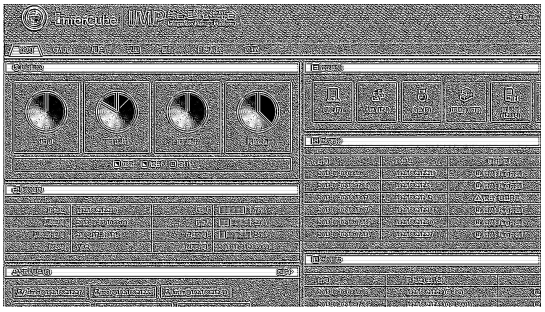
where $M_r(x, y)$, $N_r(x, y)$ and $M_d(x, y)$, $N_d(x, y)$ are the chrominance components of the reference SCI and the distortion SCI, respectively. Likewise, $c_2$ is a positive

(a) Reference SCI "Dashboard"

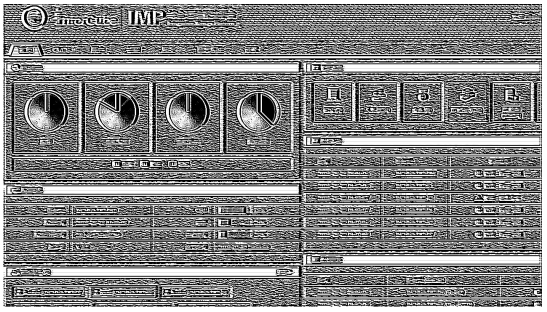
(e) Distorted SCI "Dashboard"

(b) Luminance component of (a)

(f) Luminance component of (e)

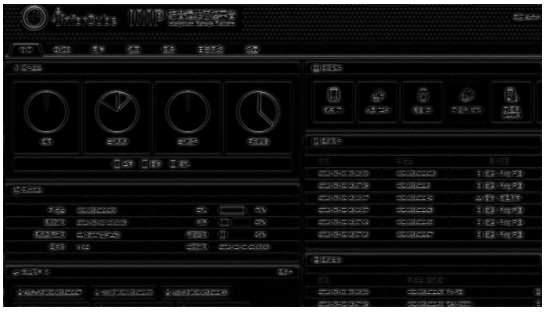


(c) Gabor feature map of (b)



(g) Gabor feature map of (f)

(d) Log-Gabor feature map of (b)



(h) Log-Gabor feature map of (f)

FIGURE 3.5: An illustration of the extracted log-Gabor features of SCI "Dashboard" and comparison to their Gabor features. Left column: reference image. Right column: distorted image.

constant, playing the same role as $c_1$.

After obtaining the similarity measurement on the luminance and two chrominance components, the local quality score $S_Q(x, y)$ is then computed as

$$S_Q(x, y) = [S_{G_l}(x, y)]^{\alpha_1} \cdot [S_C(x, y)]^{\alpha_2}, \tag{3.13}$$

where $\alpha_1$ and $\alpha_2$ are two positive constants that are used to adjust the relative importance of $S_{G_l}(x, y)$ and $S_C(x, y)$. In this chapter, the values of these two parameter are empirically determined as $\alpha_1 = 1$ and $\alpha_2 = 0.04$.

### 3.2.5 Log-Gabor Feature Pooling Strategy

Intuitively, our HVS pays more attention on those pixels with stronger features and therefore results in stronger visual perception on those locations. This suggests that pixels that yield a larger log-Gabor feature value should be given more consideration. The log-Gabor feature pooling strategy is thus designed to accommodate such intuition as follows.

At each pixel location $(x, y)$, the log-Gabor filter generated features from the reference SCI (i.e., $G_{lr}(x, y)$) and the distorted SCI (i.e., $G_{ld}(x, y)$) will be considered; that is, the larger value of these two quantities will be exploited as the weighting factor to reflect the above-mentioned phenomenon. Thus, a *weight map* $\omega(x, y)$ can be generated according to

$$\omega(x, y) = \max\{|G_{lr}(x, y)|, |G_{ld}(x, y)|\}, \tag{3.14}$$

where $|\cdot|$ means the absolute value of the argument will be performed. The final image quality assessment measurement, *LGFM Score*, can be obtained by performing the weighted average over all the pixel locations $(x, y)$ on the local quality map

$S_Q(x, y)$; that is,

$$LGFM\ Score = \frac{\sum_{(x,y)} \omega(x, y) \cdot S_Q(x, y)}{\sum_{(x,y)} \omega(x, y)}. \tag{3.15}$$

## 3.3 Experimental Results

### 3.3.1 Performance Comparison

In our experiments, the most recent and largest publicly available database for screen content image quality assessment, namely SCID [83], is used for conducting performance evaluation of SCIs using our proposed LGFM IQA model and several state-of-the-art models for comparisons.

Besides SCID, there are many IQA related datasets such as CSIQ [84], TID2008 [85], TID2013 [86], LIVE [87], and some more recent databases such as Waterloo Exploration Database [88] and KonIQ-10k [21]. But they are all natural scene images instead of screen content images. SIQAD [89] is a dataset that focuses on screen content images. But comparing to SCID, it contains only 20 source and 980 distorted SCIs, approximately half of SCID.

Subsequently, a database called quality assessment of compressed SCI (QACS) [90] is created. Compared to SIQAD, QACS added the HEVC and HEVC SCC compression artifacts, used higher-resolution SCIs, and included more variations on SCI content types and application scenarios. However, the QACS dataset contains 24 reference SCIs and 528 distorted SCIs of only 2 distortion types.

The SCID contains 40 reference SCIs and 1,800 distorted versions rendered from the reference SCIs. There are nine different types of image distortions in the distorted SCIs in SCID, including Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), JPEG compression (JPEG), JPEG2000 compression


FIGURE 3.6: GUI for collecting user subjective opinion scores for distorted SCIs.

(J2K), color saturation change (CSC), distortion by the screen content coding extension to High Efficiency Video Coding (HEVC-SCC), and color quantization with dithering (CQD). Each of the distorted SCI comes with a *mean opinion score* (MOS) describing the quality of the distorted SCI and is used as the "ground truth" for evaluating the performance of various IQA models.

To obtain such MOS, 186 human assessors' were involved in a double stimulus subjective test as specified in ITU-R BT.500-13 [91] and the double stimulus impairment scale (DSIS) method is employed. Among the assessors, there were 104 males and 82 females. All of them were students with ages ranging from 18 to 27, and they did not have any background or experience in image processing to avoid bias. For each pair of a reference SCI and one of its distorted versions, the reference SCI and the distorted SCI were presented to the assessors for 10 seconds each with a midgrey presentation inserted in between. The assessors were then instructed to give a subjective opinion score of their own to the distorted SCI. There were five opinion scores for the assessor to select and their corresponding definitions were given together as shown in Fig.3.6. The higher the score, the better the perceptual quality is deemed.

After obtaining all the assessors subjective opinion scores, outliers in the raw opinion scores were then identified and discarded according to the checking procedures described in [91]. Finally, the MOS value of each distorted SCI was computed as the mean of the re-scaled normalized opinion scores of every qualified assessors subjective opinion scores. The MOS value of each distorted SCI was then included in SCID, which served as the ground truth to evaluate the performance of the IQA models and the MOS can be considered as a real reflection of humans perception experiences.

Three standard performance evaluation procedures as suggested in [92] are used in this work. The *Pearson linear correlation coefficient* (PLCC) reflects the linear relationship between our model's predicted score and the ground truth *mean opinion score* (MOS); the closer the value of PLCC to 1, the stronger the linear correlation between the model's predicted score and MOS. The *Spearman rank order correlations coefficient* (SROCC) measures how well our model ranks the images in the same order as MOS; the closer the value of SROCC to 1, the better the ranking consistency between the model's predicted score and MOS. The *root mean square error* (RMSE) quantifies how far the model predicted scores are away from MOS on average. A smaller value of RMSE indicates a closer prediction, indicating higher accuracy of the model.

The proposed LGFM model is compared with the classical and state-of-the-art IQA models and documented in Table 3.1. In the experiment, the same data, SCID, are used to compare all IQA models, and the same parameters as the models were proposed in their original publications are used. Note that all the methods, except PSNR and SSIM [2], are designed for the evaluation of the SCIs. In this table, the first-ranked, the second-ranked, and the third-ranked performance evaluation results using each measurement criterion are boldfaced in red, blue, and black, respectively.

TABLE 3.1: PLCC, SROCC AND RMSE COMPARISON OF VARIOUS IQA MODELS UNDER DIFFERENT DISTORTION TYPES ON THE SCID DATABASE [1].

| | Distortions | PSNR | SSIM | SCQI | SIQM | SVQI | MDOGS | ESIM | GFM | LGFM |
|---|---|---|---|---|---|---|---|---|---|---|
| PLCC | GN | **0.9530** | 0.9354 | 0.9319 | 0.9269 | 0.9362 | 0.9418 | **0.9563** | 0.9497 | **0.9599** |
| | GB | 0.7772 | 0.8711 | 0.8244 | **0.9266** | 0.9130 | 0.9117 | 0.8700 | **0.9156** | **0.9309** |
| | MB | 0.7615 | 0.8794 | 0.8147 | **0.9152** | 0.8997 | 0.8565 | 0.8824 | **0.9023** | **0.9002** |
| | CC | 0.7435 | 0.6903 | **0.8353** | 0.7821 | 0.8266 | 0.8101 | 0.7908 | **0.8787** | **0.8336** |
| | JPEG | 0.8393 | 0.8581 | 0.9036 | 0.9226 | 0.9356 | **0.9375** | **0.9421** | **0.9392** | 0.9318 |
| | J2K | 0.9176 | 0.8586 | 0.9312 | 0.9076 | **0.9513** | 0.9063 | **0.9457** | 0.9226 | **0.9453** |
| | CSC | 0.0622 | 0.0890 | **0.8393** | 0.0683 | 0.0919 | 0.0699 | 0.0694 | **0.8728** | **0.8759** |
| | HEVC-SCC | 0.7991 | 0.7914 | 0.8708 | 0.8316 | 0.8496 | 0.8048 | **0.9108** | **0.8740** | **0.9036** |
| | CQD | **0.9210** | 0.7810 | 0.8823 | 0.8385 | **0.9047** | 0.7950 | 0.9005 | 0.8928 | **0.9203** |
| | Overall | 0.7622 | 0.7343 | 0.7489 | 0.8303 | 0.8604 | 0.8384 | **0.8630** | **0.8760** | **0.9023** |
| SROCC | GN | 0.9424 | 0.9171 | **0.9556** | 0.9133 | 0.9191 | 0.9256 | **0.9460** | 0.9370 | **0.9511** |
| | GB | 0.7702 | 0.8698 | 0.8638 | **0.9232** | 0.9079 | 0.9055 | 0.8699 | **0.9081** | **0.9287** |
| | MB | 0.7375 | 0.8588 | 0.8587 | **0.9006** | 0.8842 | 0.8247 | 0.8608 | **0.8892** | **0.8933** |
| | CC | 0.7265 | 0.6564 | 0.7465 | 0.7435 | **0.7705** | 0.7636 | 0.6182 | **0.8225** | **0.8575** |
| | JPEG | 0.8321 | 0.8490 | 0.9171 | 0.9158 | **0.9287** | 0.9331 | **0.9455** | 0.9281 | **0.9324** |
| | J2K | 0.9074 | 0.8439 | 0.9270 | 0.8935 | **0.9367** | 0.9097 | **0.9359** | 0.9085 | **0.9342** |
| | CSC | 0.0908 | 0.0963 | **0.8970** | 0.0617 | 0.0790 | 0.0607 | 0.1037 | **0.8736** | **0.8901** |
| | HEVC-SCC | 0.8074 | 0.8263 | **0.8721** | 0.8517 | 0.8665 | 0.8423 | **0.9036** | 0.8712 | **0.8914** |
| | CQD | **0.9080** | 0.7766 | **0.9099** | 0.8301 | 0.8957 | 0.8236 | 0.8868 | 0.8907 | **0.9152** |
| | Overall | 0.7512 | 0.7146 | 0.7814 | 0.8086 | 0.8386 | 0.8142 | **0.8478** | **0.8759** | **0.9046** |
| RMSE | GN | **3.8093** | 4.4458 | 4.5600 | 4.8222 | 4.4179 | 4.2257 | **3.6760** | 3.9378 | **3.5248** |
| | GB | 6.6633 | 5.1998 | 5.9943 | **4.0989** | 4.3194 | 4.3507 | 5.2213 | **4.2566** | **3.8679** |
| | MB | 7.0843 | 5.2044 | 6.3394 | **4.7388** | 4.7709 | 5.6415 | 5.1431 | **4.6121** | **4.7625** |
| | CC | 5.9867 | 6.4767 | 4.9217 | 6.1281 | **5.0374** | 5.2487 | 5.4790 | **4.2732** | **4.9446** |
| | JPEG | 8.1718 | 7.7179 | 6.4390 | 6.7341 | **5.3053** | 5.2319 | **5.0373** | **5.2011** | 5.4587 |
| | J2K | 6.3222 | 8.1562 | 5.8002 | 7.2951 | **4.9058** | 6.7242 | **5.1695** | 6.1385 | **5.1898** |
| | CSC | 9.8203 | 9.8003 | **5.3503** | 9.8394 | 9.7977 | 9.8153 | 9.8156 | **4.8031** | **4.7744** |
| | HEVC-SCC | 8.4009 | 8.5037 | 6.8407 | 8.1970 | 7.3381 | 8.2575 | **5.7446** | **6.7590** | **5.9596** |
| | CQD | **4.9814** | 7.9855 | 6.0188 | 7.1976 | **5.4481** | 7.7554 | 5.5607 | 5.7592 | **5.0151** |
| | Overall | 9.1682 | 9.6133 | 9.3846 | 7.8920 | 7.2178 | 7.7186 | **7.1552** | **6.8310** | **6.1052** |

Overall, the proposed LGFM model achieves the highest PLCC and SROCC values and the lowest RMSE value. This indicates that the proposed LGFM model is most consistent with subjective judgments agreed by the HVS, compared with other state-of-the-art IQA models.

TABLE 3.2: COMPUTATIONAL COMPLEXITY COMPARISON, MEASURED IN AVERAGE RUNNING TIME (IN SECONDS) PER IMAGE

| Model | Time (Seconds) | Model | Time (Seconds) |
|---|---|---|---|
| PSNR | 0.0352 | SQMS | 0.2047 |
| SSIM | 0.1298 | MDOGS | 0.0488 |
| SCQI | 0.3681 | ESIM | 3.0126 |
| SIQM | 0.2504 | GFM | 0.1668 |
| SVQI | 2.4323 | LGFM | 0.2936 |

### 3.3.2 Computational Complexity Comparison

Besides the accuracy of the IQA model, the computational complexity is another figure of merit for consideration. In our experiment, the average running time per image (in seconds) is compared. The proposed model is implemented in MATLAB R2017b, and is run on a computer with Intel i5-4590 CPU@3.30GHz and 8GBs of RAM. Table 3.2 shows a comparison of the run-time results among various IQA models. It can be observed that the proposed LGFM model achieved the best accuracy without sacrificing much computational complexity.

## 3.4 Summary

In this chapter, a novel image quality assessment model for the evaluation of screen content images (SCIs), called the *log-Gabor feature-based model* (LGFM), is developed. The novelty lies in the use of the log-Gabor filter to generate salient features from the SCIs under comparison, and these features are fairly consistent with the perception of HVS on the image quality judgement of these SCIs. Extensive simulation results have clearly demonstrated that the proposed LGFM model is more accurate than other state-of-the-art IQA models on the evaluation of the SCIs.

# Chapter 4

# A Local Feature Descriptor for Image Blurriness Estimation

## 4.1 Introduction

Various image enhancement techniques have been developed to improve the perceptual quality of images. One of the most popular techniques is unsharp masking. Unsharp masking extracts the details of an image, then amplifies the details with a scaling factor, and finally add the amplified details back to the input images. With these operations, the sharpness of the image is improved. In the conventional unsharp masking approach, a fixed scaling factor is employed across the entire image. However, such fixed scaling practice is not always welcomed because in many cases, because some blurred regions in the image are purposely created by the photographer or the creator, to build a "pop out" effect on the main subject. Such effect is commonly known as the "bokeh" effect. Many camera manufactures, especially mobile phone manufactures, nowadays are spending a lot of efforts to strengthen such blurriness effect in their camera products via post-processing.

We have observed that enhancing a highly-blurred image region often leads to unsatisfactory results. Usually, noise in those region in the enhanced image could become more distinct than the original textures in those regions, leading to a distractive feeling of the perception. Secondly, the blurred regions are created on purpose as mentioned above. In this case, it is completely undesirable to enhance the detailed textures or features in the blurred regions. On the country, what is expected is to further blur those regions. Based on these observations, the need to separate the blurred and sharp regions of an image has come into our view so that appropriate post-processing algorithms can be applied without affecting the desired blurred regions. During the past years, one noticeable work is the *blurriness-guided unsharp masking* (BUM) [93] which exploited this idea as well.

The purpose of blurriness estimation is to separate the blurred and non-blurred regions so that any aforementioned post-processing can be facilitated. To achieve such objective, it is obvious that local feature descriptors should be incorporated, rather than global features. Herein, in this chapter, a blurriness estimation method based on a novel *cyclic rotational invariant local binary patterns* is proposed.

The rest of this chapter is organized as follows. Section 4.2 describes a proposed cyclic rotational invariant local binary pattern and demonstrates its application in blurriness estimation. Section 4.3 provides performance evaluations and demonstrates the use cases of the blurriness estimation method based on the proposed cyclic rotational invariant local binary pattern. Section 4.4 concludes this chapter.

## 4.2 Proposed Cyclic Rotation Invariant Local Binary Pattern for Blurriness Estimation

### 4.2.1 Motivation by Rapid Transform

The Rapid transform (RT) [94] was initially proposed half century ago and is a fast transform which is useful for shift invariant signal classification and recognition. Given an input vector $\{X_0,\ X_1,\ ...,\ X_{N-1}\}$ with $N = 2^M$ elements, numbered from 0 to $N-1$, it takes $M$ iteration steps to complete the RT process. The iterative calculation procedures for an 8-element input vector with 3 iteration steps is demonstrated in Fig 4.1.

There are three important properties of Rapid transform that are worth mentioning: cyclic shift invariance, reflection invariance and feature dimension reduction.

