ON THE USE OF IMAGE SHARPNESS MEASURE TO JPEG2000 NO-REFERENCE IMAGE QUALITY ASSESSMENT

By

PHONG VAN VU

Bachelor of Arts in Electronics and Telecommunications
Posts and Telecommunications Institute of Technology
Hanoi, Vietnam
2004

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2013

ON THE USE OF IMAGE SHARPNESS MEASURE TO JPEG2000 NO-REFERENCE IMAGE QUALITY ASSESSMENT

Thesis Approved:

Damon M. Chandler

Thesis Advisor

Dr. Martin Hagan

Dr. Guoliang Fan

ACKNOWLEDGMENTS

I would like to thank my advisor, Damon M. Chandler, for his encouragement, interest, and patience. Personally, I would like to thank him for sharing his knowledge which has enriched my study in topology.

This material is based upon work supported by, or in part by, the National Science Foundation, "Content-Based Strategies of Image and Video Quality Assessment," Award 0917014, the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number W911NF-10-1-0015.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: Phong Van Vu

Date of Degree: December, 2013

Title of Study: ON THE USE OF IMAGE SHARPNESS MEASURE TO JPEG2000 NO-REFERENCE IMAGE QUALITY ASSESSMENT

Major Field: Electrical and Computer Engineering

Abstract:

Digital images play an important role in social communities today. Many applications and devices have been developed to capture images, compress and store them, and transfer from/to servers to/from end-users over broadband connection, wireless communications, etc. In most applications, the goal is to maintain the high quality of the images, but how to assess the quality of an image remains a challenge task. Therefore, an effective and robust method of image quality assessment is crucial and required.

One attribute that contributes to the quality of an image is its sharpness level. It is easy and effortless for human subjects to judge sharpness level within an image or cross images. However, the task is still challenging for a computer; only a handful of algorithms can generate a local image sharpness map. Here, we present a simple, yet effective wavelet-based algorithm for estimating both global and local image sharpness (FISH, *F*ast *I*mage *SH*arpness). FISH operates by first decomposing the input image via a three-level separable discrete wavelet transform (DWT). Next, the log-energies of the DWT sub-bands are computed. Finally, a scalar index corresponding to the image's overall sharpness is computed via a weighted average of these log-energies. Testing on several image databases demonstrates that, despite its simplicity, FISH is competitive with the currently best-performing techniques both for local sharpness estimation and for no-reference image quality assessment of blurred images.

It is also known that the destruction of sharp regions due to JPEG2000 encoding reduces visual quality. Therefore, a sharpness/blurriness estimator can be used to estimate quality of JPEG2000-compressed images. In Chapter 4, we propose the EDIQ algorithm, (*ED*ge-based *I*mage *Q*uality), that estimate quality of JPEG2000-compressed images via the edge/near-edge regions, which are defined by applying edge detection and edge-pixel dilation. Then, perceived blurring is estimated by the FISH algorithm and perceived ringing is estimated by the local variance of Laplacian coefficients in the edge/near-edge regions. These local values are combined and collapsed into the final quality index of the image. Testing on various subsets of JPEG2000-compressed images demonstrates the efficacy of EDIQ in predicting the quality of JPEG2000-compressed images.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Digital images play an important role in social communities today. Many applications and devices have been developed to capture images, compress and store them, and transfer from/to servers to/from end-users over broadband connection, wireless communications, etc. In most applications, the owners/producers want to maintain the high quality of the images, but how to assess the quality of an image remains a challenge for the producers. An algorithm of image quality assessment (IQA) can be used to validate the quality of a new compression standard or a new type of transmission before delivery to the end-user. A service system can implement an IQA algorithm to be able to realize and quantify the quality degradations of the images that occur in the system, so that it can maintain, control, and possibly enhance the quality of the data. Therefore, an effective and robust image quality assessment method is crucial and required for many applications.

There are two types of quality assessment methods: the subjective and objective methods. The subjective method is a reliable way of assessing the quality of an image or a video, because human beings are the ultimate receivers in most applications and the method collects ratings from human subjects. These subjective ratings are then carefully processed using outlier rejection, score normalization, etc. to obtain the statistically accurate quality score of the image/video in terms of Mean Opinion Score (MOS) or Difference Mean Opinion Score (DMOS). However, despite the advantages of high reliability and accuracy, the subjective methods have many disadvantages: it is inconvenient, very time-consuming, and too expensive to deploy frequently because

of system settings, subject fees, etc. Therefore, it is difficult and almost impossible to collect customers' ratings about the products and customize them to make it better. Because of these disadvantages, the subjective assessment method is mainly necessary in barely situations where quality must be assured.

The second IQA method is using a computer with a built-in algorithm, this is so-called the objective method. Specifically, the goal of objective method in IQA research is to design algorithms that can predict the perceived image quality automatically and in a manner that agrees with human subjects. A reliable objective quality metric can be employed to monitor quality for quality control systems, and optimize the algorithms and the parameter settings of the processing system such as image compression, image enhancement, etc.

Objective IQA algorithms can be classified according to the availability of the original image, which is often considered to have perfect quality. The original image may be used as a reference to compare with a distorted image; this type is called the full-reference method. Nevertheless, in many practical applications, the reference images are often not accessible but human observers usually can effectively and reliably assess the quality of distorted images without using any reference. Therefore, it is highly desirable to develop algorithms that can evaluate image quality without the knowledge of the reference. These algorithms are called blind or no-reference quality assessment algorithms; developing a no-reference method is even more difficult than a full-reference one.

For image quality assessment, one attribute of the image that contributes to its quality is the sharpness level of the image. Although the term "sharpness" lacks a precise technical definition, it is easy and effortless for human subject to point out the sharp regions in an image and/or compare the sharpness level of the two images. Intuitively, perceived sharpness is a combination of the captured resolution, which cannot be changed in processing, and of acutance, which can be so changed. A sharp

region in the image is the region in which fine details are resolvable (high resolution) and in which edges and object boundaries appear to be of high contrast (high acutance [2]). Indeed, most professional photographers attempt to maximize perceived sharpness by using a high-resolution camera and employing digital retouching to increase acutance (e.g., via unsharp masking). A sharper image often has higher quality than the less sharp image, and the sharpness can be measured without the knowledge of the reference image. Therefore, sharpness can be used to develop a no-reference image quality assessment. In addition, auto-enhancement algorithms can use sharpness information to sharpen images in a spatially adaptive fashion [3]; sharpness can also be an useful factor for main subject detection in photographs [4].