**Cyclic Shift Invariance**

The RT transform is invariant under cyclic shift of the input sequence. With an input sequence of N numbers,

$$\begin{aligned} &RT\{X_{(i\ mod\ N)},\ X_{(i+1\ mod\ N)},\ ...,\ X_{(i+N-1\ mod\ N}\} \\ &= RT\{X_{(j\ mod\ N)},\ X_{(j+1\ mod\ N)},\ ...,\ X_{(j+N-1\ mod\ N}\} \end{aligned}$$

where $i$ and $j$ are any positive integers and mod performs modulo operation. This shift invariant property is mathematically proved in [94].

**Reflection Invariance**

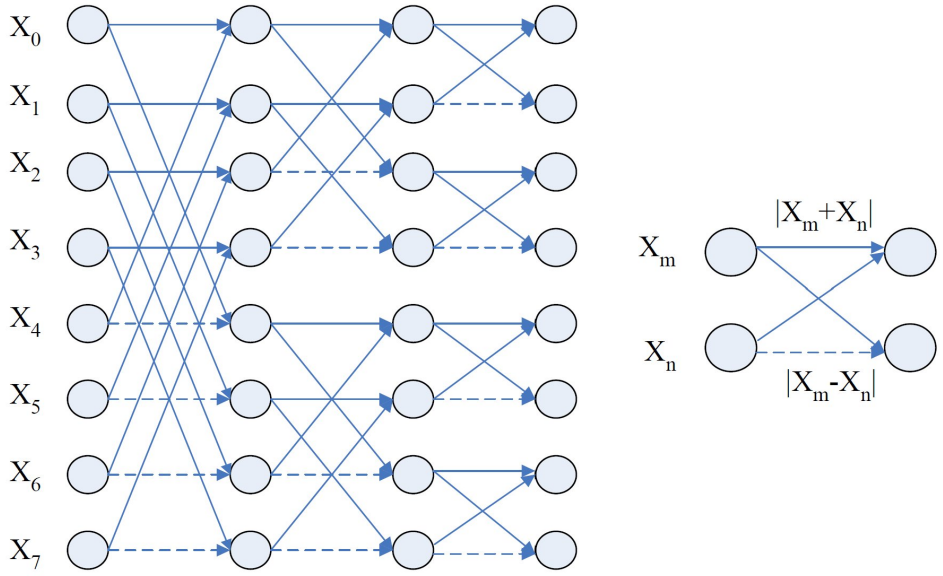



FIGURE 4.1: Rapid transform algorithm for 8-element input vector. The unit operation consist of an addition operation and an absolute value operation on a subtraction.

The RT transform is invariant under reflection of the input sequence. For an input sequence with $N$ elements,

$$\begin{aligned} &RT\{X_1,\ X_2,\ ...,\ X_{N-1}\} \\ &= RT\{X_{N-1},\ X_{N-2},\ ...,\ X_1\} \end{aligned}$$

### Feature Dimension Reduction

Given an input of 8-bit binary pattern, the number of possible 8-bit binary patterns is 256. After Rapid transform, because of the above-mentioned two properties, the possible output patterns after RT will be reduced to 21.

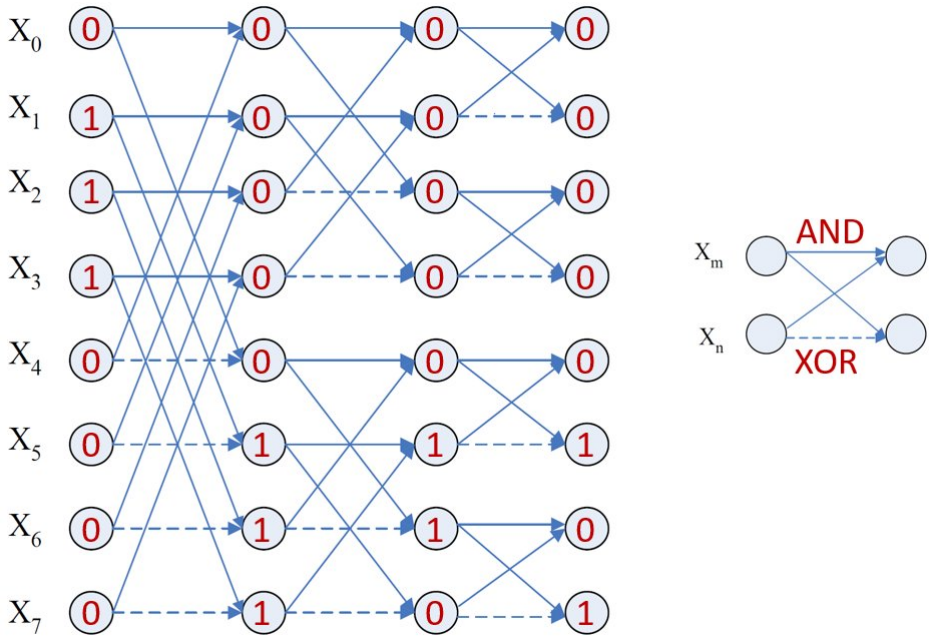



FIGURE 4.2: Proposed calculation method for binary domain Rapid transform.

### 4.2.2 Proposed Cyclic Rotation Invariant Local Binary Pattern

Before we introduce the proposed cyclic rotation invariant LBP, inspired by the appealing properties brought by Rapid transform, we firstly propose a binary domain Rapid transform since we will only take binary numbers as the input to the proposed RT, and the desired output is also binary.

In binary domain, the addition, subtraction and taking absolute value operations do not make sense any more. To achieve the similar effects to those operations in decimal calculation, we propose to use AND operation to substitute the ordinary add operation, and use XOR (exclusive OR) to substitute the ordinary subtraction and absolute value operations. Hence, all calculations are done with Boolean operators in the proposed binary domain Rapid transform. Given an input in the binary form, the output from the proposed binary domain Rapid transform remains in binary. Fig. 4.2 shows the calculation procedure of the proposed binary domain Rapid transform with an example of 01110000 as the input pattern.

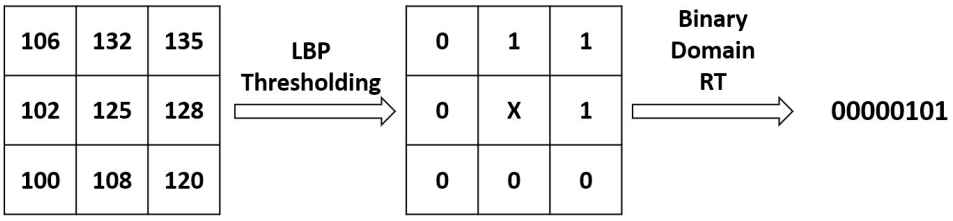


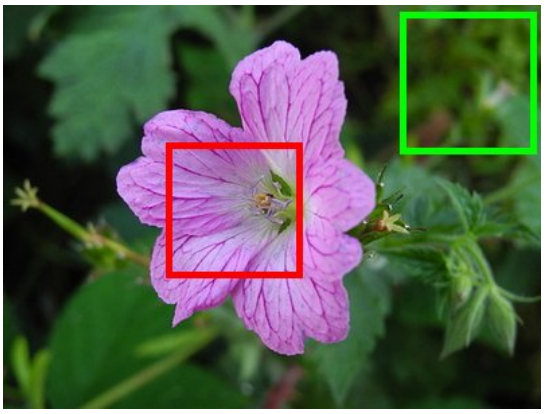

FIGURE 4.3: Proposed calculation method for cyclic rotation invariant local binary pattern.

With the proposed binary domain Rapid transform as a powerful tool to achieve cyclic rotation invariance, refection invariance and dimension reduction, now we can define our proposed cyclic rotation invariant local binary pattern. The first step in the proposed cyclic rotation invariant LBP is the same as conventional LBP descriptors that thresholding the neighbouring pixels by the center pixel's value in a 3×3 patch. After the ordinary LBP pattern is obtained, it will go through the proposed binary domain Rapid transform and hence will we get the cyclic rotation invariant LBP, named as *RT-LBP*. This process is demonstrated with an example in Fig. 4.3.

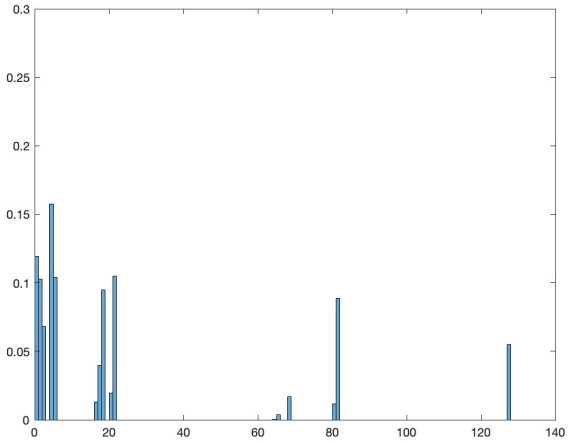
### 4.2.3 Blurriness Estimation Based on Proposed Cyclic Rotation Invariant Local Binary Pattern

From experiments and observation, the statistical properties of the RT-LBP pattern distribution on the sharp region of images and the blurred region of images are strongly distinguishable. This is consistent with the observations in [37].

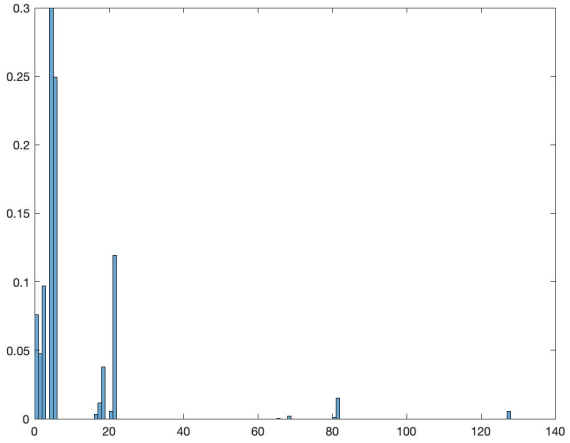
It is observed that the frequency of patterns with RT-LBP values smaller than 60 in the sharp regions is noticeably less than that of blurred regions. Fig. 4.4 shows the histogram of the RT-LBP patterns appearing in the blurred and non-blurred regions of an image. Based on this, the proposed blurriness metric exploits these



(a) Selected sharp region (red squared region) and blurred region (green squared region) in an image.




(b) Histogram of RT-LBP patterns in sharp region.




(c) Histogram of RT-LBP patterns in blurred region.

FIGURE 4.4: RT-LBP code distribution in blurred and sharp regions.

observations and is expressed as:

$$m_{RT-LBP} = \frac{1}{N} \sum_{i<60} n(RT - LBP_i) \tag{4.1}$$

where $n(RT-LBP_i)$ is the number of RT-LBP pattern with its decimal value equal to $i$, and $N$ is the total number of pixels in the selected local region which serves to normalize the metric to the range $[0, 1]$. The final blurriness map generated by the proposed RT-LBP based method is illustrated in Fig. 4.7 (f). With this approach, the proposed RT-LBP based blurriness estimation method can successfully solve the blurriness estimation on the "almost-smooth" regions of images as well, which particularly satisfies our interest.

## 4.3 Experimental Results

Extensive simulations have been conducted on test images to evaluate the performance of the RT-LBP based blurriness estimation. Some representative results are shown in this section.

Since SCIs are a mixture of *natural* image content and *non-natural* image content as described in Chapter 1, therefore blurriness can be easily brought in by the natural image parts due to motion blur and out-of-focus blur. Blurriness on SCIs could also be inevitably introduced during the acquisition, processing, compression, transmission, as well as display stages. The overall blurriness score is calculated as the sum of the blurriness scores of every local region. A lower blurriness score indicates a more blurred image. The proposed blurriness estimation results on two SCIs and their distorted versions are shown in Fig. 4.5 and Fig. 4.6 respectively.

Secondly, the quality of the blurriness map was compared visually with one of the excellent previous work [36].

As shown in Fig. 4.7 (a), this is a kind of image that has a single object with a distinct level of blurriness from its background. For this given example, previous method by Ye *et al.'s* [93] under estimated the blurriness of the background leaves and treated some of the background regions as sharp region, as shown in Fig. 4.7 (c). In contrast, as shown in 4.7 (f), the proposed method gives an accurate estimation and separated the sharp "flower" region very well.

Another example is shown in Fig. 4.8 and Fig. 4.9. In these two images, the "lemon" and the yellow "road sign" region are the above-mentioned "almost-smooth" region. Compared to Fig. 4.8 (c) which fails to detect the lemon as sharp region, and Fig. 4.9 (c) which regards the road sign region as blurred region, our proposed method successfully identifies this region as sharp and in-focus.

Fig. 4.10 gives an example of the case that two main objects are in the image, but with different level of blurriness. Refer to Fig. 4.10 (a), the "dog" in front is the sharpest object in this image while the "cat" behind is a little bit blurred. The most blurred region is the background curtain. In this case, the proposed method detects the "dog" as sharp, and the boundary of the "cat" is sharp as well. Ye *et al.'s* [93] method detects both of the main objects as sharp and in focus. This may be more welcomed by users as both of them are the main objects in this image. However, in the application that we use the blurriness maps as a guidance to enhance the image, the visual difference between the two final enhanced results is barely distinguishable. This will be further illustrated below.

Thirdly, we fed the blurriness maps obtained by our proposed RT-LBP based blurriness estimation method to the blurriness-guided unsharp masking (BUM) [93], to test how helpful it is in the image enhancement applications.

As expected, our proposed blurriness map prevented the over enhancement problem in the result shown in Fig. 4.7 (d) and (e). Also it prevented the under enhancement problem in the result shown in Fig. 4.8 (d) and (e), and Fig. 4.9 (d) and (e). Both

of the credit should be given to the accuracy of the proposed blurriness estimation method.

In Fig. 4.10, the test image has two main objects but with different level of blurriness. Refer to Fig. 4.10 (d), it can be observed that our program does not handle the "slightly blurred" object as good as the result in Fig. 4.10 (b). But the difference in the final enhanced images using the different blurriness map is hardly visually distinguishable, as compared in Fig. 4.10 (c) and (e).

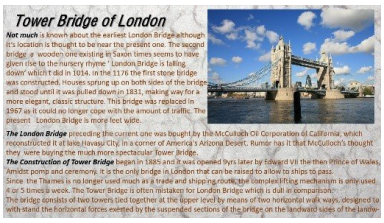
Fourthly, the proposed blurriness estimation technique was tested to evaluate the blurriness score of some recent image deburring algorithms and to see which can do a better job on reducing the blurring of the image. The chosen algorithms are from Nah *et al.* [95], SDNet4 [96], Tao *et al.* [97] and Liang *et al.* [98]. The results are shown in Fig. 4.11 below. In general, the trend and monotonicity of the estimated blurriness scores are consistent with our subjective perception. This experiment further demonstrates the usefulness of our proposed blurriness estimation method in benchmarking the performance of image deblurring algorithms.

Last but not least, the computational speed of the proposed RT-LBP based blurriness estimation method is way much faster than the previous state-of-the-art method as prosed in [93]. Take the image in Fig. 4.10 (a) as example. The resolution of the image is 800×530 pixels. Using a computer with Intel Core i7-7660 CPU and Matlab 2020a, our proposed method only takes *0.73* second to perform blurriness estimation, while Ye *et al.'s* requires *11.02* second.

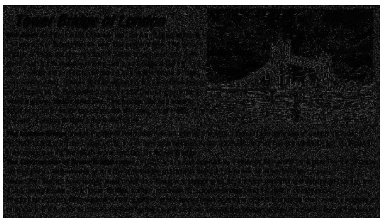
## 4.4 Summary

In this chapter, firstly a binary domain Rapid transform is proposed that brings the cyclic rotation invariance, reflection invariance and feature dimension reduction to conventional LBP descriptors to form the proposed RT-LBP. The proposed

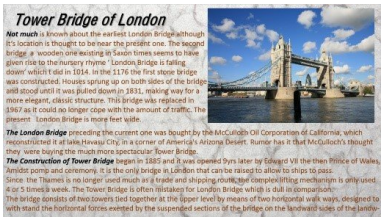
RT-LBP is then used in an image blurriness estimation method to generate an accurate pixel-wise blurriness map of each input image, followed by utilizing the obtained blurriness map to impose a pixel-wise adaptive enhancement, such that those highly-blurred regions are further blurred or not enhanced, and those sharp regions are highly enhanced. Extensive simulations on test images have demonstrated that the proposed method is able to deliver superior quality of the blurriness map and can serve a pre-processing algorithm for image enhancement, image segmentation, etc.



(a) Reference SCI "Tower Bridge" No blur.

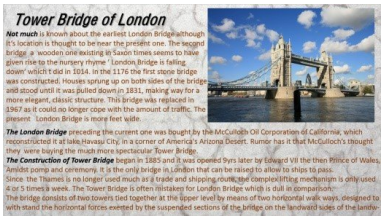


(b) Blurriness map of (a). Blurriness score = 18.97

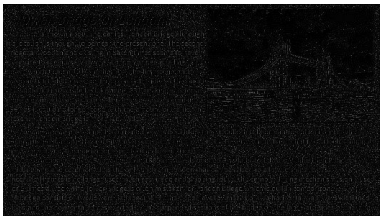


(c) Distorted SCI "Tower Bridge" Blurriness level 1.

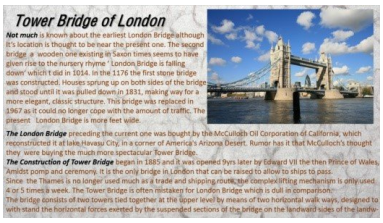
(d) Blurriness map of (c). Blurriness score = 16.43



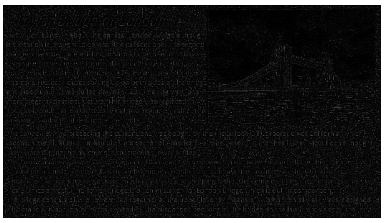
(e) Distorted SCI "Tower Bridge Blurriness level 2.