Despite the ease with which that human subjects can judge the sharpness in an image both locally and globally, this task remains quite challenging for a computer. Knowing the important of a sharpness/blurriness estimator, many researchers have developed their own algorithms to estimate sharpness of the image but only a handful of algorithms can generate a local sharpness map. Here, we present our simple yet effective approach to measure image sharpness both globally and locally and the application of our sharpness measure to estimate quality of the JPEG2000-compressed images. The next chapters in this thesis report are organized as follows:

- In Chapter 2, a literature review of the sharpness/blurriness estimators that have been done in previous works is presented. Those methods have been classified into four different categories of algorithms based on their working mechanisms. Those categories are the edge-based methods, pixel-based methods, transformed-based methods, and hybrid methods, which utilize the combination of working mechanisms from the three former methods.

- In Chapter 3, we present our *F*ast *I*mage *Sh*arpness algorithm (FISH) that predict both the global and local image sharpness by examining the log-energy

of the high-frequency components via the wavelet transformation. FISH operates by first decomposing the input image via a three-level separable discrete wavelet transform (DWT). Next, the log-energies of the DWT subbands are computed. Finally, a scalar index corresponding to the image's overall sharpness is computed via a weighted average of these log-energies. The performance of the proposed algorithm is validated in terms of local sharpness estimation and no-reference image quality assessment of blurred images.

- A typical application of the FISH image sharpness algorithm to construct a no-reference image quality assessment algorithm for the JPEG2000-compressed images is proposed in Chapter 4. The algorithm bases on the assumption that the quality of JPEG2000-compressed images can be evaluated by computing the perceived blurring and ringing artifacts in the edge/near-edge regions of the image. The edge/near-edge regions are defined by applying the Canny edge detection and edge-pixel dilation. The local perceived blurring artifact is measured by the FISH algorithm and the local perceived ringing artifact is measured by the local variance of the Laplacian coefficients. These local values are combined and collapsed via root mean square to yield a single scalar that represents the quality of the input JPEG2000-compressed image. Performance evaluation is performed on the JPEG2000-compressed subsets of various image-quality databases.

- Conclusions and potential future studies will be presented in Chapter 5 of the thesis report.

CHAPTER 2

LITERATURE REVIEW

It is important to note that sharpness is not necessarily the inverse of blurriness; an unsharp region might not be blurred, and a blurred region might not be considered as unsharp. However, for most applications on which researchers have tested their algorithms, sharpness and blurriness have been used as two contrary concepts. Therefore, in this chapter, we will review both sharpness and blurriness estimators.

Various algorithms have been proposed to estimate the sharpness/blurriness of the images. These estimators can generally be classified into four main categories: 1) edge-based methods, which involve measuring the spread of edges; 2) pixel-based methods, which work with the activities of image pixels in the spatial domain; 3) transform-based methods, which work in the spectral domain; and 4) hybrid methods, which utilize the combination of the three former methods. This chapter provides a literature review of these four categories of sharpness/blurriness methods.

2.1 Edge-Based Methods

A common technique of sharpness/blurriness estimation involves measuring the spread of edges: the narrower the spread of edges, the sharper the image. Marziliano *et al.* [5] proposed a technique based on the smoothing effects of blur on edges. First, they identify vertical edges in an image by using Sobel edge detector. Then, they scan each row of the image to find local extreme locations closest to edge, which are defined as the start and the end positions of the edge. The local blur measure is defined as the difference between the end and the start position. Finally, the global blur measure for

the whole image is obtained by averaging the local blur values over all edge locations.

Ong *et al.* [6] proposed a different approach for measuring spread of edges using gradient direction. The proposed algorithm involves four following steps. In the first step, gradients direction is defined as the angle between vertical and horizontal gradients. The second step detects edge pixels using Canny edge detector. The third step measures the edge-spread in the direction of the gradients and in the opposite direction of the gradients. The final step computes the image quality measure as a function of average edge-spread, which is computed by dividing the total amount of edge-spread on all edge pixels by the number of edge pixels extracted from the image.

Edge spread can be measured from the line/edge profile [7, 8]. Dijk *et al.* [7] determined the location and orientation of the lines and edges. The profiles of these lines/edges are modeled by using different Gaussian functions. The width and amplitude of the line/edge are then estimated from the response function, which is measured from the Gaussian derivatives at several scales. Sharpness measured is formed by the 5^{th} percentile of Gaussian variances at all points and the number of pixels for which the variance is smaller than one (1). Chung *et al.* [8] presented a technique for a non-parametric blur measure based on edge analysis from edge profile and gradient. For any edge point, the blur is measured by the standard deviation of the edge magnitude profile around that point and the value of the edge gradient magnitude. These values are weighted by the contrast of the image and averaged to yield the perceptual blur measure of the image.

Wu *et al.* [9] constructed a blurriness estimator using the point spread function (PSF). They first detect edge locations using Sobel edge detector and then, extract a sequence of points orthogonal to the edge. The derivative of this sequence is used to obtain the line spread function (LSF); these LSFs are averaged across different locations on the edges. Finally, the PSF parameters are characterized from the LPF and the radius of the PSF is used as the degree of blurriness of the image. In Ref.

10, Liu *et al.* also employs the Sobel edge detector to extract local edge features. A statistical description of an image feature is used to unify these edge features values to a single vector using eleven percentiles. This global descriptor is trained with a circular back-propagation neural network system for blur estimation.

Ferzli and Karam [11] introduced the notion of just noticeable blur (JNB), which is defined as a threshold that human subject can perceive blurriness around an edge, given a contrast higher than the just noticeable difference. The JNB concept is used to estimate the perceived blur distortion within each 64×64 edge block of the image (an edge block is a block that contains a number of edge pixels greater than a threshold). The distortion for each block is then used as the input of a probabilistic model to estimate the perceived blur distortion of the whole image. The image sharpness index is given by the ratio between the total number of edge blocks in the image and the computed perceived blur distortion.