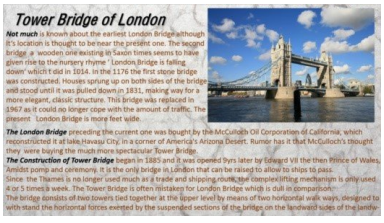
(f) Blurriness map of (e). Blurriness score = 13.97



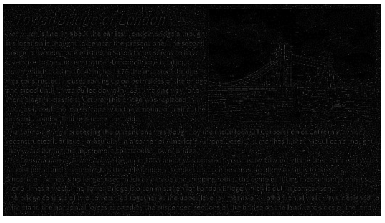
(g) Distorted SCI "Tower Bridge Blurriness level 3.



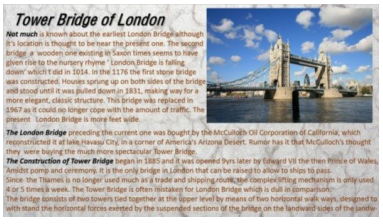
(h) Blurriness map of (g). Blurriness score = 11.98



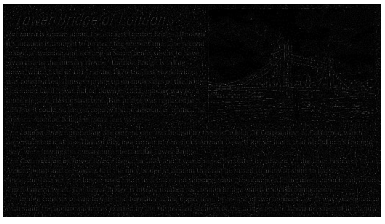
(i) Distorted SCI "Tower Bridge Blurriness level 4.



(j) Blurriness map of (i). Blurriness score = 10.47



(k) Distorted SCI "Tower Bridge Blurriness level 5.



(l) Blurriness map of (k). Blurriness score = 9.04

FIGURE 4.5: Proposed blurriness estimation result on the reference SCI "Tower Bridge" and its blurred versions with five different levels of Gaussian blur.

(a) Reference SCI "Factory" No blur.

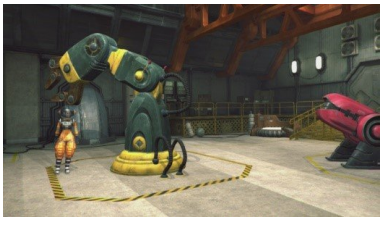
(b) Blurriness map of (a). Blurriness score = 16.87

(c) Distorted SCI "Factory" Blurriness level 1.

(d) Blurriness map of (c). Blurriness score = 11.87

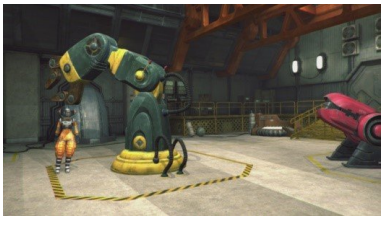
(e) Distorted SCI "Factory" Blurriness level 2.

(f) Blurriness map of (e). Blurriness score = 9.18



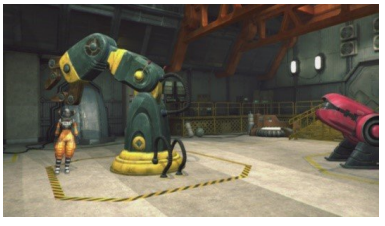
(g) Distorted SCI "Factory" Blurriness level 3.

(h) Blurriness map of (g). Blurriness score = 7.54



(i) Distorted SCI "Factory" Blurriness level 4.

(j) Blurriness map of (i). Blurriness score = 6.48



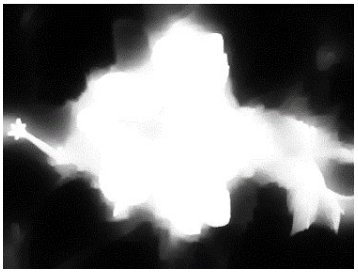
(k) Distorted SCI "Factory" Blurriness level 5.

(l) Blurriness map of (k). Blurriness score = 5.77

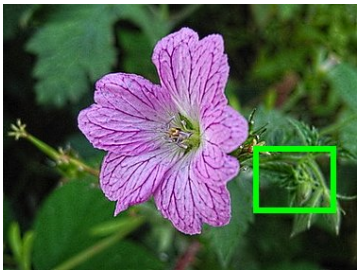
FIGURE 4.6: Proposed blurriness estimation result on the reference SCI "Factory" and its blurred versions with five different levels of Gaussian blur.

(a) Input image "flower"

(b) Zoom-in of (a)

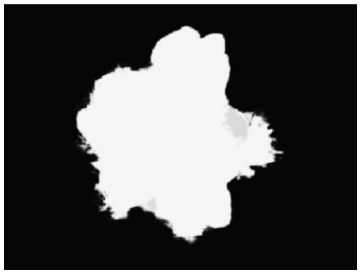


(c) Blurriness map by Ye [29]

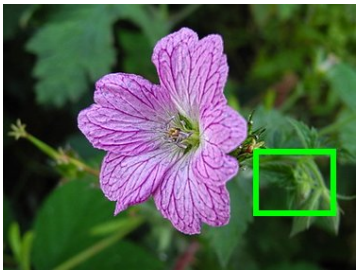


(d) BUM enhanced result

(e) Zoom-in of (d)



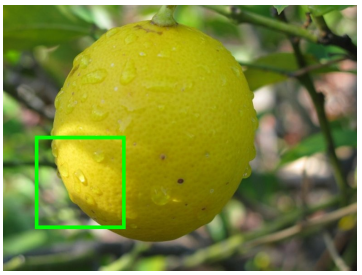
(f) Ours blurriness map



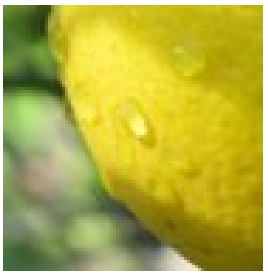
(g) Ours enhanced result

(h) Zoom-in of (g)

FIGURE 4.7: Comparison of the blurriness estimation and BUM result with the "flower" image.

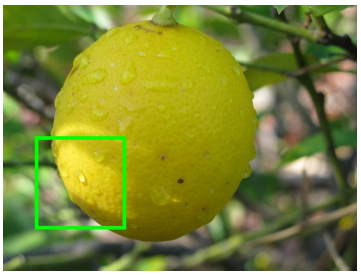


(a) Input image "lemon"

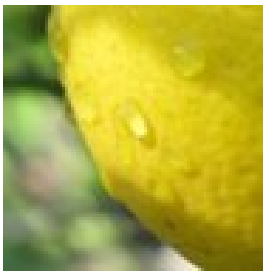


(b) Zoom-in of (a)

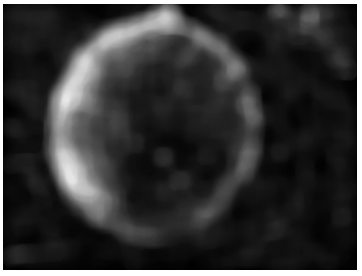
(c) Blurriness map by Ye [29]



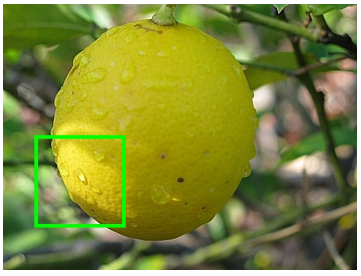
(d) BUM enhanced result



(e) Zoom-in of (d)



(f) Ours blurriness map



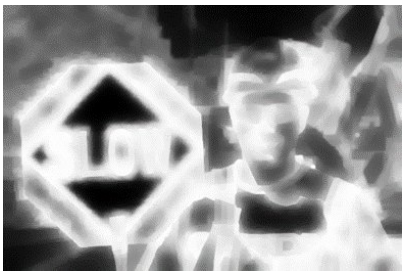
(g) Ours enhanced result

(h) Zoom-in of (g)

FIGURE 4.8: Comparison of the blurriness estimation and BUM result with the "lemon" image.


(a) Input image "road sign"

(b) Zoom-in of (a)



(c) Blurriness map by Ye [29]


(d) BUM enhanced result

(e) Zoom-in of (d)

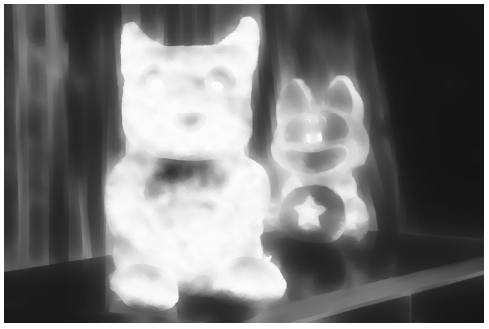
(f) Ours blurriness map


(g) Ours enhanced result

(h) Zoom-in of (g)

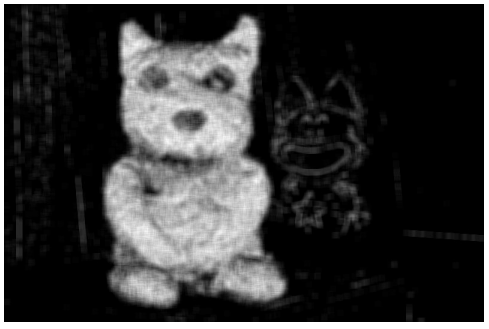
FIGURE 4.9: Comparison of the blurriness estimation and BUM result with the "road sign" image.

(a) An image with 2 objects



(b) Blurriness map by Ye [29]

(c) BUM enhanced result



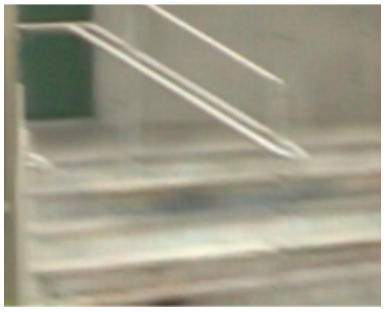
(d) Ours blurriness map

(e) Ours enhanced result

FIGURE 4.10: Comparison of the blurriness estimation with 2 main objects of different level of blurriness.

(a) Input image "Stair". Score = 3.35

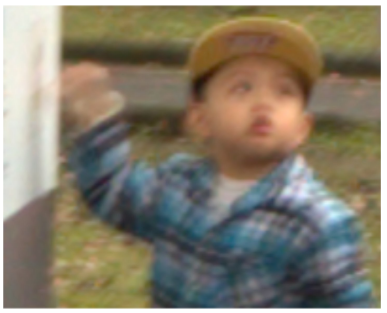
(b) Nah *et al.* [95]. Score = 3.18



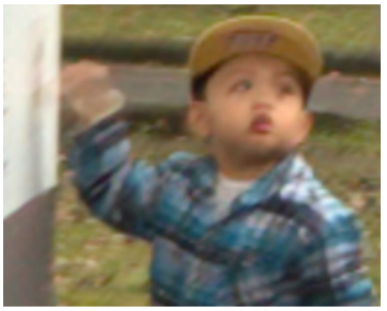
(c) SDNet4 [96]. Score = 2.89

(d) Tao *et al.* [97]. Score = 3.80

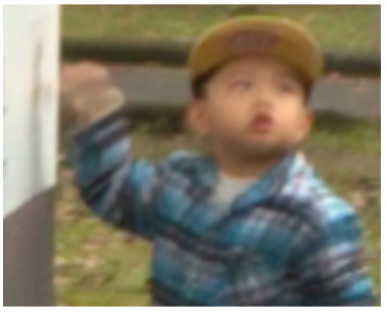
(e) Liang *et al.* [98]. Score = 3.61



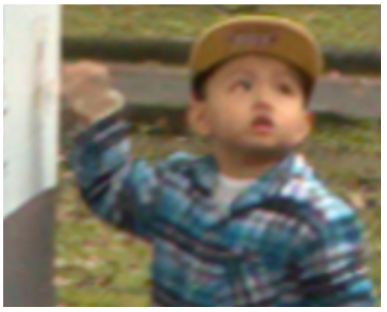
(f) Input image "Child". Score = 3.21



(g) Nah *et al.* [95]. Score = 4.19



(h) SDNet4 [96]. Score = 3.85



(i) Tao *et al.* [97]. Score = 4.26

(j) Liang *et al.* [98]. Score = 4.31

FIGURE 4.11: An evaluation of some recent image deblurring algorithms with proposed blurriness estimation method.

# Chapter 5

# A Color Space Study for Lightness Perception

## 5.1 Introduction

In the past decades, a lot of attempts have been made to create color space models that can define color relationships the same way as the human eye perceives, leading to the creation of HSL color space, HSV color space, CIELAB color space, and so on. Lightness has become one of the most important definitions of color-making attributes since then [60]. In this chapter, a novel estimation method for lightness is proposed, which takes into account the understanding of human reaction to color. The proposed method can both benefit applications such as image quality assessment, and being adopted in image enhancement applications. The novelty of the proposed lightness estimation method is that the proposed strategy takes human's physiological perceptual phenomenon into consideration, and utilizes the most recent survey result from real human's physiological response as useful reference, creating a promising physiologically-relevant lightness estimation methods with superior performance over the state-of-the-art models.

The remaining content of this chapter is organized as follows. In Section 5.2, the proposed human's physiological-perception-based algorithm for lightness estimation is described in detail. In Section 5.3, performance evaluation on the proposed lightness estimation with several comparable methods are presented, followed by two applications/use-cases demonstration of the proposed lightness estimation method. Section 5.4 summarizes this chapter.

## 5.2 Proposed Color Model for Screen Content Images

### 5.2.1 Motivation

In the era multimedia information processing and communications, digital image processing and display technologies have gone through many stages of technical advancements, both software (algorithms) and hardware. As a result, various kinds of display devices will deliver different perceptual experiences for the same image due to different physical properties and limitations brought by different display models and manufactures. Furthermore, different profile settings for the same type of display device can also produce different presentation result on the display as perceived by our human visual system. The above-mentioned variations could become even much more severe for displaying screen content images on a device, since such category of images tends to have high-dynamic image contents—i.e., more sharp edges, wider range of color gamut covered, highly sophisticated graphic shading requirements, to name a few.

It has been observed that the *lightness* of color images, which is the most important attribute of color images, tends to change drastically, depending on the color space used to represent the image under display. According to our observations, these

changes lie in two aspects: image local *details* and entire image *contrast*. To demonstrate, the lightness (perceived as gray-scale image) delivered by each of the above-discussed color spaces are shown in Fig. 5.1 for comparison. For example, Fig. 5.1 (e), the value component $V$ in HSV color space delivers the highest contrast (with brightest lightness) for the entire image compared with that of other color spaces. This is because the value $V$ is obtained by taking the maximum value of R, G, and B component for each pixel; thus, the value tends to be high. As a result, some structural details are lost, such as the fire's texture disappeared from the fire region, for example.

Two more examples are shown in Fig. 5.2 and Fig. 5.3; many lightness components are poorly presented. In another example as shown in Fig. 5.3 (c) and (f), although the edges between the color bar and its gray background are not completely lost in those two lightness components, the edge contrast is not correctly preserved and they are not consistent with human observations. Note that edges are not clear enough to detect, and the color bar cannot be as easily distinguished from the background as the one in the color image. Fig 5.12 (b) and Fig. 5.13 (b) gives more example on the cases with screen content images. The red text "GAMES" in Fig. 5.12 (a) and the object teapot in purple color on the far left in Fig. 5.13 (a) are quite vivid in the color image, but not much so in the L component of the LMN color space, as its edge contrast is suppressed, making the text and object appear less distinct, as shown in Fig. 5.12 (b) and Fig. 5.13 (b) respectively.

This reveals a fundamental issue that one must be very careful when choosing the type of color space to work with and such choice is highly application oriented. For example, whether it is for computer vision purpose as an image *recognition* issue, or human visual perception as an image *fidelity* issue.

Now, the issue is boiled down to how can we calculate the lightness of images with the most consistency with human eyes perception. For that, we shall highlight

what the technical issues we observed first before proposing our adaptive lightness enhancement (or compensation) technique.

### 5.2.2 Observations

It is important to note that most of the color spaces are calculated based on the *photopic luminous efficiency function* $V(\lambda)$, a CIE standard proposed in the 1924. This function describes human eyes relative sensitivity to light of different wavelengths and is used in the CIE standard to map the incident color light onto the values in *lightness*. However, this function is *not* always consistent with human eyes response and hence introduces error to the lightness calculation [99]. To address this issue, two fundamental discoveries could be utilized to compensate the inaccurate measurement of the lightness computed by the function $V(\lambda)$. They are: (1) Helmholtz-Kohlrausch Effect model [100] and (2) inability to discern image details presented in *blue* color [101]. They are described as follows.