Using the concept of JNB in a different way, Narvekar and Karam [12] estimated the sharpness of an image as the cumulative probability of detecting blur at an edge. The image is divided into 64×64 blocks, each block is classified as edge block or non-edge block depending on the edge information. For each edge pixel in an edge block, the corresponding edge width and the JNB edge width are obtained from the local contrast of the block. The probability of blur at the edge pixel is then computed and compared with the probability of the JNB. As the amount of blur in the image increases, the spread of the edges increases and results in a higher value and a higher probability of blur detection at the considered edge.

The sharpness measure proposed by Corchs *et al.* [13] uses a region-based segmentation algorithm to obtain the segmented image. This image is used to extract and collect all the boundaries between two adjacent regions that are distinct segments. The segments have been selected by considering various features and sharpness is estimated by the median of the spread of edge segments.

2.2 Pixel-Based Methods

Some sharpness/blurriness estimators operate in the spatial domain by examining the activities of the image pixels. Crete *et al.* [14] generated a blurred version of the input image by using a low-pass filter. Then, they computed the intensity variation between neighboring pixels of both the input image and the low-pass filtered (blurred) image. By comparing the two intensity variations, the authors estimate the blur annoyance of the original image: a high variation difference between the original and the blurred image means that the original image was sharp whereas a slight variation difference means that the original image was already blurred.

Sharpness can be estimated from the activities of adjacent pixels or blocks of pixels. Tsomko and Kim [15] proposed a blur estimator that introduces the prediction residue as the difference between adjacent pixels. The proposed algorithm classifies the images into three different levels of quality : globally sharp, globally blurry, and average quality based on the variance of the prediction residue. Debing *et al.* [16] proposed a block-based blur detection for the blurring artifacts caused by the H.264/AVC compression. The blur index of the image is computed as the average of all the local blur values calculated at the boundaries of macro blocks. A weighting scheme based on video content is employed to reduce the influence from the texture.

Wee and Paramesran [17] argued that the eigenvalues of an image contain statistical information about that image: the most dominant eigenvalues contain image information while the least dominant ones contain noise information. Based on this, they proposed a method that estimates the image sharpness by using the dominant eigenvalues of the covariance matrix of the image. The final image sharpness index is determined by the trace of the first several eigenvalues. An approach similar to [17] has been used in [18] by Zhu and Milanfar. The authors quantified the amount of both blur and noise based on the singular value decomposition of the local image gradient matrix. They argue that the singular values are sensitive to sharpness be-

cause they reflect the strength of the gradients along the dominant direction and its perpendicular direction. Sharpness metric is then constructed by using the singular values of the local image gradient matrix.

2.3 Transformed-Based Methods

A number of sharpness/blurriness algorithms work in the spectral domain using transforms such as the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), and/or Discrete Wavelet Transform (DWT). For example, the algorithm developed by Marichal *et al.* [19] estimated sharpness based on the histogram of nonzero DCT coefficients among all 8×8 blocks of the transformed image. Kristan *et al.* [20] estimate image sharpness based on the Bayes entropy computed from DCT coefficients. In order to reduce computational complexity, the authors divided the image into 8×8 blocks and then used the DCT to estimate the spectra with the ignorance of the coefficients of an order higher than some predefined threshold. Next, sharpness values are calculated separately for each sub-image and their mean is taken as a measure of the overall image focus level.

Zhang *et al.* [21] argued that the sharpness of an image is closely related to the peakedness of its power spectral density. Therefore, they proposed a statistical measure using the bi-variate kurtosis. The DFT is computed first for the whole image, and the resultant matrix is treated as a two-dimensional probability density function that is used to compute the kurtosis. Sharpness is estimated based on the kurtosis of the power spectral density of the image.

Shaked and Tastl [3] proposed an image sharpness measure based on localized frequency analysis. The ideal high-pass and band pass filters are used to estimate energy in different frequency bands. Sharpness is then estimated based on the ratio of high-pass to low-pass energy computed from the spatial derivative of each line/column of the image.

In Ref. [22], Hassen *et al.* follow the idea that the phases of complex wavelet coefficients exhibit a consistent relationship across scales in the vicinity of sharp image features, such as edges and lines. The authors proposed a new measure to quantify the degree of local phase coherence (LPC) at each spatial location to build an LPC map. The overall sharpness index is then computed via a weighted sum of the values in the LPC map using weights determined by the rank order of these values.

Hsin *et al.* [23] argued that a high-frequency component preserving the full information of a given image. The authors developed a sharpness measure based on the output of a multi-channel filter bank, which is constructed from the Gaussian derivative wavelets. The output of the filter bank does not only contain the complete information of the input but also manifests prominent image features.

2.4 Hybrid Methods

In addition to the previous three categories, some of the sharpness/blurriness estimators are implemented via a hybrid approach, which employs a combination of edge-/pixel-based and transform-based methods [24–26]. Methods with hybrid approach have generally proven to perform better than edge-only based or transform-only based methods, though at the expense of added computational complexity.

Caviedes and Gurbuz [24] proposed a sharpness algorithm, which is a combination of the spatial domain edge profile acutance and the kurtosis of the frequency spectrum algorithms. The algorithm built a block-based sharpness estimator using the kurtosis of the DCT coefficients of each block; the overall sharpness estimate is given by the average of the sharpness values computed for edge profiles.

Chen and Bovik [25] proposed a blur metric that employs the statistics of the image gradient histogram and a wavelet-based detail map. First, a probabilistic support vector machine (SVM) is applied as a rough image quality evaluator. Then the detail image is used to refine the blur measurements. Finally, the blur information

is pooled to predict the blur quality of images.

Vu *et al.* [26] used a block-based approach to develop the first method specifically designed to measure local sharpness. Their method estimates the spatial and spectral sharpness of local image regions using the slope of the local magnitude spectrum [27] and the local total variation [28]; these values are then combined to generate an image sharpness map.

Chen *et al.* [29] proposed an algorithm that models the gradient of the given image as Markov chain and utilizes transition probabilities to compute a blurriness measure. The authors first compute the transition probabilities for selected pairs of gradient values and then combine these probabilities, using a pooling strategy, to formulate the blurriness measure.

2.5 Summary

This chapter presented a literature review of the previous sharpness/blurriness estimators that attempt to estimate the image sharpness both globally and locally. These methods can be roughly classified into four different categories: edge-based methods, pixel-based methods, transform-based methods, and hybrid methods according to their working mechanisms.

CHAPTER 3

A FAST WAVELET-BASED ALGORITHM FOR GLOBAL AND LOCAL IMAGE SHARPNESS ESTIMATION

3.1 Introduction

A useful goal in image processing is to determine whether one image (region) appears sharper than another does. Many applications find sharpness measure a crucial factor for tasks such as main-subject detection, image quality assessment, and image restoration. Previous methods of sharpness/blurriness estimation have employed a wide variety of approaches [see Chapter 2].

In this chapter, we present a sharpness estimator, called *FISH* (*F*ast *I*mage *SH*arpness), which offers the simplicity of a spectral-based method but with the improved predictive performance of a hybrid method. Following Refs. 3, 21, FISH operates under the assumption that perceived sharpness can be estimated by examining the energy in high-frequency bands. Here, we use a three-level separable discrete wavelet transform (DWT) and measure the log-energy of each DWT subband. Sharpness is estimated based on a weighted geometric mean of these log-energies. As we will demonstrate, despite its simplicity, FISH is competitive with the currently best-performing techniques. In addition, by clustering DWT coefficients, we show how FISH can be easily modified to yield a map indicating the relative sharpness of each image region. Thus, unlike most existing methods and similar to S_3 [26], FISH can generate maps of local image sharpness.

This chapter is organized as follows: In Section 3.2, we provide a brief introduction of the discrete wavelet transform. Section 3.3 provides details of the FISH algorithm.

Section 3.4 presents performance of FISH in terms of within-image and across-image sharpness estimation, and no-reference quality assessment of blurred images; this section also includes a discussion of run-time requirements. Brief summary is presented in Section 3.5.

3.2 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found to yield a fast computation of wavelet transform for which the wavelets are discretely sampled. In the DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales.

3.2.1 Multi-Resolution Analysis using Filter Banks

The DWT can be realized by the iteration of filters with rescaling. The resolution of the signal, which is a measure of the amount of detail information in the signal, is determined by the filtering operations, and the scale is determined by the up-sampling and down-sampling (subsampling) operations.

The DWT is computed by successive low-pass and high-pass filtering of the discrete time signal as illustrated in Figure 3.1. This is called the Mallat algorithm or Mallat-tree decomposition [30]. Its significance is in the manner it connects the continuous-time multi-resolution to discrete-time filters. Figure 3.1 shows the DWT decomposition of the $1D$ signal $X[n]$, where n is an integer. The low-pass filter is denoted by G_0 while the high-pass filter is denoted by H_0. At each level, the high-pass filter produces detail information, $d[n]$, while the low-pass filter associated with scaling function produces coarse approximations, $a[n]$.

At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in

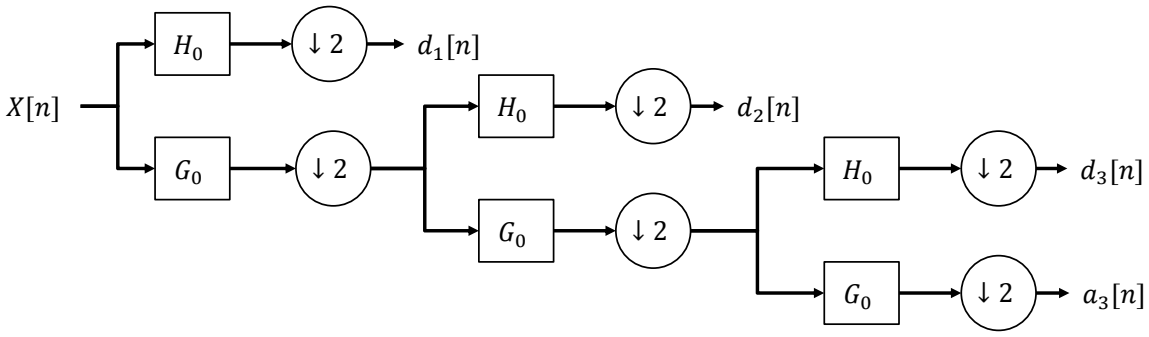


Figure 3.1: DWT decomposition, the $1D$ signal $X[n]$ is filtered by the low-pass filter G and the high-pass filter H. At each level, the high-pass filter produces detail information, $d[n]$, while the low-pass filter associated with scaling function produces coarse approximations, $a[n]$.

frequency is reduced by half. In accordance with Nyquist's rule if the original signal has a highest frequency of ω, which requires a sampling frequency of 2ω radians, then it now has a highest frequency of $\omega/2$ radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the low-pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale.

With this approach, the time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, $a[n]$ and $d[n]$, starting from the last level of decomposition.

Figure 3.2 shows the reconstruction of the original signal from the wavelet coefficients. The reconstruction is the reverse process of decomposition. The approximation and detail coefficients at every level are up-sampled by two, passed through the

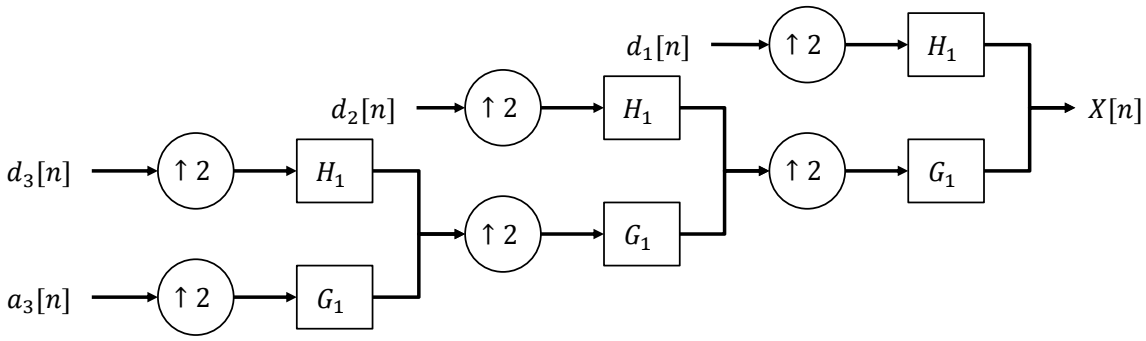


Figure 3.2: Original signal is reconstructed using IDWT from the wavelet coefficients. The reconstruction is the reverse process of the decomposition.

low pass and high pass synthesis filters and then added. This process is continued through the same number of levels as in the decomposition process to obtain the original signal. The Mallat algorithm works equally well if the analysis filters, G_0 and H_0, are exchanged with the synthesis filters, G_1 and H_1.