#### 5.2.2.1 Helmholtz-Kohlrausch Effect Model

In the heterochromatic brightness matches test, observers were provided two colors stimuli, for example red and green. By adjusting the relative luminance of the red verses that of the green until the two colors appeared equally bright to the observer's eye. It has been found out that the stimuli luminance ratio is different for each such color pairs for making them appear as the same lightness to our human visual system. In this experiment, the luminous efficiency function $V(\lambda)$ of the CIE has been proved to be invalid for handling such heterochromatic brightness matches [102]. This issue was also noted by Helmholtz and experimentally justified by Kohlrausch, concluding that: *the perceived lightness of a stimulus increases as its chroma increases* [103]. In other words, by keeping the same amount of luminance,

(a) Color photograph (sRGB color space)

(b) Intensity $I$ in sRGB

(c) Luminance $Y$ in YUV

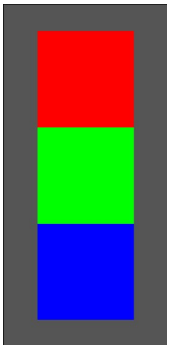
(d) Lightness $L$ in HSL

(e) Value $V$ in HSV

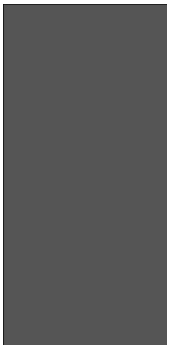
(f) $L^*$ in CLELAB

(g) $L$ in the Gaussian color model

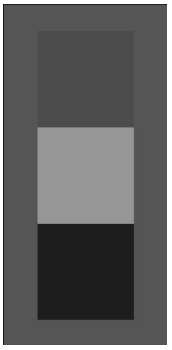
FIGURE 5.1: Comparison of the lightness component in different color spaces. The color image (a) is taken from https://upload.wikimedia.org/wikipedia/commons/0/02/_breathing_2_Luc_Viatour.jpg



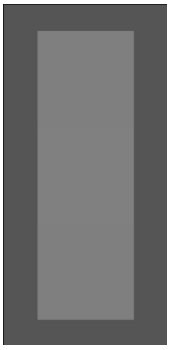
(a) Color photograph (sRGB color space)



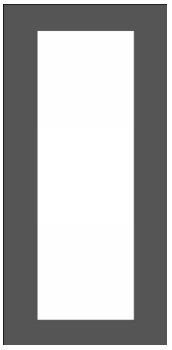
(b) Intensity $I$ in sRGB



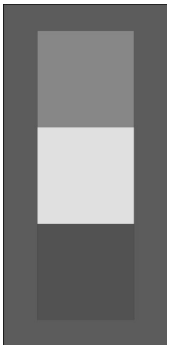
(c) Luminance $Y$ in YUV



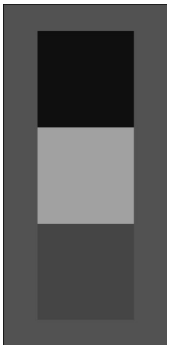
(d) Lightness $L$ in HSL



(e) Value $V$ in HSV

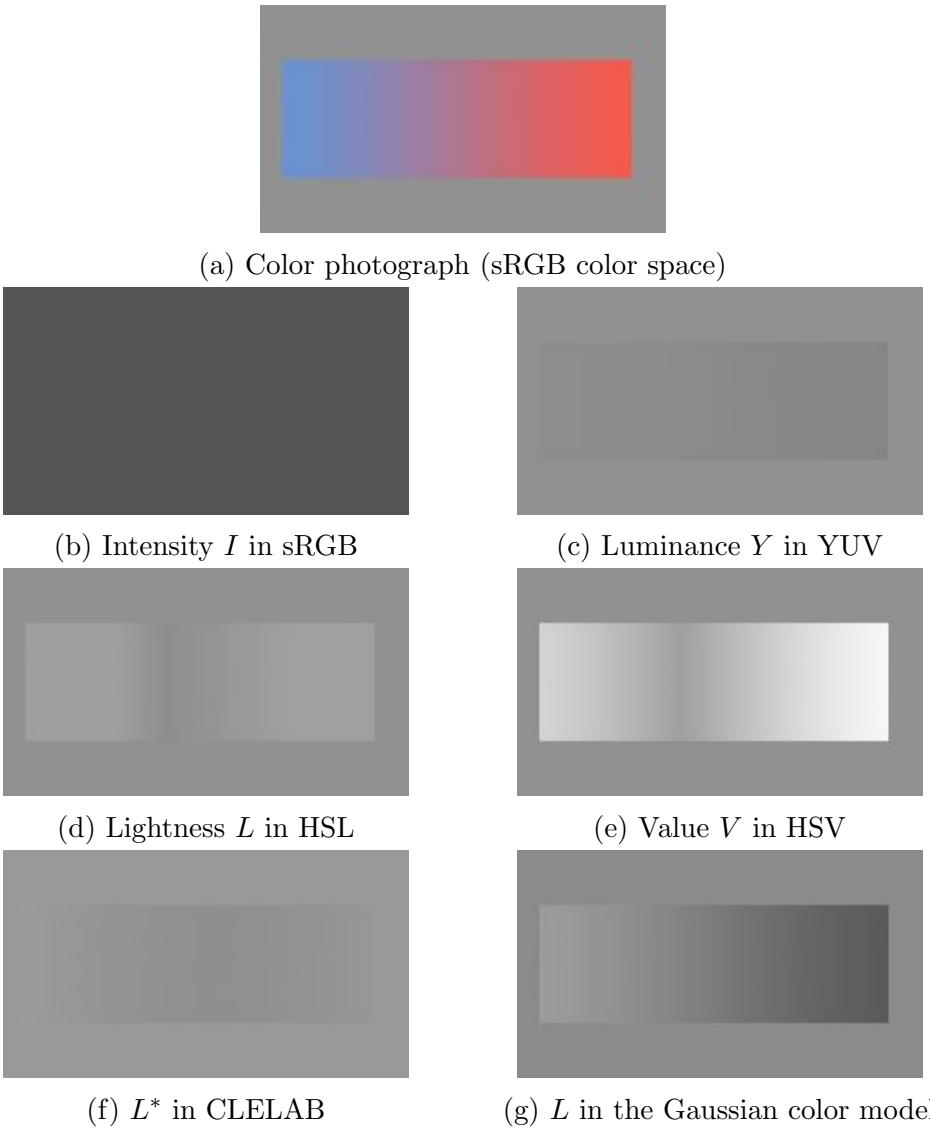


(f) $L^*$ in CLELAB



(g) $L$ in the Gaussian color model

FIGURE 5.2: Comparison of the lightness component of image "RGB" in different color spaces.



(a) Color photograph (sRGB color space)

(b) Intensity $I$ in sRGB

(c) Luminance $Y$ in YUV

(d) Lightness $L$ in HSL

(e) Value $V$ in HSV

(f) $L^*$ in CLELAB

(g) $L$ in the Gaussian color model

FIGURE 5.3: Comparison of the lightness component of image "color ramp" in different color spaces.

a highly chromatic or saturated color appears brighter than that of a less chromatic color.

Independently, Sharpe *et al.* [102] also pointed out that the photopic luminous efficiency function, $V(\lambda)$, used in CIELAB is not always consistent enough in accordance with human eyes perception. In [102], extensive experiments involving human observers for arriving at a more consistent relationship between color and lightness has been reported. In Fig. 5.4, the horizontal line represents the ideal

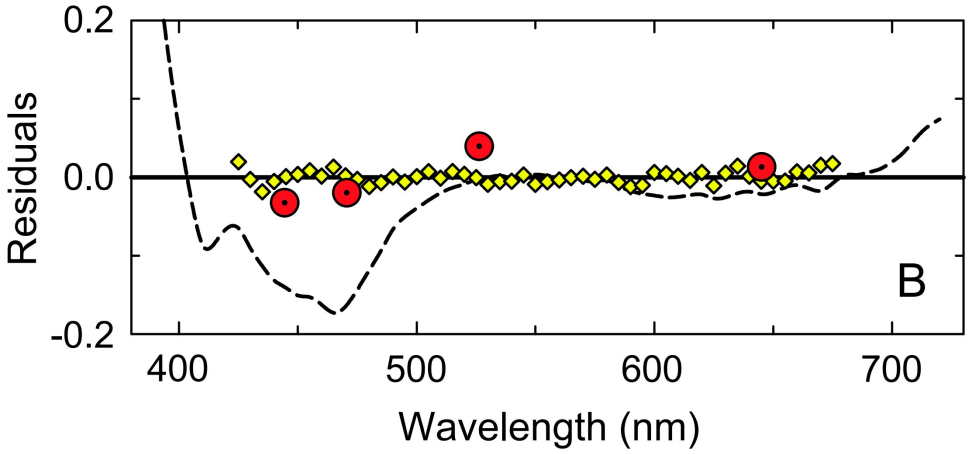



FIGURE 5.4: The dash line represents the difference yielded between the lightness measured by Sharpe *et al.* [102] and the lightness calculated according to CIELAB $L^*$ .

case that a zero difference between the human eyes actual sensitivity and the photopic luminous efficiency function $V(\lambda)$ is observed. However, it is just the ideal case. The difference between the CIELAB standard $V(\lambda)$ and the experimental result from Sharpe *et al.s* [102] observation is plotted as the dotted line. It can be found that the CIELAB has inaccurately *over* measured the lightness for the wavelength regions of violet (380 $nm$ to 420 $nm$ wavelength), and dark red color (700 $nm$ to 750 $nm$), but *under* measured for the region of 420 $nm$ to 500 $nm$ (covering blue to cyan colors) and of 580 $nm$ to 700 $nm$ (covering yellow, orange, and light red colors).

To overcome the above-mentioned inaccurate measurements, Bartleson *et al.* [104] proposed a new lightness measurement scheme to measure the perceived brightness from luminance at an early age. He proposed a formula describing how the human visual systems perceive lightness and the lights luminance. However, no further explanation was given to explain the form of the equation except that the equation fitted the data well. Later on, Fairchild *et al.* [100] proposed their model to predict the perceived lightness by taking the CIELAB *chroma* color space into consideration as follows.

Considering the fact that hue and chroma can affect the lightness perceived by human visual system, the *Helmholtz-Kohlrausch effect* (H-K effect, in short) was proposed in [100], describing that: *the perceived lightness will increase when the chroma increases, even though the physical lightness remains the same.* Based on the feedbacks collected from the human observers involved in their experiments, Fairchild *et al.* [100] proposed a simple but effective model to predict the human-perceived lightness $L^{**}$ as

$$L^{**} = L^{*} + 0.143 \times C^{*}, \tag{5.1}$$

where $L^{*}$ is is the lightness component expressed in the CIELAB color space and $C^{*}$ is the chroma calculated according to Equation (2.23). However, this model does not always solve the consistency problem between the human perceived lightness and CIELAB's lightness $L^{*}$. It can be seen that all the $L^{*}$ values receive some extent of positive compensation, and its corresponding $L^{**}$ values will become larger, since $C^{*}$ is positive. As a result, when an image is viewed on its $L^{**}$ component, the entire image looks brighter than $L^{*}$ and this may be considered as an additional distortion. Another drawback of this model is that in the cases that $L^{*}$ overestimates the human perceived lightness, for example in the wavelength regions of violet (380 $nm$ to 420 $nm$ wavelength), and dark red color (700 $nm$ to 750 $nm$), with such non-negative compensation proposed by Fairchild, the over-estimation error will become even larger.

Although the Fairchild's Helmholtz-Kohlrausch effect model does not solve the lightness estimation problem completely, it has indeed provided us with a direction that color components can be exploited for estimating the lightness of images more properly.

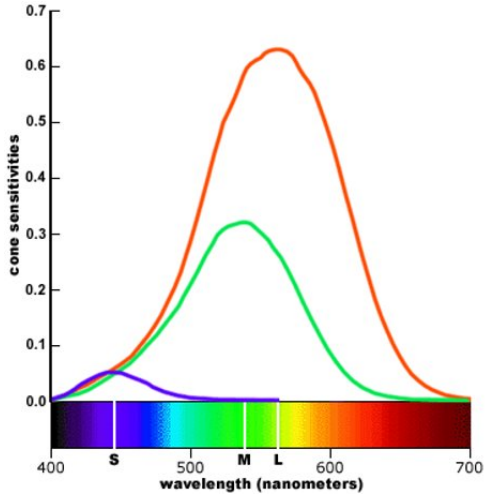



FIGURE 5.5: An comparison of the L, M and S cone sensitivities [105].

#### 5.2.2.2 Inability to Detect Image Details in Blue Color

As mentioned earlier, there are three types of cone photo-receptors in our human eyes, and each has a different peak spectral sensitivities to the long-, middle- and short-wavelengths. Thus, these cones are denoted by L, M, and S cones, respectively. This forms the foundations of our trichromatic color vision. According to the retinal anatomy, there are 63% L cones, 31% M cones, and 6% S cones in human's retina [105] and the population weighted linear cone sensitivity functions are plotted in Fig. 5.5. As illustrated in the figure, human eyes are the least sensitive to short wave light since S cone has the smallest number. This makes it harder for human eyes to observe blue features compared to others. Fig. 5.6 is an animation that can be viewed online. In this animated picture, the blue and yellow color squares are moving up and down at the same speed and pace. But human observers will see the blue squares making sudden starts and stops along the way and not moving smoothly. The reason is the low density of S cones in the retina makes human eyes more trouble distinguishing fine detail in blue against black or dark background.

To further address the aforementioned problem and especially make it more suitable in image quality assessment applications, we propose to incorporate the hue

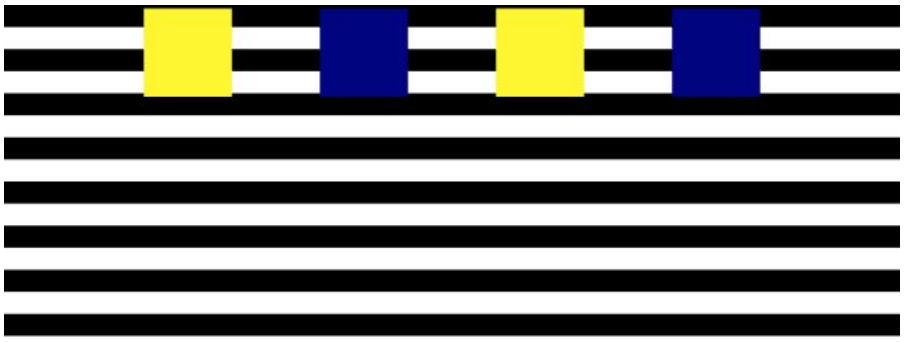


FIGURE 5.6: Illustration on human eye's slow response on blue light[1].

value $h$ in the CIELAB color space as an important factor and together with physiological observations from human eyes to obtain some form of color-dependent compensation in image lightness estimation. The proposed lightness estimation method can better preserve the structural and lightness information in images and can benefit to applications like color-to-grayscale transform, serving as the first step in various image processing applications.

[1]https://gizmodo.com/believe-it-or-not-these-four-squares-move-at-the-same-1559629417

### 5.2.3 Proposed Adaptive Lightness Compensation Method

The proposed lightness calculation is still based on CIELAB color space, since in general this color space over performs others in lightness estimation and in the consistency with human eyes. On the parts that CIELAB does not perform well, we propose to provide a compensation to CIELAB. The proposed lightness compensation method is described as:

$$L = (1 + k) \times L^* \tag{5.2}$$

where $k$ a wavelength-depend compensation factor based on human observers.

The compensation factor is calculated based on the latest experimental observation from human observers on human's spectral sensitivities. The experimental results are taken from the Colour & Vision Research laboratory database[2]. Fig. 5.7 plots the difference between the observed human's spectral sensitivity and the spectral sensitivity calculated from the photonic luminous efficiency function $V(\lambda)$ adopted by CIELAB. Using polynomial regression up to the power of 6 on the plot in Fig. 5.7, we modelled the relationship between the compensation factor $k$ and wavelength $\lambda$ as

$$\begin{aligned} k =& 1 \times 10^{14} \times \lambda^6 - 4 \times 10^{-11} \times \lambda^5 + 6 \times 10^{-8} \times \lambda^4 \\ & - 5 \times 10^{-5} \times \lambda^3 + 0.0236 \times \lambda^2 - 5.7907 \times \lambda + 583.79 \end{aligned} \tag{5.3}$$

The wavelength $\lambda$ of the light is strongly related to hue angle $h$ defined by CIE, because hue angle was originally created to describe the dominant wavelength of a color. However the relationship between the wavelength $\lambda$ and hue angle $h$ is not simply a linear relationship. According to a description from Ostwald in [106], the numerical relationship between hue and wavelength is plotted in Fig. 5.8.

[2]http://www.cvrl.org

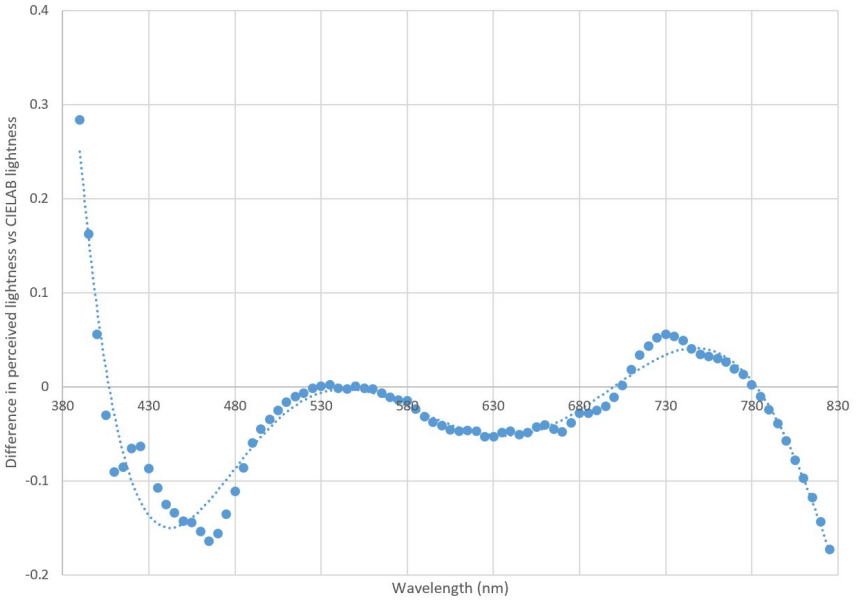



FIGURE 5.7: Error between perceived lightness and CIELAB $L^*$.

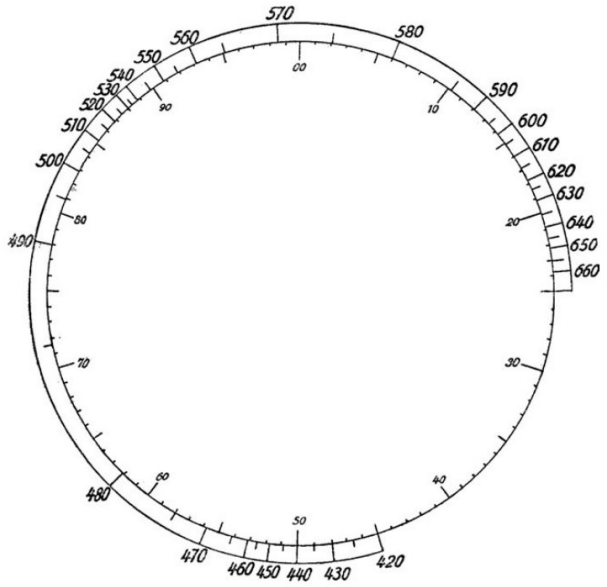



FIGURE 5.8: Ostwald's representation of hue circle with the related wavelengths plotted on the outside and corresponding hue on the inside.