For the image, which is the 2D signal, the filtering is performed along rows and along columns of the image alternatively. Figure 3.3 shows the filtering operation applied to one level of the image. The coarse approximation at each level of the decomposition (LL) will be decomposed until the number of levels is reached.

3.2.2 Cohen-Daubechies-Feauveau wavelet

Cohen-Daubechies-Feauveau (CDF) wavelet is the historically first family of biorthogonal wavelets, which was made popular by Ingrid Daubechies. In these wavelets, the low-pass and the high-pass filters do not have the same length. The low-pass filter is always symmetric, while the high-pass filter could be either symmetric or asymmetric.

CDF wavelets are the most popular wavelets. They represent the foundations of wavelet signal processing and have numerous applications. The JPEG2000 compression standard uses the biorthogonal CDF 5/3 wavelet (also called the LeGall 5/3 wavelet) for lossless compression and a CDF 9/7 wavelet for lossy compression. The

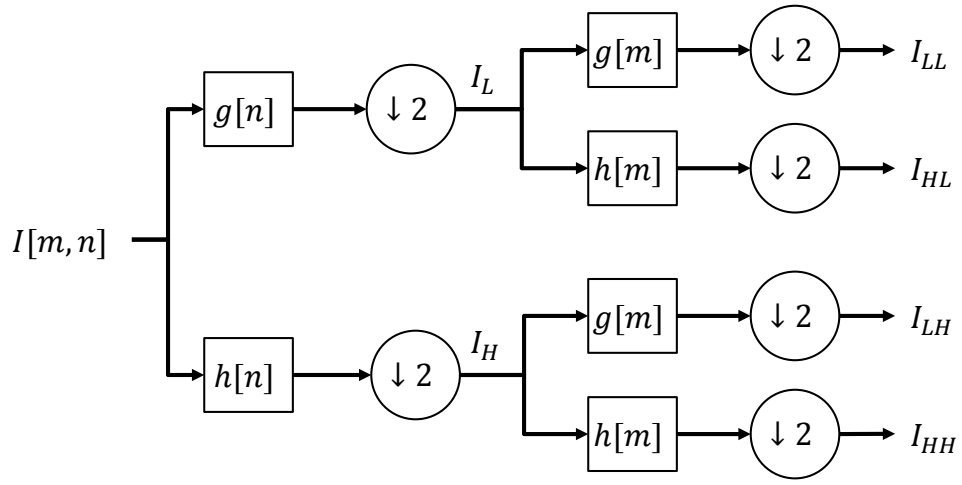


Figure 3.3: DWT decomposition applied to the image signal. The figure shows the filtering performed in one level of the decomposition.

notation 9/7 (or 5/3) represents the length of the discrete low-pass/high-pass filters used in the tree decomposition.

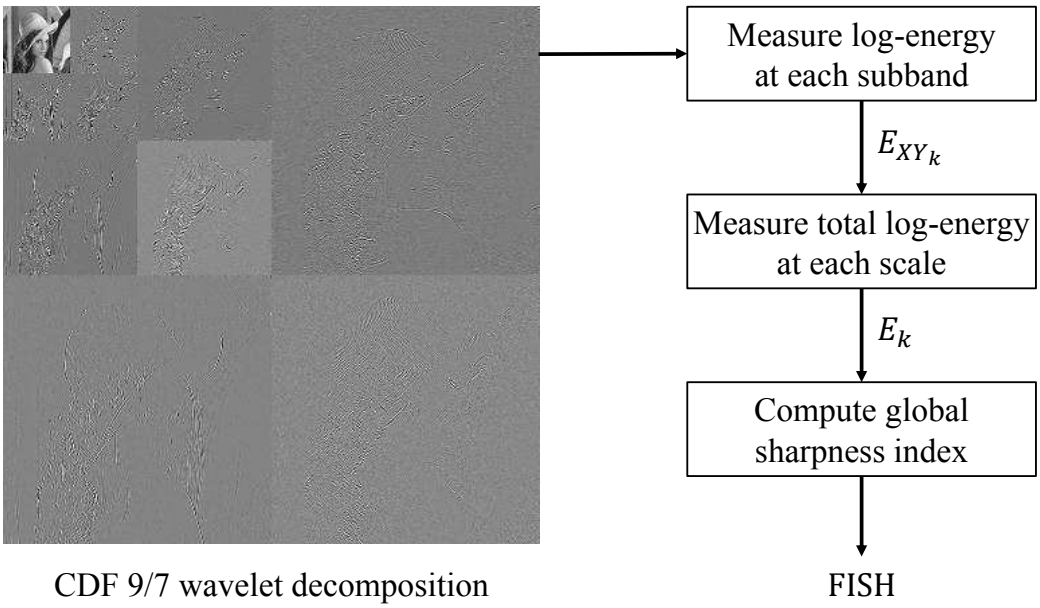
Table 3.1 shows the coefficients of the centered scaling and wavelet sequences of the filters used for CDF 9/7.