Thus, a mapping function $\lambda = f(h)$ to map the hue angle $h$ to wavelength $\lambda$ could be found through this look-up table, and the compensation factor $k$ can be then calculated accordingly. By applying polynomial regression on the plotted data, we can mathematically express the relationship between wavelength and hue angle as:

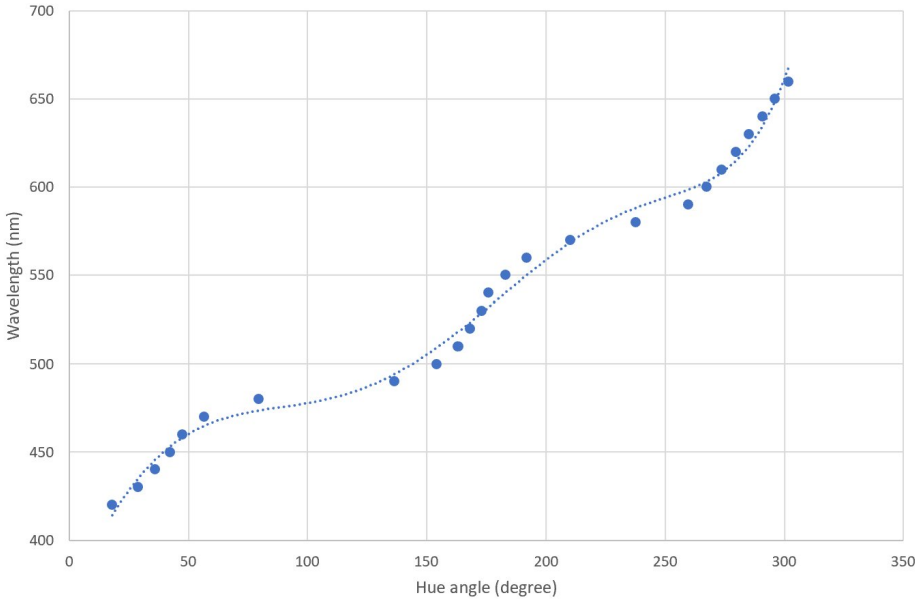



FIGURE 5.9: Plot of hue vs wavelength and regression.

$$\begin{aligned}\lambda =& 1 \times 10^{-11} \times h^6 - 1 \times 10^{-8} \times h^5 + 2 \times 10^{-6} \times h^4 \\ & - 9 \times 10^{-5} \times h^3 - 0.0246 \times h^2 + 3.1122 \times h + 366.63\end{aligned} \tag{5.4}$$

where $h$ is the hue angle that can be calculated using equation (2.23).

Combing equations (3.24), (3.25) and (3.26) together, the proposed lightness of images can then be calculated.

### 5.2.4 Use Case 1: Proposed Color Model for Image Quality Assessment Applications

In various image processing and computer vision algorithms dealing with color images, the first step is to convert the color images to grayscale images, followed by a series of feature extraction operations, feature enhancement operations, etc. This also applies to the image quality assessment applications. The main reason why

grayscale representations are often used for extracting features instead of operating on full color images directly is that a well-captured grayscale image can preserve all structural information in the color image while simplifies the algorithm and reduces computational requirements.

YIQ color space is the most commonly used color space during color to grayscale transform. One possible reason is that this is the default method used by Matlab® built-in *rgb2color* function. Some IQA algorithms such as [12, 19, 83] adopted this color space as their working color space.

The Gaussian color model was very popular and has been adopted in many image quality assessment models such as [15, 16, 74, 107–109] over the past few years. The $L$ component has been used as the luminance component in the feature extraction steps in those IQA models. However, this physical reflectance model may not be suitable for image quality assessment on screen content images (SCIs), since most objects in SCIs are charts, texts, computer generated graphs, etc., which are not real objects that reflect light. It does not make sense if a physical model for real-world object light reflection is applied on SCIs. In our experiment, it is observed that $L$ in this color space sometimes fails to capture a lot of edge information in SCIs during the transformation from RGB to the luminance.

The first observation on the Gaussian color model is that it has huge difficulties handling the difference between red and black, which is a very common color scheme used in the dashboard in stock market, as shown in Fig. 5.10 (a). Fig. 5.10 (b) is the $L$ component from the Gaussian color model. It can be seen that the contrast/edge between red and black is almost lost completely. Also, the luminance component in the Gaussian color model also has difficulties to distinguish the edge between light blue and yellow either, as illustrated in Fig. 5.11 (a) and (b). Since a lot of IQA methods use the edge features as the critical assessment criteria [74, 83, 110], it is important to preserve all edge information during color space transform. Besides

edge width and edge direction, edge contrast is also an important property to be preserved and modeled similar to human perceptions.

The second observation on the Gaussian color model is that although in some cases the edges in the luminance component are not completely lost, the edge contrast is not correctly preserved and is not consistent with human observations. For example, the red text "GAMES" in Fig. 5.12 (a) and the purple "teapot" in Fig. 5.13 (a) are strongly distinguishable against its surroundings in the color image, but the edge contrast in the L component is suppressed, as shown in Fig. 5.12 (b) and Fig. 5.13 (b) respectively.

Fig. 5.10 (c), Fig. 5.11 (c), Fig. 5.12 (c) and Fig. 5.13 (c) shows the $L^*$ component of the images in CIELAB color space. Edges in the $L^*$ component can be better preserved compared to those in the Gaussian color model. However, in the case in Fig. 5.13, the line pointed by the red arrow is almost lost in the $L^*$ component. The reason is that the photopic luminous efficiency function, CIE 1924 $V(\lambda)$, used for calculate the luminance is invalid for heterochromatic brightness matches [102].

Utilizing our proposed color space for image lightness estimation can solve the above mentioned problems successfully. Fig. 5.10 (d), Fig. 5.11 (d), Fig. 5.12 (d) and Fig. 5.13 (d) shows our proposed method can better preserve the structural information including the edge contrast in the lightness image. Significant performance improvements can be seen in the existing image quality assessment models by only incorporating our proposed color space while all other algorithms remaining unchanged. Detailed experimental results will be illustrated in Section 3.4.1.

### 5.2.5 Use Case 2: Proposed Color Model for Image Enhancement Applications

Image enhancement is used to improve the quality of an image for visual perception of human beings. It can be also used as preprocessing procedures to enhance the performance of many computer vision tasks such as objection.

A difficult situation in image enhancement algorithms is to keep the image's original color away from distortion. Thus, such image enhancement task is typically done by converting the image to a color space that has image luminosity as one of its components, such as the CIELAB color space. Contrast adjustment and detail enhancement is then performed on the luminosity layer only. After that, the image is converted back to the RGB color space for storage and display. Manipulating luminosity only affects the intensity of the pixels, while preserving the original colors.

Unsharp masking (UM) is one of the most widely-used image enhancement techniques for improving the sharpness of image. This method first extracts the detail layer from the input image, such as edges by applying a high-pass filter, followed by subtracting it from the original image to generate a base layer. Then the detail layer is amplified by multiplying a scaling factor and then added back to the base layer to generate the enhanced image.

Blurriness-guided unsharp masking (BUM) [93] is the most state-of-the-art variation of UM with the best visual performance. In the BUM operation, the first step is to convert the color image into CIELAB color space and then such enhancement will be applied on the luminance component $L^*$ only. The final enhanced image is generated by combining the enhanced luminance component with the original chrominance components $a^*$ and $b^*$, followed by converting them back to the RGB color space.

In this chapter, a further improvement in the image enhancement performance is demonstrated by applying BUM using the proposed color space. Visual comparison of the enhancement results using different color spaces is shown in Section 3.4.2.

## 5.3 Experimental Results

### 5.3.1 Evaluations on Image Quality Assessment Use Case

The same setup and database as described in Section 2.4 was used here. Two state-of-the-art IQA models, namely GFM [74] and LGFM model proposed in Chapter 2 is chosen here as representative IQA models to evaluate the performance improvement brought by the proposed color space.

In this experiment, firstly, the proposed lightness component $L$ with color-dependent compensation was extracted from both the reference SCI and the distorted SCI. Secondly, either the Gabor filter or the log-Gabor filter was used to generate the corresponding feature maps on the compensated lightness component $L$. Thirdly, the degree of similarity between the set of the feature maps will be measured. At last, the feature similarity score is fused with the chrominance component similarity score, to arrive at a final IQA evaluation score.

In both IQA models, their performance with the originally used Gaussian color model, the CIELAB color space as well as the proposed color space based on human perception were evaluated. The performance results with the proposed color model are compared with the classical and state-of-the-art color models and documented in Table 5.1. Overall, the proposed color model achieves the highest PLCC and SROCC values and the lowest RMSE value. This indicates that the proposed color model can benefit to the existing IQA models to make them more consistent with subjective judgments agreed by the human visual system, compared with other state-of-the-art color models.

TABLE 5.1: INFLUENCE FROM COLOR SPACE ON IQA PERFORMANCE

| | Distortions | GFM | | | LGFM | | |
|---|---|---|---|---|---|---|---|
| | | LMN Space | CLELAB $L^*$ | Proposed $L$ | LMN Space | CLELAB $L^*$ | Proposed $L$ |
| PLCC | GN | 0.9497 | 0.9639 | 0.9612 | 0.9599 | 0.9654 | 0.9661 |
| | GB | 0.9156 | 0.9129 | 0.9075 | 0.9309 | 0.9135 | 0.9190 |
| | MB | 0.9023 | 0.8917 | 0.8903 | 0.9002 | 0.8939 | 0.8954 |
| | CC | 0.8787 | 0.8744 | 0.8705 | 0.8336 | 0.8888 | 0.8897 |
| | JPEG | 0.9392 | 0.9266 | 0.9231 | 0.9318 | 0.9229 | 0.9313 |
| | J2K | 0.9226 | 0.9314 | 0.9233 | 0.9453 | 0.9355 | 0.9434 |
| | CSC | 0.8728 | 0.8987 | 0.8920 | 0.8759 | 0.9037 | 0.9062 |
| | HEVC-SCC | 0.8740 | 0.8918 | 0.8929 | 0.9036 | 0.9120 | 0.9114 |
| | CQD | 0.8928 | 0.9185 | 0.9184 | 0.9203 | 0.9296 | 0.9351 |
| | Overall | 0.8760 | 0.8873 | **0.8895** | 0.9023 | 0.9015 | **0.9067** |
| SROCC | GN | 0.9370 | 0.9530 | 0.9506 | 0.9511 | 0.9651 | 0.9577 |
| | GB | 0.9081 | 0.9074 | 0.9026 | 0.9287 | 0.9122 | 0.9179 |
| | MB | 0.8892 | 0.8843 | 0.8805 | 0.8933 | 0.8849 | 0.8845 |
| | CC | 0.8225 | 0.8263 | 0.8156 | 0.8575 | 0.8490 | 0.8496 |
| | JPEG | 0.9281 | 0.9251 | 0.9235 | 0.9324 | 0.9205 | 0.9261 |
| | J2K | 0.9085 | 0.9224 | 0.9176 | 0.9342 | 0.9341 | 0.9372 |
| | CSC | 0.8736 | 0.9014 | 0.8968 | 0.8901 | 0.9049 | 0.9097 |
| | HEVC-SCC | 0.8712 | 0.8838 | 0.8850 | 0.8914 | 0.8962 | 0.8963 |
| | CQD | 0.8907 | 0.9127 | 0.9122 | 0.9152 | 0.9199 | 0.9223 |
| | Overall | 0.8759 | 0.8878 | **0.8905** | 0.9046 | 0.9061 | **0.9094** |
| RMSE | GN | 3.9378 | 3.3477 | 3.4686 | 3.5248 | 3.2798 | 3.2465 |
| | GB | 4.2566 | 4.3228 | 4.4470 | 3.8679 | 4.3086 | 4.1750 |
| | MB | 4.6121 | 4.9467 | 4.9785 | 4.7625 | 4.9002 | 4.8661 |
| | CC | 4.2732 | 4.3426 | 4.4062 | 4.9446 | 4.1036 | 4.0866 |
| | JPEG | 5.2011 | 5.6552 | 5.7810 | 5.4587 | 5.7867 | 5.4773 |
| | J2K | 6.1385 | 5.7956 | 6.1212 | 5.1898 | 5.6640 | 5.2950 |
| | CSC | 4.8031 | 4.3165 | 4.4544 | 4.7744 | 4.2174 | 4.1603 |
| | HEVC-SCC | 6.7590 | 6.2956 | 6.2635 | 5.9596 | 5.7056 | 5.7263 |
| | CQD | 5.7592 | 5.0609 | 5.0645 | 5.0151 | 4.7220 | 4.5321 |
| | Overall | 6.8310 | 6.5304 | **6.4712** | 6.1052 | 6.1294 | **5.9744** |

### 5.3.2 Evaluations on Image Enhancement Use Case

Extensive experiments have been conducted on a variety of test images to evaluate the performance improvement brought by the proposed color space. As previously mentioned, the proposed color space is a general color space that can be deployed with various image enhancement methods. In this work, we tested with the most representative ones using blurriness-guided unsharp masking (BUM) as the benchmark enhancement method.

Firstly we tested such technique with some general color images. Fig. 5.14 (a) and (b) are two images with colors of similar luminance across its entire image but different color in each region of the image, taken from Cadik's dataset [111]. Fig. 5.14 (c) and (d) show the enhancement results on the test image produced by

applying BUM on its $L^*$ component CIELAB color space as originally proposed by the author. Fig. 5.14 (e) and (f) show the enhancement results by applying BUM on the proposed color space in this thesis. It can be observed from Fig. 5.14 (c) and (d) that applying BUM on the $L^*$ component as proposed in the previous work hardly brings any improvement in the image sharpness. Visually they are almost the same as the original images. While in Fig. 5.14 (e) and (f), the contrast between the color fragments in both images becomes stronger, making both images looks sharper.

Another interesting and promising finding on the proposed color space is that it is specifically suitable to be used in fluoresce microscopy image postprocessing, an imaging technique widely used in biology science. The reason is that in fluoresce microscopy images, dark blue color against pure black background is a very common color scheme appears in such images. This is exemplified by the blue DAPI dye used to stain nucleic acids in biological samples. As we mentioned before in Section 3.2, blue details against black are hard to be resolved by human eyes. Advantageously, the proposed color space can provide compensation for color lightness at the blue range. Applying BUM using the proposed color space can relief this problem to a great extent.

Fig. 5.15 (a) is a fluoresce microscopy image without further postprocessing. The region of interest (ROI) were those cells stained with blue color and a zoom-in image is shown in Fig. 5.15 (b). Biologists are interested to observe the shape of the blue cells as well as counting the number of the cells. However, the blue cells are extremely difficult to be distinguished from the background and each other. Not to mention that in the center part of the ROI, the cells are hardly visible. Fig. 5.15 (c) and (d) show the enhancement results on the test image by applying BUM on its $L^*$ component CIELAB color space as originally proposed by the authors in their previous study. Fig. 5.15 (e) and (f) show the enhancement results by applying BUM on the proposed color space in this thesis. It can be observed

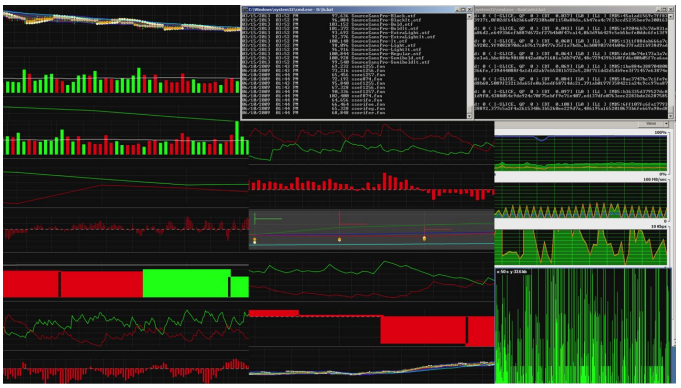
that applying BUM on the $L^*$ component as proposed in the previous work hardly brings any improvement. Visually they are almost the same as the original images. While in Fig. 5.15 (e) and (f), the contrast between the blue cells becomes stronger against the background and its neighbor cells, making such images much easier to be analyzed by biologists. Fig. 5.16 and Fig. 5.17 give two more examples in this use case.

Thirdly, it is observed that applying BUM with the proposed color space can improve the performance of object detection algorithms. This proves that the proposed color space is a good candidate for image preprocessing. Fig. 5.18 (a) is a microscopy images of tissue with haematoxylin and eosin (H&E) stain, one of the most fundamental techniques used in histology. Biologists are interested to segment the regions stained with dark blue or dark purple color (nucleus) and researchers in computer vision have developed object detection and segmentation algorithms for this purpose [3]. The detection result on the original microscopy image is shown in Fig. 5.18 (a) with detection bounding boxes around the nucleus. Fig. 5.18 (b) is the detection result using the same detection algorithm but the image was enhanced by applying BUM on its $L^*$ component. It can be seen that although the nuclei at the bottom left corner was successfully detected as an improvement over the original image, two nucleus on the top and middle right of the image were missed as sacrifice. As a result, the overall detection performance dropped. Fig. 5.18 (c) shows the detection result on the enhanced image by applying BUM on the proposed color space in this thesis. In this case, two more nucleus were detected as a bonus.

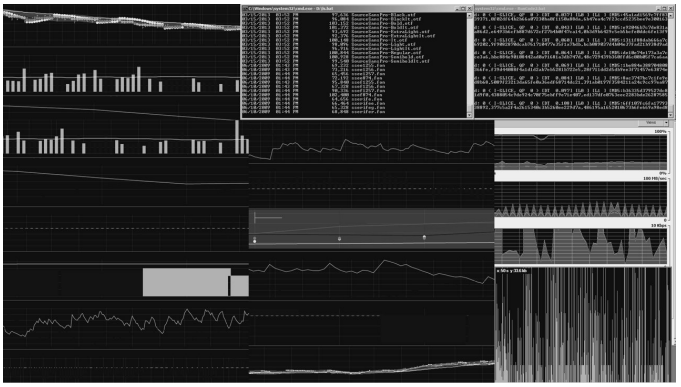
[3] https://github.com/matterport/Mask-RCNN/tree/master/samples/nucleus

## 5.4 Summary

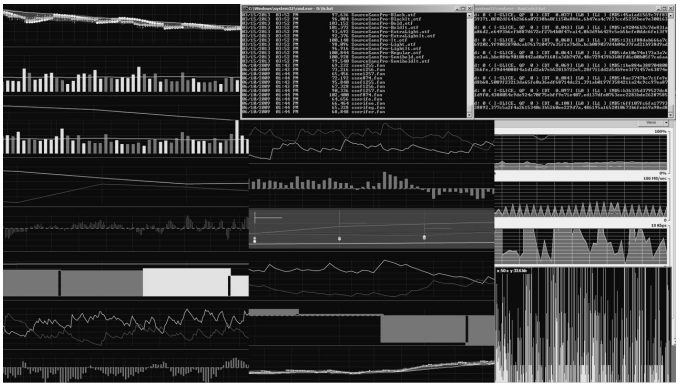
In this chapter, a novel color space for image lightness estimation based on human perception is proposed that incorporates the physiological visual response from real human observers. The proposed color space is motivated by several key observations which suggest that color has a huge impact on human's perceived lightness. Based on such observations, a color-dependent lightness compensation method is developed to generate an accurate lightness estimation of each image pixel. Extensive simulations on various test images have demonstrated that the proposed method can be adopted in many image processing algorithms to deliver superior perceptual quality compared with those algorithms using the conventional and other existing color spaces.



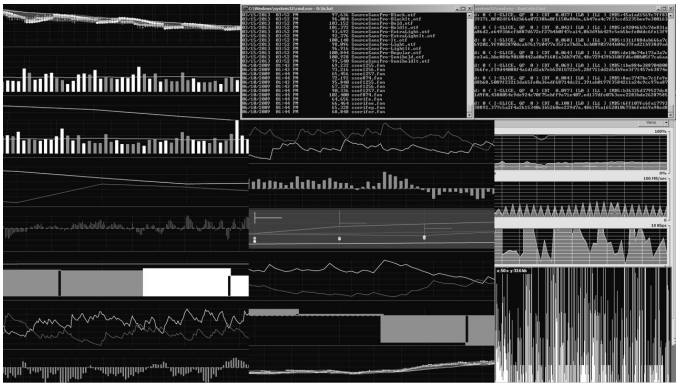
(a)



(b)

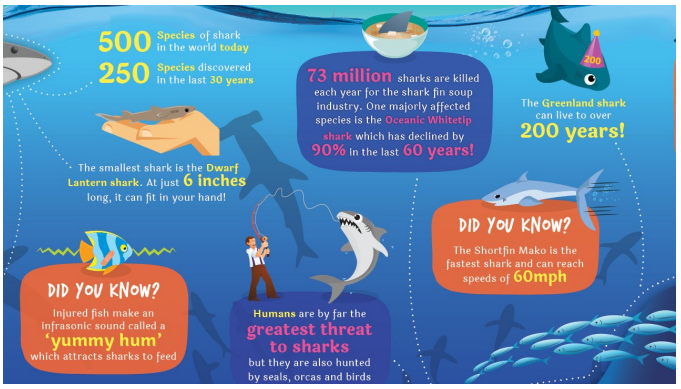


(c)

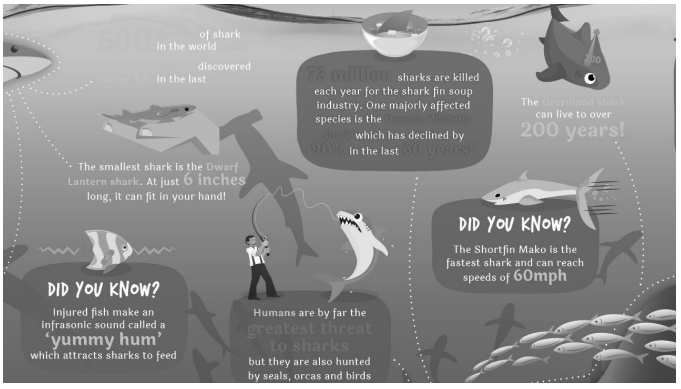


(d)

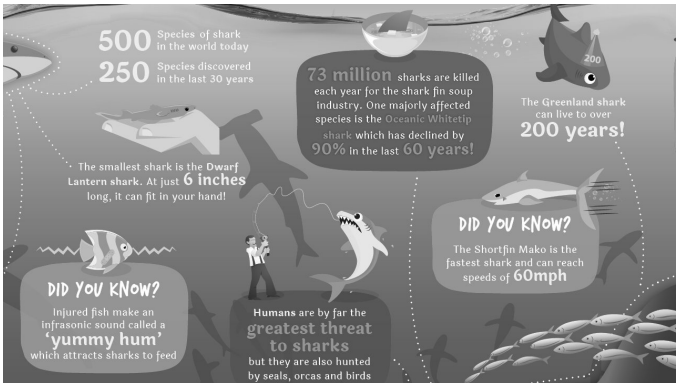
FIGURE 5.10: SCI example 1: (a) RGB color SCIs from SCID database. (b) the $L$ component of the LMN color space. (c) the $L^*$ component of the CIELAB color space. (d) proposed $L$ with color-dependent compensation.



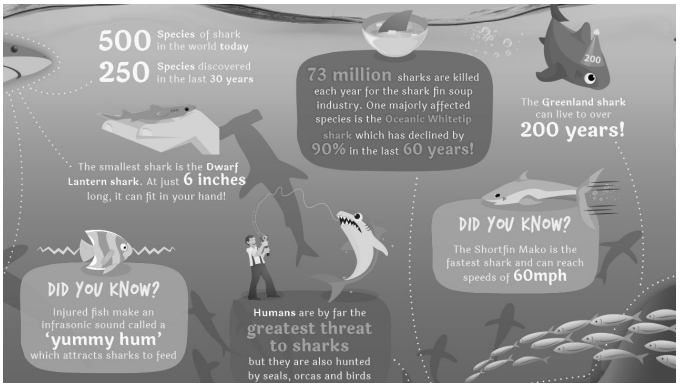

(a)




(b)



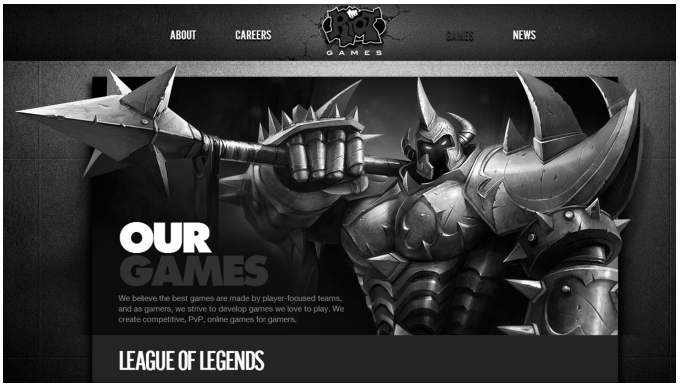

(c)



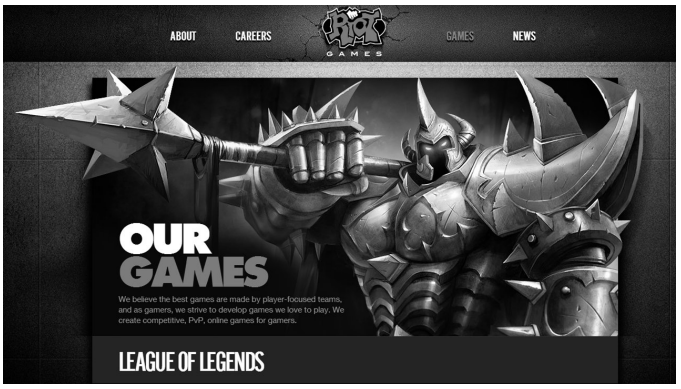

(d)

FIGURE 5.11: SCI example 2: (a) RGB color SCIs from SCID database. (b) the $L$ component of the LMN color space. (c) the $L^*$ component of the CIELAB color space. (d) proposed $L$ with color-dependent compensation.


(a)



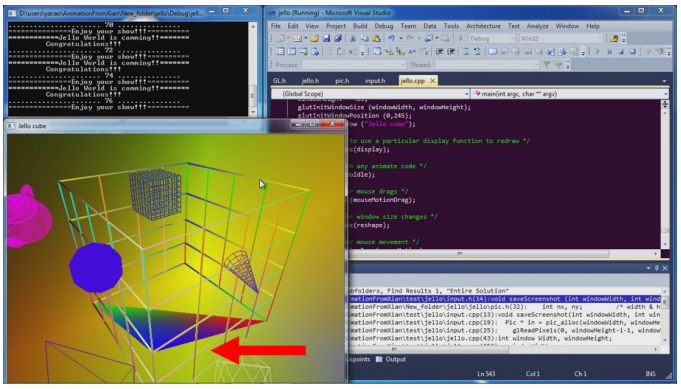

(b)



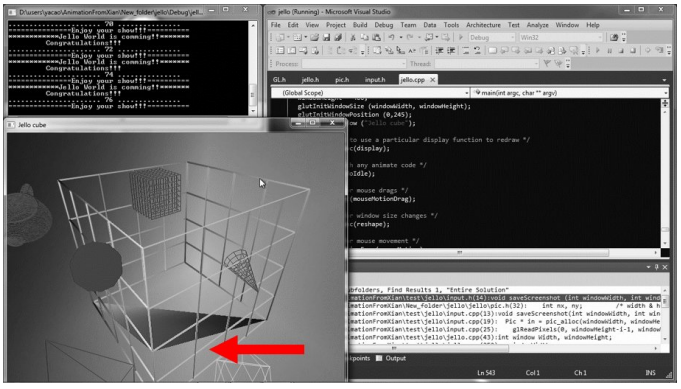

(c)


(d)

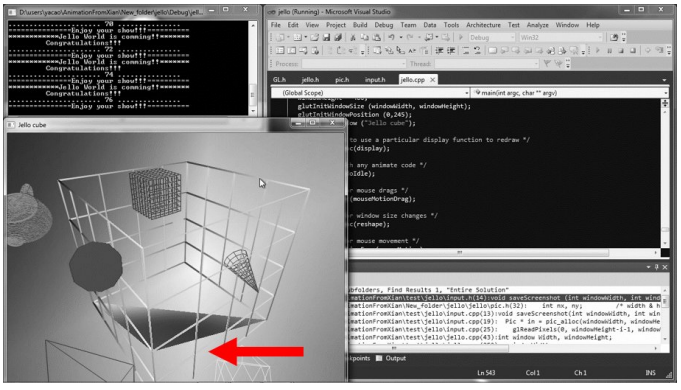
FIGURE 5.12: SCI example 3: (a) RGB color SCIs from SCID database. (b) the $L$ component of the LMN color space. (c) the $L^*$ component of the CIELAB color space. (d) proposed $L$ with color-dependent compensation.



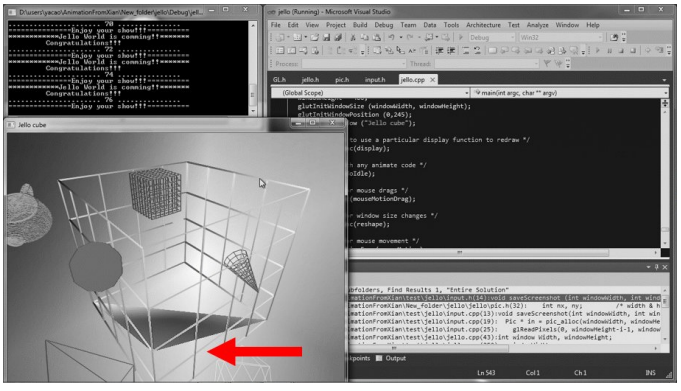
(a)



(b)

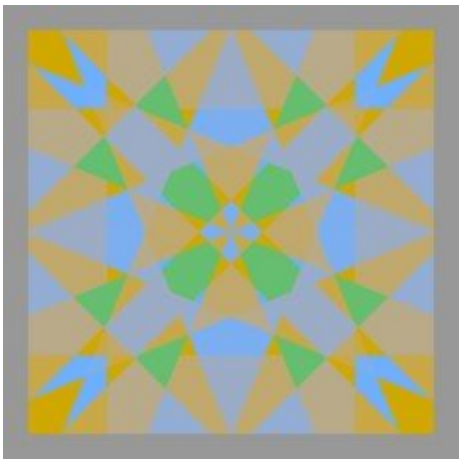


(c)



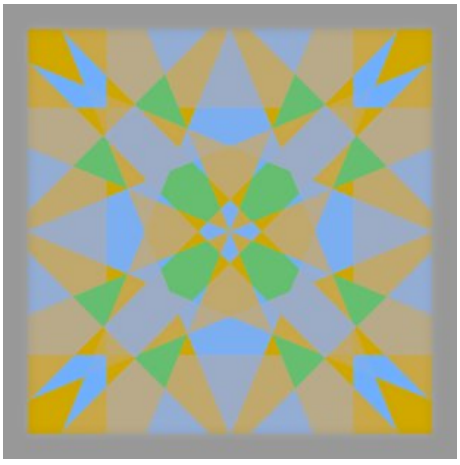
(d)

FIGURE 5.13: SCI example 4: (a) RGB color SCIs from SCID database. (b) the $L$ component of the LMN color space. (c) the $L^*$ component of the CIELAB color space. (d) proposed $L$ with color-dependent compensation.

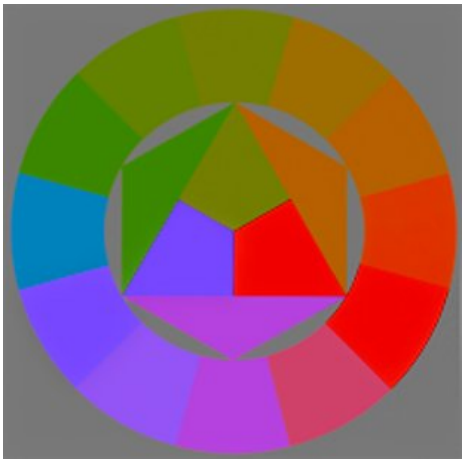


(a) Original image “color fragment”.

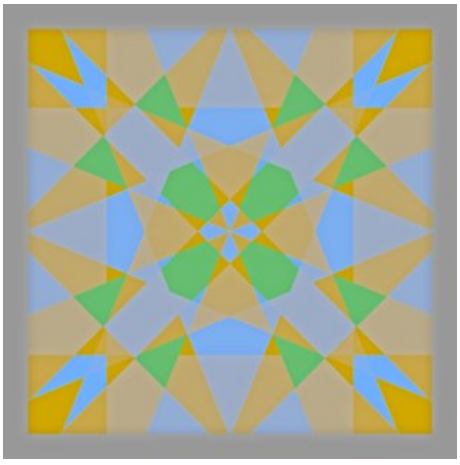
(b) Original image “color wheel”.



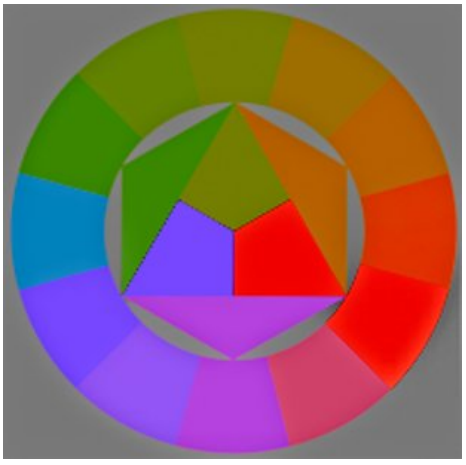
(c) BUM with CIELAB color space



(d) BUM with CIELAB color space

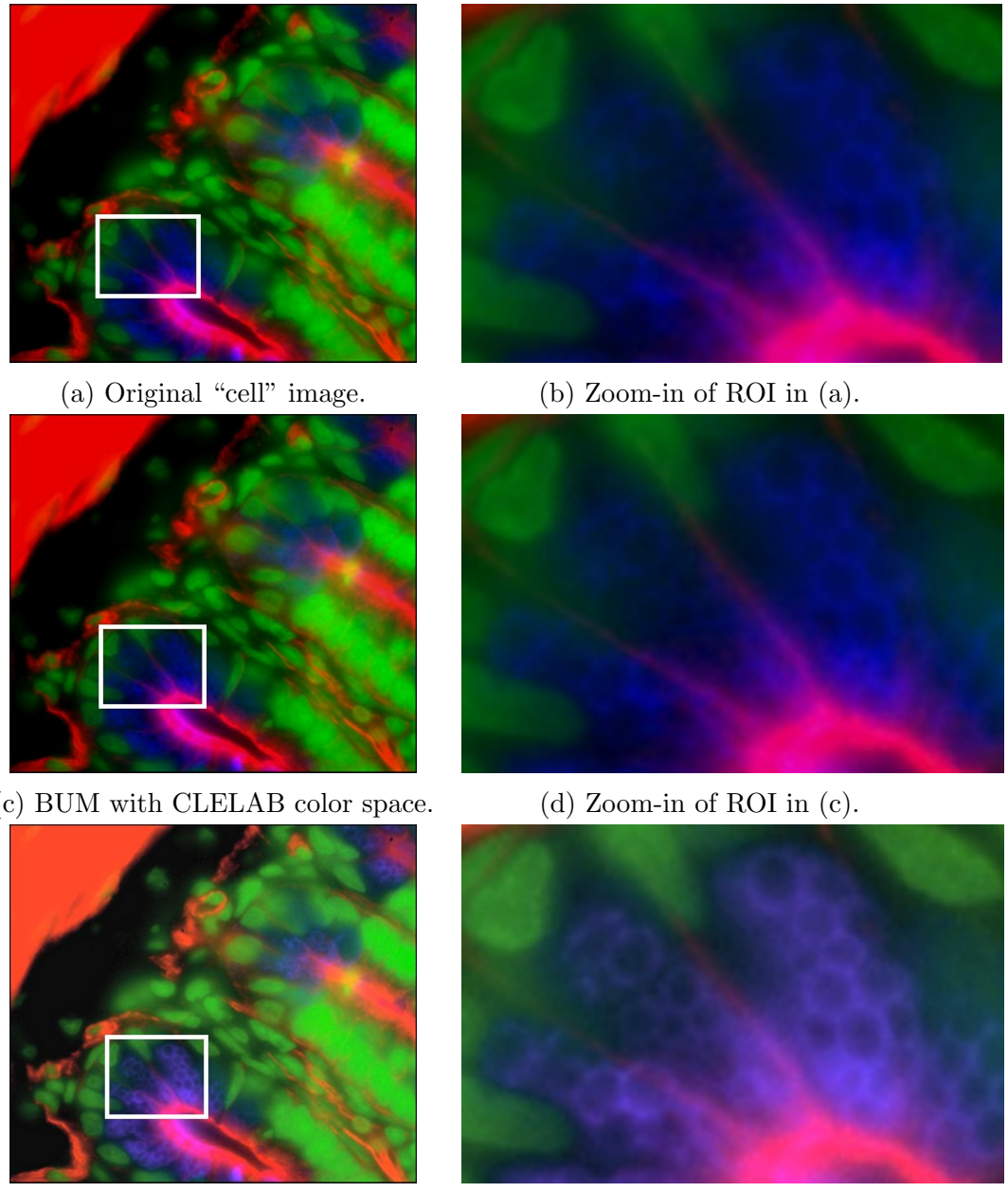


(e) BUM with proposed color space



(f) BUM with proposed color space

FIGURE 5.14: Comparison of the BUM enhancement on color fragment images using different color spaces.