3.3 Fast Image SHarpness - FISH

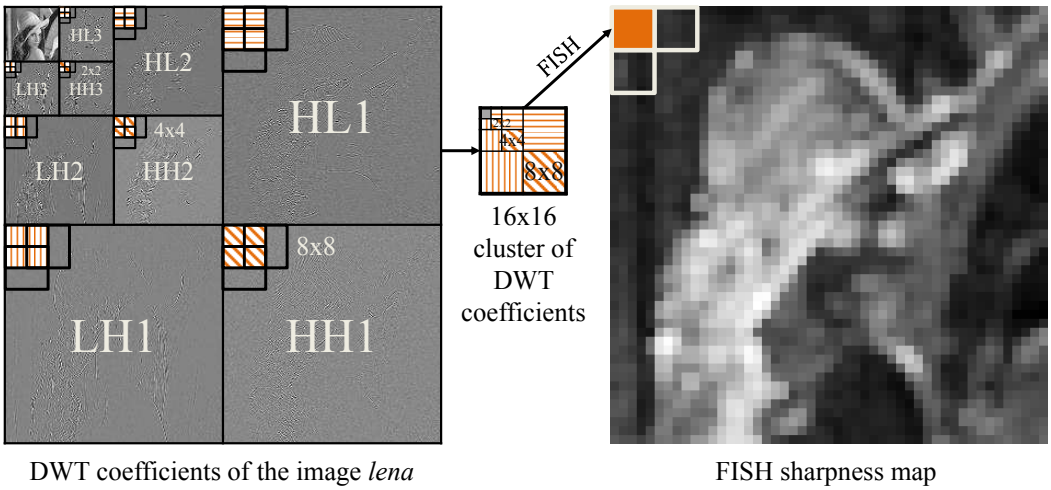
The FISH algorithm uses the assumption that for the sharpness regions in the image, most of its energy will be distributed in the high-frequency subbands of the wavelet decomposition. The lower the level of wavelet subband or the higher the frequency, the more energy is distributed in that level. Thus, the sharpness measure is estimated from the log-energy in the high-frequency subbands via wavelet coefficients.

3.3.1 Global Image-Based FISH

Figure 3.4(a) shows a block diagram of the FISH algorithm to compute the global sharpness index, which performs the following three steps for a given image I.



(a) Global sharpness estimation



(b) Clustering of wavelet coefficients to construct a local sharpness map

Figure 3.4: Illustration of global FISH algorithm (a) and DWT coefficients clustering into a wavelet block of size 16×16 to construct a local perceived sharpness map (b). The orange pixel and its two adjacent pixels in the sharpness map are shown according to the orange stripe set of DWT coefficients and two adjacent sets of DWT coefficients with 50% overlap. Note that, to promote visibility, the size of the blocks and sharpness map are not drawn to scale. (Figure from Ref. 1.)

Table 3.1: Coefficients of the filters used for CDF 9/7

	Analysis LPF	**Analysis HPF**	**Synthesis LPF**	**Synthesis HPF**
k	$1/2\ a_{dual}$	b_{dual}	a_{prim}	$1/2\ b_{prim}$
-4	0.026748757411	0	0	0.026748757411
-3	-0.016864118443	0.091271763114	-0.091271763114	0.016864118443
-2	-0.078223266529	-0.057543526229	-0.057543526229	-0.078223266529
-1	0.266864118443	-0.591271763114	0.591271763114	-0.266864118443
0	0.602949018236	1.11508705	1.11508705	0.602949018236
1	0.266864118443	-0.591271763114	0.591271763114	-0.266864118443
2	-0.078223266529	-0.057543526229	-0.057543526229	-0.078223266529
3	-0.016864118443	0.091271763114	-0.091271763114	0.016864118443
4	0.026748757411	0	0	0.026748757411

1. Decompose the input image into a three-level DWT using the lifting scheme of the CDF 9/7 filters [31].

2. Compute the log-energy of the DWT coefficients at each level.

3. Compute the overall sharpness index via a weighted sum of the per-level log-energy values.

The details of each step are described in the following subsections.

Step 1: Compute the DWT using CDF 9/7 filters

If the input image is the color image, then it will be converted to grayscale using the pixel-wise transformation as follows

$$gray\_img = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \tag{3.1}$$

where R, G and B stand for the red, green and blue channels of the color images.

The grayscale input image is then decomposed into wavelet subbands by using the CDF 9/7 filters [31] with three levels of decomposition. Let S_{LH_n}, S_{HL_n}, S_{HH_n} denote the LH, HL, and HH subbands at DWT level $n \in [1, 3]$. (The LL_3 subband, which contains the low-frequency components, is not used.)

Step 2: Compute the Log-Energy at each DWT Level

Images that appear sharp generally contain more high-frequency content than images which appear smooth/blurred. To quantify this effect, we first measure the log-energy of each subband at each decomposition level as follows:

$$E_{XY_k} = \log_{10}(1 + \frac{1}{N_k} \sum_{i,j} S^2_{XY_k}(i, j)), \tag{3.2}$$

where $XY \in \{LH, HL, HH\}$ is the wavelet subband. The quantity N_k, $k \in \{1, 2, 3\}$ is the number of DWT coefficients in the subband at level k. The addition of one is used to prevent negative values of E_{XY_k}.

Next, we measure the total log-energy at each decomposition level via

$$E_k = (1 - \alpha)\frac{E_{LH_k} + E_{HL_k}}{2} + \alpha E_{HH_k}, \tag{3.3}$$

where the parameter $\alpha = 0.8$ was chosen empirically to give greater weight to the energy in the HH subband; this band can be regarded to span a higher radial spatial frequency (by a factor of $\sqrt{2}$) than the LH and HL bands.

Step 3: Compute the Sharpness Index

Finally, the three per-level log-energy values E_1, E_2, and E_3 are combined as follows to determine a scalar sharpness index representing the image's overall sharpness:

$$FISH = \sum_{k=1}^{3} 2^{3-k} E_k. \tag{3.4}$$

Here, FISH ≥ 0, is the overall sharpness index; the larger the index, the greater the perceived sharpness. The factor $2^{3-k} = \{4, 2, 1\}$ when $k = \{1, 2, 3\}$ is used to provide greater weight to the finer scales (higher-frequency bands).

3.3.2 Local Block-Based FISH

The previous section described the FISH algorithm applied to the entire image. It is also possible to apply the algorithm in a block-based fashion to determine a map denoting local perceived sharpness.

To generate a local sharpness map, we compute a collection of local FISH values in a block-based fashion by using the clusters of DWT coefficients corresponding to each block b of size 16×16 in the image. As shown in Figure 3.4(b), each DWT subband is divided into small blocks of size 8×8, 4×4, and 2×2 for levels 1, 2, and 3, respectively with 50% overlap between neighboring blocks. These small blocks are assembled in clusters of size 16×16 to estimate local perceived sharpness.

Specifically, as seen in Figure 3.4(b), a cluster of size 16×16 is assembled by taking one 8×8 block from each level-1 subband, one 4×4 block from each level-2 subband, and one 2×2 block from each level-3 subband at the same spatial location. This cluster corresponds to a region of size 16×16 in the input image. Equations (3.2), (3.3), and (3.4) are then applied to this cluster of DWT coefficients to yield a FISH value that represents the perceived sharpness of the corresponding image region.

The FISH value is computed for every cluster of 16×16 DWT coefficients generated with 50% overlap between neighboring blocks of DWT coefficients in each subband, yielding a local sharpness map $FISH(x, y)$. Because we use 50% of overlap between neighboring blocks, each pixel in the sharpness maps corresponds to a block size of 8×8 in the input image. Figure 3.4 (*right*), illustrates the sharpness map computed by the FISH algorithm for the image *lena*.

It is also possible to collapse the sharpness map into a scalar sharpness index representing the image's overall sharpness. This index, $FISH_{bb}$, is computed by taking the root mean square of 1% largest values of local sharpness (FISH) indices

(following from [26]):

$$FISH_{bb} = \sqrt{\frac{1}{T}\sum_{i=1}^{T} FISH_i^2}, \tag{3.5}$$

where T denotes the number of blocks which received the 1% largest FISH indices of the sharpness map; and where $FISH_i$, $i = 1, 2, .., T$ denotes the FISH indices of these blocks. The value of 1% is used because, as argued in [26], the overall perceived sharpness of an image is largely determined by the image's sharpest regions.

Examples of the local perceived sharpness maps computed from two different images are shown in Figure 3.5. Located in Figure 3.5(a) are the image *lena* and its corresponding FISH sharpness map. The sharp regions of the image corresponding to the subject's hair are well captured by the FISH sharpness map and are represented by the brighter regions in the map. In the other hand, the smooth regions of the image corresponding to the subject's shoulder, which are the less sharp regions in the image, are represented by the darker regions in the sharpness map.

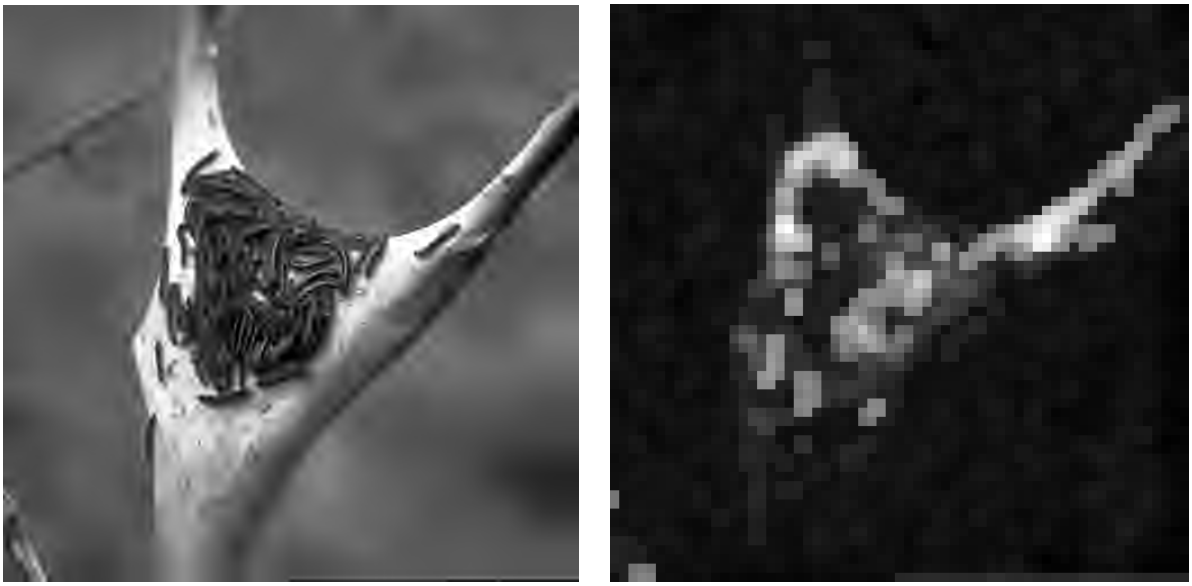
In Figure 3.5(b), the image *swarm*, compressed by the JPEG2000 compression standard, and its corresponding FISH sharpness map are shown. The blur regions corresponding to the background of the image are well captured by the FISH sharpness map and are represented by the darker regions in the map. The regions contain the swarm are sharper than the background and are represented by the brighter regions in the sharpness map.

3.4 Results and Discussion

3.4.1 Representative Results

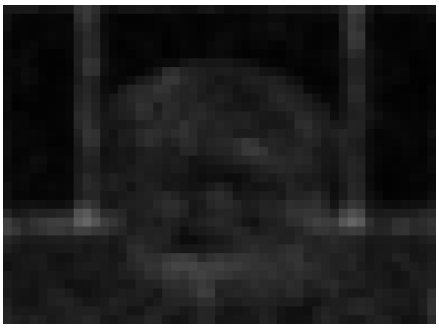
Figures 3.6 and 3.7 show representative results that demonstrate the ability of FISH/FISH$_{bb}$ to accurately estimate across-image and within-image sharpness (FISH$_{bb}$ only) for a variety of images containing different sharpness levels. The images are ordered based on subjective ratings of sharpness [26].