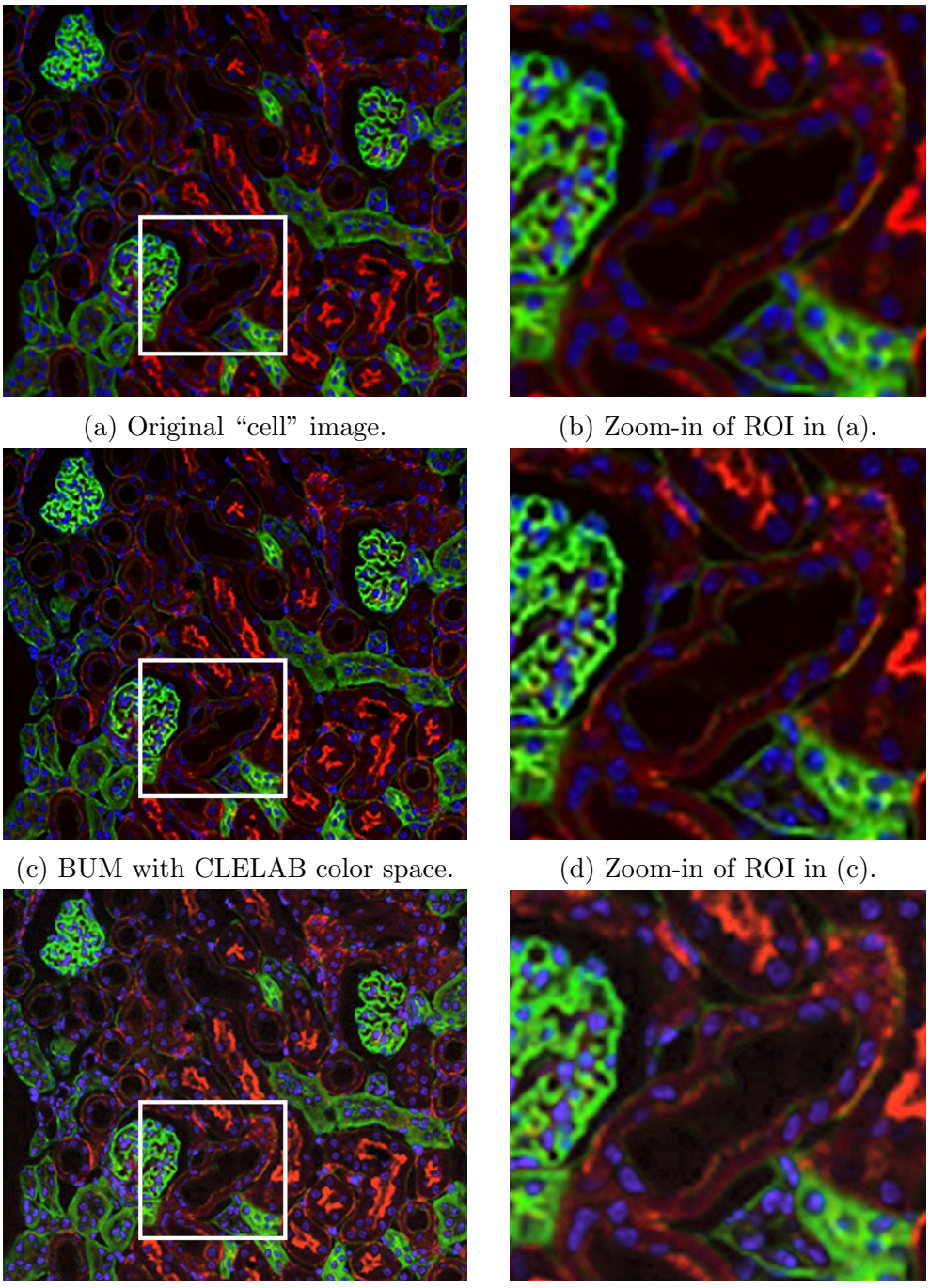


(a) Original "cell" image. (b) Zoom-in of ROI in (a).

(c) BUM with CLELAB color space. (d) Zoom-in of ROI in (c).

(e) BUM with proposed color space. (f) Zoom-in of ROI in (e).

FIGURE 5.15: Comparison of the BUM enhancement on fluoresce microscopy images using different color spaces.

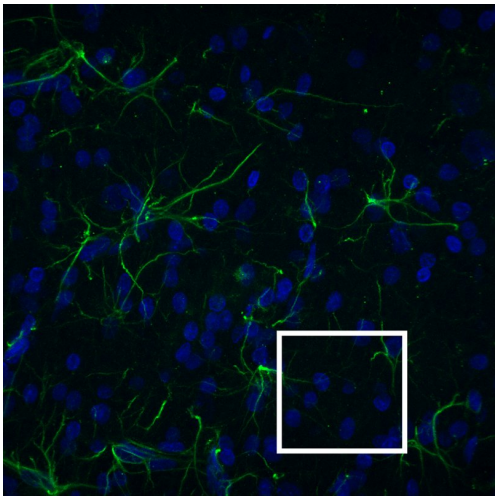


(a) Original "cell" image. (b) Zoom-in of ROI in (a).

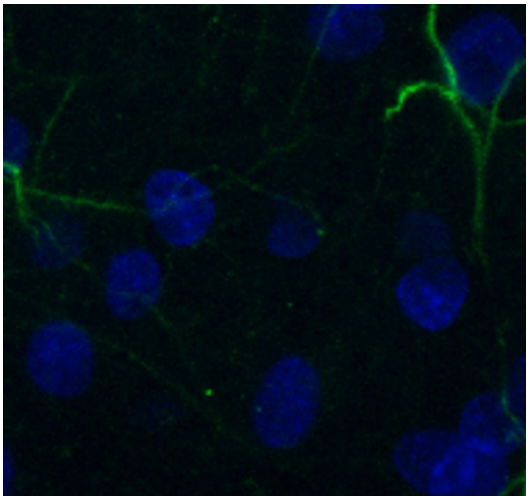
(c) BUM with CLELAB color space. (d) Zoom-in of ROI in (c).

(e) BUM with proposed color space. (f) Zoom-in of ROI in (e).

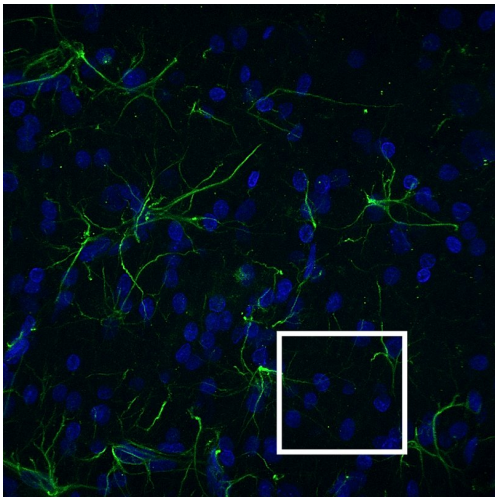
FIGURE 5.16: Comparison of the BUM enhancement on fluoresce microscopy images using different color spaces.



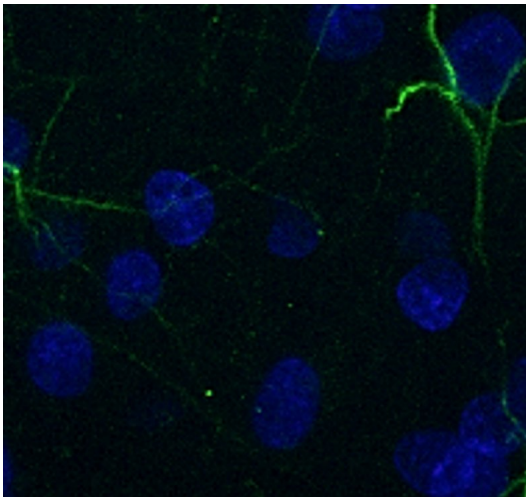
(a) Original "cell" image.



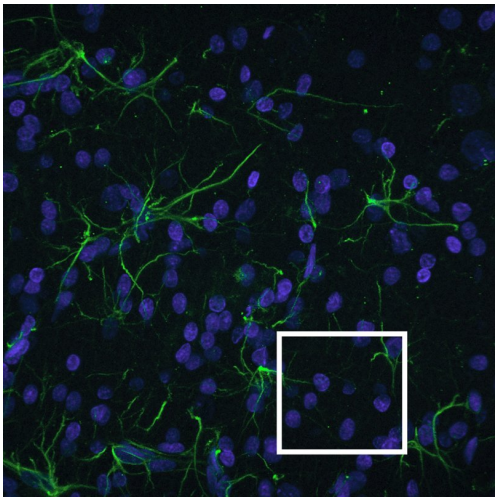
(b) Zoom-in of ROI in (a).



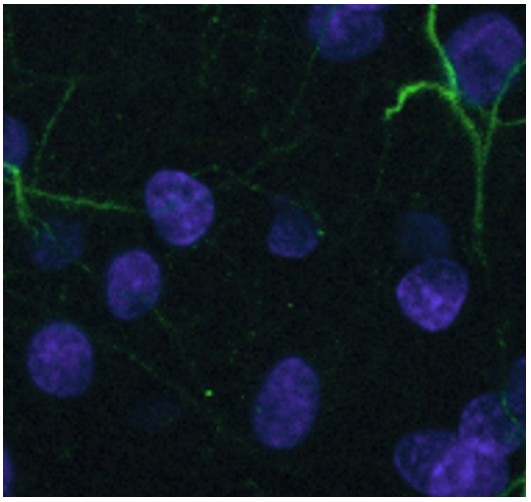
(c) BUM with CLELAB color space.



(d) Zoom-in of ROI in (c).

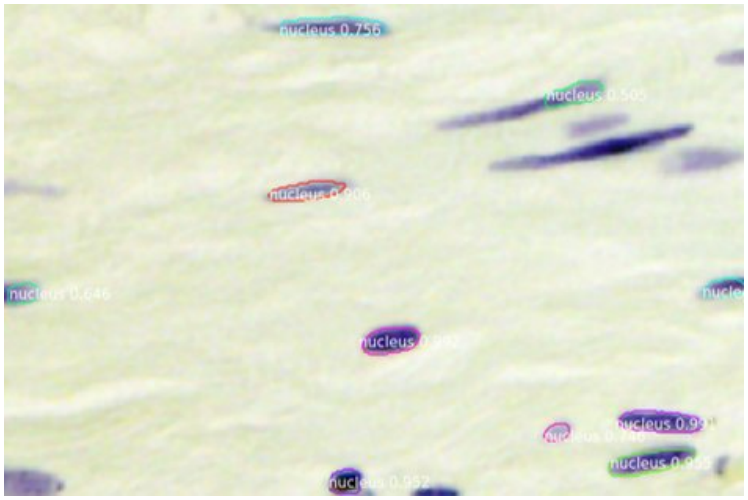


(e) BUM with proposed color space.

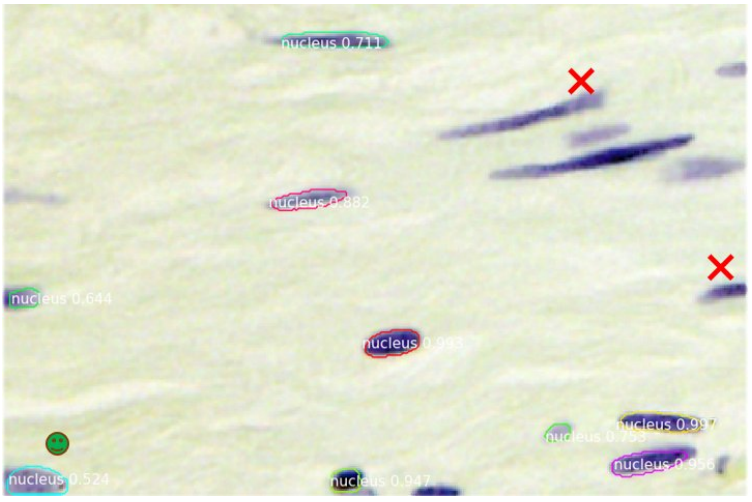


(f) Zoom-in of ROI in (e).

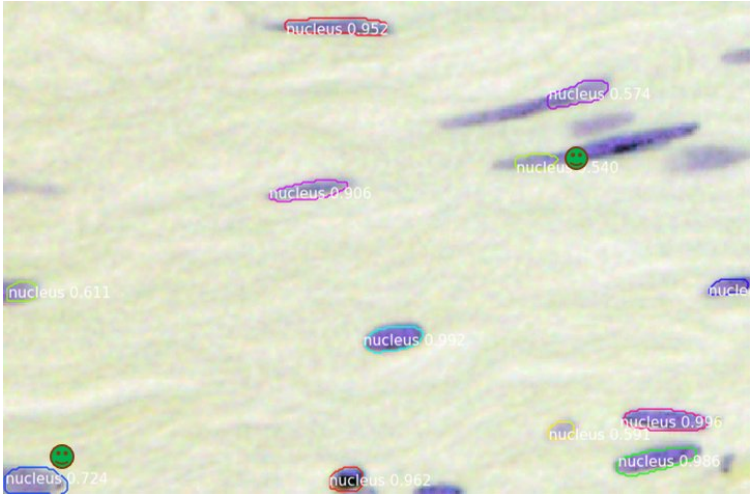
FIGURE 5.17: Comparison of the BUM enhancement on fluoresce microscopy images using different color spaces.



(a) Original "cell" image.



(b) BUM with CLELAB color space.



(c) BUM with proposed color space.

FIGURE 5.18: Comparison of the BUM enhancement on improving object detection rate. Source image is provided by 2018 Data Science Bowl Challenge in https://www.kaggle.com/c/data-science-bowl-2018.

# Chapter 6

# Conclusion and Recommendations

## 6.1 Conclusion

In this thesis, three fundamental topics on the quality of screen content color images have been investigated in this thesis, namely image quality assessment, image blurriness estimation, and image lightness estimation. The contributions in each topic were presented in Chapters 2, 3, and 4 in this thesis respectively and summarized as follows.

- Image Quality Assessment for Screen Content Images

  The proposed log-Gabor feature-based image quality assessment model (L-GFM) is a novel extension of the recently-developed Gabor feature-based IQA model (GFM). LGFM has demonstrated much superior performance over its predecessor and over various conventional image quality assessment approaches.

  Unlike the existing Gabor feature-based IQA model [74] which have exploited the Gabor feature of the images, the proposed log-Gabor feature-based model applies the log-Gabor filter instead because of the following two remarkable

and desirable characteristics of the log-Gabor filters: (i) zero response at the DC, and (ii) wider bandwidth hence stronger response at high frequencies. Both of these characteristics are highly beneficial to develop such IQA model that is more accurate and consistent with the perception of visual cortex in our human visual system.

Extensive simulations conducted on the largest screen content images dataset for quality assessment, SCID, have demonstrated that our model has achieved the best performance, compared to the existing state-of-the-art IQA methods.

- Image Blurriness Estimation for Image Enhancement

  In this work, a fast and accurate image blurriness estimation method has been proposed. Firstly, a binary domain Rapid transform was invented, followed by incorporating it into the proposed cyclic rotational invariant local binary pattern feature descriptor, named RT-LBP. With this novel RT-LBP local feature descriptor, the blurriness of different regions in an image can be accurately estimated. The obtained blurriness map is further proved to be useful in image enhancement applications.

  Experiments conducted on various test images have demonstrated that the proposed method is able to deliver superior blurriness estimation results with extremely low computational complexity.

- Color-Adaptive Image Lightness Estimation

  In this work, a color-adaptive compensation method for image lightness estimation is proposed, which takes the understanding of human's reaction to light and color into account.

  The proposed method can benefit both applications of image quality assessment and image enhancement applications as mentioned in the above two points. Just replacing the working color spaces used in the above-mentioned

IQA and image enhancement works with the proposed lightness estimation method in this work, significant improvement in performance can be obtained.

Extensive experiments and simulations showed that the proposed lightness estimation method is able to deliver state-of-the-art image lightness estimation accuracy and can serve as the fundamentals to various higher level image processing tasks.

## 6.2 Recommendations for further research

Two possible future research directions are discussed in this section:

- Reduced reference and no-reference image quality assessment methods

  In general, the IQA models can be roughly categorized into three classes according to the availability of the reference image, or the so-called ground truth image. They are full-reference (FR), reduced-reference (RR), and no-reference (NR). The related work presented in this thesis focuses on FR-IQA models, as the reference image of screen content are easier to obtain and the FR methods can deliver the best assessment accuracy. However, in many practical scenarios, the reference image may not exist or cannot be easily obtained due to various limitations. Hence the need for reduced-reference or even no-reference IQA methods emerges. Therefore, such RR- and NR-IQA methods are worth to be further investigated.

- Video quality assessment and enhancement

  The algorithms in this thesis are all targeting on still images. Although they can be applied in videos processing tasks without any modifications just by processing the video frame by frame, this is not an efficient approach. The spatial and temporal correlations across the video frames are not utilized.

Video streaming services such as online classes, tele-conferences, live video broadcasting surged in the year of 2020. According to a news from Microsoft, its MS Teams software, an online chat and video call software, had an increment of 31 million active users in just over a month from March 2020 to April 2020 due to the influence of the COVID-19 pandemic. Yet many users had experienced annoyance due to low video quality in such online video chat or video conference services.

Therefore, developing more advanced fundamental video quality assessment and enhancement techniques would always be worthwhile as a future work.

# List of Publications

## Conference Proceedings

- H. Guo, K-K. Ma and H. Zeng, "A Log-Gabor Feature-Based Quality Assessment Model for Screen Content Images," *2019 IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019, pp. 4499-4503, doi:10.1109/ICIP.2019.8803491.

# Bibliography


[1] Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1): 98–117, 2009. ISSN 1053-5888.

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

[3] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.

[4] Hugh R Wilson and James R Bergen. A four mechanism model for threshold spatial vision. *Vision Research*, 19(1):19–32, 1979.

[5] David Marr. Vision: A computational investigation into the human representation and processing of visual information. *Free-man, San Francisco, CA*, 1982.

[6] Andrew B Watson. Detection and recognition of simple spatial forms. In *Physical and Biological Processing of Images*, pages 100–114. Springer, 1983.

[7] Fergus W Campbell and JG Robson. Application of fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3):551–566, 1968.

[8] Colin Blakemore and Fergus W Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, 203(1):237–260, 1969.

[9] David Marr and Ellen Hildreth. Theory of edge detection. *Proceedings of Royal Society of London B*, 207(1167):187–217, 1980.

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

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

[12] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. Fsim: A feature similarity index for image quality assessment. *IEEE transactions on Image Processing*, 20(8):2378–2386, 2011.

[13] Anmin Liu, Weisi Lin, and Manish Narwaria. Image quality assessment based on gradient similarity. *IEEE Transactions on Image Processing*, 21(4):1500–1512, 2012.

[14] Wufeng Xue, Lei Zhang, Xuanqin Mou, and Alan C Bovik. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2):684–695, 2014.

[15] L. Zhang, Y. Shen, and H. Li. VSI: A visual saliency-induced index for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 23(10):4270–4281, 2014.

[16] Sung-Ho Bae and Munchurl Kim. A novel image quality assessment with globally and locally consilient visual quality perception. *IEEE Transactions on Image Processing*, 25(5):2392–2406, 2016.

[17] Ke Gu, Leida Li, Hong Lu, Xiongkuo Min, and Weisi Lin. A fast reliable image quality predictor by fusing micro-and macro-structures. *IEEE Transactions on Industrial Electronics*, 64(5):3903–3912, 2017.

[18] Ke Gu, Junfei Qiao, Xiongkuo Min, Guanghui Yue, Weisi Lin, and Daniel Thalmann. Evaluating quality of screen content images via structural variation analysis. *IEEE transactions on visualization and computer graphics*, 24 (10):2689–2701, 2017.

[19] Ying Fu, Huanqiang Zeng, Lin Ma, Zhangkai Ni, Jianqing Zhu, and Kai-Kuang Ma. Screen content image quality assessment using multi-scale difference of gaussian. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(9):2428–2432, 2018.

[20] Lingxuan Zuo, Hanli Wang, and Jie Fu. Screen content image quality assessment via convolutional neural network. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2082–2086. IEEE, 2016.

[21] Vlad Hosu, Hanhe Lin, Tamas Sziranyi, and Dietmar Saupe. Koniq-10k: An ecologically valid database for deep learning of blind image quality assessment. *IEEE Transactions on Image Processing*, 29:4041–4056, 2020.

[22] Le Kang, Peng Ye, Yi Li, and David Doermann. Convolutional neural networks for no-reference image quality assessment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1733–1740, 2014.

[23] Kwan-Yee Lin and Guanxiang Wang. Hallucinated-iqa: No-reference image quality assessment via adversarial learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 732–741, 2018.

[24] Hongyu Ren, Diqi Chen, and Yizhou Wang. Ran4iqa: restorative adversarial nets for no-reference image quality assessment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[25] Sebastian Bosse, Dominique Maniry, Klaus-Robert Müller, Thomas Wiegand, and Wojciech Samek. Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Transactions on Image Processing*, 27(1):206–219, 2017.

[26] Ali Borji, Ming-Ming Cheng, Huaizu Jiang, and Jia Li. Salient object detection: A benchmark. *IEEE transactions on image processing*, 24(12):5706–5722, 2015.

[27] Lijun Wang, Huchuan Lu, Yifan Wang, Mengyang Feng, Dong Wang, Baocai Yin, and Xiang Ruan. Learning to detect salient objects with image-level supervision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 136–145, 2017.

[28] Chengyao Shen and Qi Zhao. Webpage saliency. In *European conference on computer vision*, pages 33–46. Springer, 2014.

[29] Pina Marziliano, Frederic Dufaux, Stefan Winkler, and Touradj Ebrahimi. A no-reference perceptual blur metric. In *Proceedings. International Conference on Image Processing*, volume 3, pages III–III. IEEE, 2002.

[30] Hantao Liu, Junle Wang, Judith Redi, Patrick Le Callet, and Ingrid Heynderickx. An efficient no-reference metric for perceived blur. In *3rd European Workshop on Visual Information Processing*, pages 174–179. IEEE, 2011.

[31] C Li, W Yuan, AC Bovik, and X Wu. No-reference blur index using blur comparisons. *Electronics letters*, 47(17):962–963, 2011.

[32] Doron Shaked and Ingeborg Tastl. Sharpness measure: Towards automatic image enhancement. In *IEEE International Conference on Image Processing 2005*, volume 1, pages I–937. IEEE, 2005.

[33] Cuong T Vu, Thien D Phan, and Damon M Chandler. A spectral and spatial measure of local perceived sharpness in natural images. *IEEE transactions on image processing*, 21(3):934–945, 2011.

[34] Nien F Zhang, Andras Vladar, Michael T Postek, and Robert D Larrabee. A kurtosis-based statistical measure for two-dimensional processes and its applications to image sharpness. *Proceedings of the 2003 Section on Physical and Engineering Sciences*, 2003.

[35] Matej Kristan, Janez Perš, Matej Perše, and Stanislav Kovačič. A bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recognition Letters*, 27(13):1431–1439, 2006.

[36] Jianping Shi, Li Xu, and Jiaya Jia. Just noticeable defocus blur detection and estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 657–665, 2015.

[37] X. Yi and M. Eramian. Lbp-based segmentation of defocus blur. *IEEE transactions on image processing*, 25(4):1626–1638, 2016.

[38] Xiaoyang Tan and Bill Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6):1635–1650, 2010.

[39] Andrew Berkovich, Michela Lecca, Leonardo Gasparini, Pamela A Abshire, and Massimo Gottardi. A 30 $\mu$w 30 fps 110$\times$ 110 pixels vision sensor embedding local binary patterns. *IEEE Journal of Solid-State Circuits*, 50(9): 2138–2148, 2015.

[40] Li Liu, Songyang Lao, Paul W Fieguth, Yulan Guo, Xiaogang Wang, and Matti Pietikäinen. Median robust extended local binary pattern for texture classification. *IEEE Transactions on Image Processing*, 25(3):1368–1381, 2016.

[41] Bin Xiao, Kaili Wang, Xiuli Bi, Weisheng Li, and Junwei Han. 2d-lbp: an enhanced local binary feature for texture image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(9):2796–2808, 2018.

[42] Qiqi Kou, Deqiang Cheng, Huandong Zhuang, and Rui Gao. Cross-complementary local binary pattern for robust texture classification. *IEEE Signal Processing Letters*, 26(1):129–133, 2018.

[43] Mohammad Alkhatib and Adel Hafiane. Robust adaptive median binary pattern for noisy texture classification and retrieval. *IEEE Transactions on Image Processing*, 28(11):5407–5418, 2019.

[44] Nuh Alpaslan and Kazım Hanbay. Multi-scale shape index-based local binary patterns for texture classification. *IEEE Signal Processing Letters*, 27:660–664, 2020.

[45] Xiaochao Zhao, Yaping Lin, Li Liu, Janne Heikkilä, and Wenming Zheng. Dynamic texture classification using unsupervised 3d filter learning and local binary encoding. *IEEE Transactions on Multimedia*, 21(7):1694–1708, 2019.

[46] Xiaochao Zhao, Yaping Lin, and Janne Heikkilä. Dynamic texture recognition using volume local binary count patterns with an application to 2d face spoofing detection. *IEEE Transactions on Multimedia*, 20(3):552–566, 2017.

[47] Jiwen Lu, Venice Erin Liong, Xiuzhuang Zhou, and Jie Zhou. Learning compact binary face descriptor for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(10):2041–2056, 2015.

[48] Weihong Deng, Jiani Hu, and Jun Guo. Compressive binary patterns: Designing a robust binary face descriptor with random-field eigenfilters. *IEEE transactions on pattern analysis and machine intelligence*, 41(3):758–767, 2018.

[49] Lunke Fei, Bob Zhang, Yong Xu, Di Huang, Wei Jia, and Jie Wen. Local discriminant direction binary pattern for palmprint representation and recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(2):468–481, 2019.

[50] Zhihua Xia, Chengsheng Yuan, Rui Lv, Xingming Sun, Neal N Xiong, and Yun-Qing Shi. A novel weber local binary descriptor for fingerprint liveness detection. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50(4):1526–1536, 2018.

[51] Xue Yuan, Xiaoli Hao, Houjin Chen, and Xueye Wei. Robust traffic sign recognition based on color global and local oriented edge magnitude patterns. *IEEE transactions on intelligent transportation systems*, 15(4):1466–1477, 2014.

[52] Meeras Salman Al-Shemarry, Yan Li, and Shahab Abdulla. An efficient texture descriptor for the detection of license plates from vehicle images in difficult conditions. *IEEE transactions on intelligent transportation systems*, 21(2):553–564, 2019.

[53] M Hassaballah, Mourad A Kenk, and Ibrahim M El-Henawy. Local binary pattern-based on-road vehicle detection in urban traffic scene. *Pattern Analysis and Applications*, 23(4):1505–1521, 2020.

[54] Min Zhang, Chisako Muramatsu, Xiangrong Zhou, Takeshi Hara, and Hiroshi Fujita. Blind image quality assessment using the joint statistics of generalized local binary pattern. *IEEE Signal Processing Letters*, 22(2):207–210, 2014.

[55] Jinjian Wu, Weisi Lin, Guangming Shi, and Long Xu. Reduced-reference image quality assessment with local binary structural pattern. In *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 898–901. IEEE, 2014.

[56] Qiaohong Li, Weisi Lin, Jingtao Xu, and Yuming Fang. Blind image quality assessment using statistical structural and luminance features. *IEEE Transactions on Multimedia*, 18(12):2457–2469, 2016.

[57] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[58] Johann Wolfgang Goethe. *Zur Farbenlehre*. BoD–Books on Demand, 2016.

[59] James Clerk Maxwell. On the theory of three primary colours. Royal Institution of Great Britain, 1861.

[60] Mark D Fairchild. *Color appearance models*. John Wiley & Sons, 2013.

[61] Thomas Smith and John Guild. The cie colorimetric standards and their use. *Transactions of the optical society*, 33(3):73, 1931.

[62] William David Wright. A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141, 1929.

[63] John Guild. The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 230(681-693):149–187, 1931.

[64] M.W. Levine. *Levine & Shefner's Fundamentals of Sensation and Perception*. Number v. 1 in Fundamentals of sensation and perception. Oxford University Press, 2000. ISBN 9780198524663.

[65] Robert William Gainer Hunt and Michael R Pointer. *Measuring colour*. John Wiley & Sons, 2011.

[66] Matthew Anderson, Ricardo Motta, Srinivasan Chandrasekar, and Michael Stokes. Proposal for a standard default color space for the internetsrgb. In *Color and imaging conference*, volume 1996, pages 238–245. Society for Imaging Science and Technology, 1996.

[67] International Electrotechnical Commission et al. Multimedia systems and equipment-colour measurement and management-part 2-1: Colour management-default rgb colour space-srgb. *IEC 61966-2-1*, 1999.

[68] Jeffrey J McConnell. *Computer graphics: theory into practice*. Jones & Bartlett Learning, 2005.

[69] Alvy Ray Smith. Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3):12–19, 1978.

[70] Mark Fairchild. Color appearance models: Ciecam02 and beyond. In *Tutorial Notes, IS&T/SID 12th Color Imaging Conference*, 2004.

[71] A Hanbury and J Serra. A 3d-polar coordinate colour representation suitable for image analysis, prip. Technical report, TU Wien, Tech. Rep. PRIP-TR-77, 2002.

[72] John W. Woods. Chapter 6 - image perception and sensing. In John W. Woods, editor, *Multidimensional Signal, Image, and Video Processing and Coding (Second Edition)*, pages 193 – 221. Academic Press, Boston, second edition edition, 2012. ISBN 978-0-12-381420-3. doi: https://doi.org/10.1016/B978-0-12-381420-3.00006-0. URL `http://www.sciencedirect.com/science/article/pii/B9780123814203000060`.

[73] J-M Geusebroek, Rein Van den Boomgaard, Arnold W. M. Smeulders, and Hugo Geerts. Color invariance. *IEEE Transactions on Pattern analysis and machine intelligence*, 23(12):1338–1350, 2001.

[74] Z. Ni, H. Zeng, L. Ma, J. Hou, J. Chen, and K.-K. Ma. A gabor feature-based quality assessment model for the screen content images. *IEEE Transactions on Image Processing*, 27(9):4516–4528, 2018. ISSN 1057-7149.

[75] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America. A, Optics and Image Science*, 2 (7):1160–9, 1985. doi: 10.1364/JOSAA.2.001160.

[76] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. A, Optics and Image Science*, 4(12):2379–2394, 1987.

[77] C. Massot and J. Hérault. Model of frequency analysis in the visual cortex and the shape from texture problem. *International Journal of Computer Vision*, 76(2):165–182, 2008.

[78] D. Boukerroui, J. Noble, and M. Brady. On the choice of bandpass quadrature filters. *Journal of Mathematical Imaging and Vision*, 21(1-2):53–80, 2004.

[79] C. Ronse. On idempotence and related requirements in edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(5):484–491, 1993. ISSN 0162-8828. doi: 10.1109/34.211468.

[80] F. Heitger, L. Rosenthaler, R. Von Der Heydt, E. Peterhans, and O. Kübler. Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Research*, 32(5):963–981, 1992.

[81] P. Kovesi. *Invariant measures of image features from phase information*. PhD thesis, University of Western Australia, 1996.

[82] S. J. Anderson and D. C. Burr. Spatial and temporal selectivity of the human motion detection system. *Vision Research*, 25(8):1147–1154, 1985. ISSN 0042-6989.

[83] Z. Ni, L. Ma, H. Zeng, J. Chen, C. Cai, and K.-K. Ma. ESIM: Edge similarity for screen content image quality assessment. *IEEE Transactions on Image Processing*, 26(10):4818–4831, 2017.

[84] Eric Cooper Larson and Damon Michael Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of electronic imaging*, 19(1):011006, 2010.

[85] Nikolay Ponomarenko, Vladimir Lukin, Alexander Zelensky, Karen Egiazarian, Marco Carli, and Federica Battisti. Tid2008-a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10(4):30–45, 2009.

[86] Nikolay Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, et al. Image database tid2013: Peculiarities, results and perspectives. *Signal processing: Image communication*, 30:57–77, 2015.

[87] Hamid R Sheikh, Muhammad F Sabir, and Alan C Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing*, 15(11):3440–3451, 2006.

[88] Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2):1004–1016, 2016.

[89] Huan Yang, Yuming Fang, and Weisi Lin. Perceptual quality assessment of screen content images. *IEEE Transactions on Image Processing*, 24(11): 4408–4421, 2015.

[90] Shiqi Wang, Ke Gu, Xiang Zhang, Weisi Lin, Li Zhang, Siwei Ma, and Wen Gao. Subjective and objective quality assessment of compressed screen content images. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 6(4):532–543, 2016.

[91] BT Series. Methodology for the subjective assessment of the quality of television pictures. *Recommendation ITU-R BT*, pages 500–13, 2012.

[92] VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment, 2015. URL `http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx`.

[93] Wei Ye and Kai-Kuang Ma. Blurriness-guided unsharp masking. *IEEE Transactions on Image Processing*, 27(9):4465–4477, 2018.

[94] H Reitboeck and TP Brody. A transformation with invariance under cyclic permutation for applications in pattern recognition. *Information and Control*, 15(2):130–154, 1969.

[95] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3883–3891, 2017.

[96] Hongguang Zhang, Yuchao Dai, Hongdong Li, and Piotr Koniusz. Deep stacked hierarchical multi-patch network for image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5978–5986, 2019.

[97] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8174–8182, 2018.

[98] Chih-Hung Liang, Yu-An Chen, Yueh-Cheng Liu, and Winston Hsu. Raw image deblurring. *IEEE Transactions on Multimedia*, 2020.

[99] Janos Schanda. Current cie work to achieve physiologically-correct color metrics. *Color vision: Perspectives from different disciplines*, pages 307–318, 1998.

[100] Mark D Fairchild and Elizabeth Pirrotta. Predicting the lightness of chromatic object colors using cielab. *Color Research & Application*, 16(6):385–393, 1991.

[101] Donald C Hood, Nicholas I Benimoff, and Vivienne C Greenstein. The response range of the blue-cone pathways: a source of vulnerability to disease. *Investigative ophthalmology & visual science*, 25(7):864–867, 1984.

[102] Lindsay T Sharpe, Andrew Stockman, Wolfgang Jagla, and Herbert Jägle. A luminous efficiency function, V*($\lambda$), for daylight adaptation. *Journal of Vision*, 5(11):3–3, 2005.

[103] R Beau Lotto and Dale Purves. The effects of color on brightness. *Nature neuroscience*, 2(11):1010, 1999.

[104] CJ Bartleson and EJ Breneman. Brightness perception in complex fields. *Josa*, 57(7):953–957, 1967.

[105] Andrew Stockman and Lindsay T Sharpe. The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision research*, 40(13):1711–1737, 2000.

[106] Wilhelm Ostwald. *Physikalische Farbenlehre*, volume 2. Verlag Unesma, 1919.

[107] Lin Zhang, Lei Zhang, and Alan C Bovik. A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 24 (8):2579–2591, 2015.

[108] Manyu Jin, Tao Wang, Zexuan Ji, and Xiaobo Shen. Perceptual gradient similarity deviation for full reference image quality assessment. *Computers, Materials & Continua*, 56(3):501–515, 2018.

[109] Zaifeng Shi, Jiaping Zhang, Qingjie Cao, Ke Pang, and Tao Luo. Full-reference image quality assessment based on image segmentation with edge feature. *Signal Processing*, 145:99–105, 2018.

[110] Wufeng Xue, Lei Zhang, Xuanqin Mou, and Alan C Bovik. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2):684–695, 2013.

[111] M Ĉadík. Perceptual evaluation of color-to-grayscale image conversions. In *Computer Graphics Forum*, volume 27, pages 1745–1754. Wiley Online Library, 2008.