(a) Image *lena* (left) and its local perceived blurring map (right)

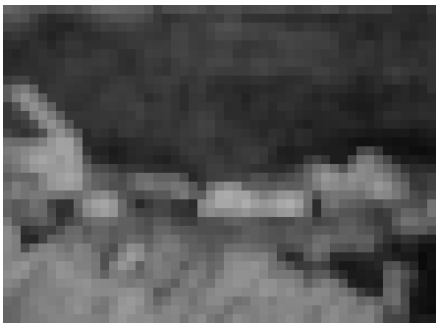


(b) Image *swarm* (left) and its local perceived blurring map (right)

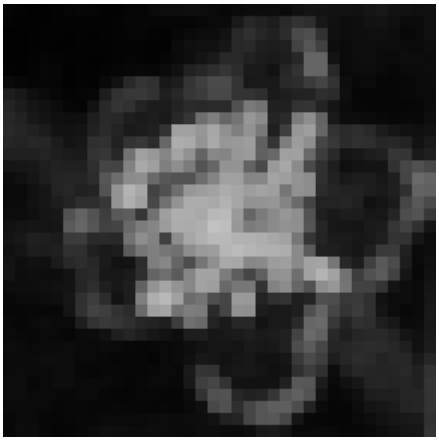
Figure 3.5: Examples of different images and their local perceived sharpness maps generated by the FISH algorithm, the maps are normalized to promote visibility. The brighter the region in the maps is, the greater the perceived sharpness is at the corresponding spatial regions of the images.



(1) *ball* 4.92 / 8.35

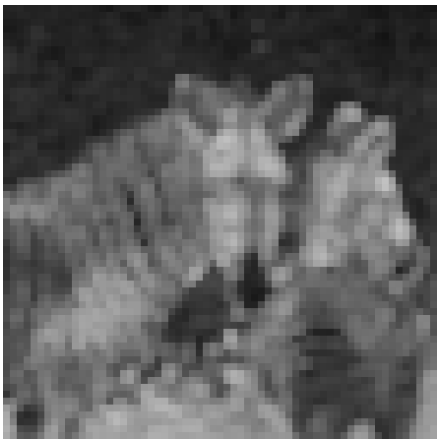


(2) *airplane* 11.91 / 17.33

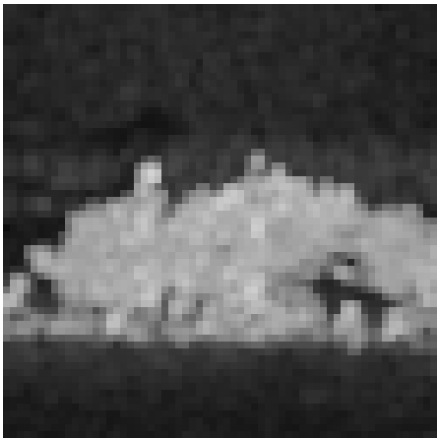


(3) *petal* 11.35 / 19.38

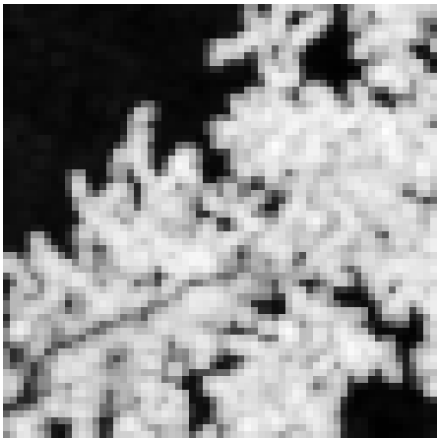
Figure 3.6: Representative maps generated by using FISH_{bb} along with sharpness indices computed via $\text{FISH}/\text{FISH}_{bb}$. The images were selected and organized in the order of overall sharpness judged by human subjects.

(4) *zebra* 13.20 / 18.86

(5) *pelicans* 14.00 / 20.59

(6) *branches* 21.20 / 25.70

Figure 3.7: Representative maps generated by using FISH_{bb} along with sharpness indices computed via $\text{FISH}/\text{FISH}_{bb}$. The images were selected and organized in the order of overall sharpness judged by human subjects.

In terms of across-image sharpness, the FISH/FISH$_{bb}$ indices generally match the relative perceived sharpness across these images. For example, two of the images, *petal* and *zebra*, are not as sharp as images *pelicans* and *branches*, but are clearly much sharper than image *ball*. FISH fails to predict the sharpness of image *petal* in comparison to image *airplane* while FISH$_{bb}$ fails to predict the sharpness of image *petal* in comparison to image *zebra*. We believe that these failure cases are attributable to the fact neither FISH nor FISH$_{bb}$ take into account local contrast. Such a measurement could be implemented, though at the expense of added complexity. However, as we will demonstrate shortly, both FISH and FISH$_{bb}$ perform very well overall in terms of across-image quality assessment of blurred images.

In terms of within-image sharpness, FISH$_{bb}$ correctly estimates the perceived sharpness of each image region. For example, in image *petal*, the flower's stamens and the edges of the petals are the sharpest regions in this image; these regions are accurately highlighted in the corresponding FISH$_{bb}$ map. Similarly, in image *branches*, the branches are much sharper than the smooth sky regions; this fact is well reflected in the corresponding FISH$_{bb}$ map.

3.4.2 Performance Evaluation Criteria

Before evaluating performance of each algorithm on each subset of JPEG2000-compressed images, we applied a four-parameter logistic transform to the raw predicted scores, recommended by VQEG [32]. The four-parameter logistic transform is given by:

$$f(x) = \frac{\tau_1 - \tau_2}{1 + \exp(-\frac{x-\tau_3}{|\tau_4|})} + \tau_2, \tag{3.6}$$

where x denotes the raw predicted score and τ_1, τ_2, τ_3, and τ_4 are free parameters selected to provide the best fit of the predicted scores to the subjective rating scores.

Following VQEQ recommendations in Ref. 32, we employ the Spearman Rank-Order Correlation Coefficient (SROCC) to measure prediction monotonicity, and employ the Pearson Linear Correlation Coefficient (CC) and the Root Mean Square

Error (RMSE) to measure prediction accuracy. The prediction consistency of each algorithm is measured by two additional criteria: the outlier ratio (OR [33]) and the outlier distance (OD [34]). The outlier ratio (OR) is the ratio of number of *false* scores predicted by the algorithm to the total number of scores. A *false* score is defined as the transformed score lying outside the 95% confidence interval of the associated subjective score [33]. In addition, the outlier distance (OD) is informative about how far the outlier falls outside of the confidence interval. The OD is measured by the total distance from all outliers to their closest edge points of the corresponding confidence interval [34].

3.4.3 No-Reference Quality Assessment of Blurred Images

To evaluate the performance of FISH in predicting quality of blurred images, we used the subsets of blurred images from four image-quality databases. They are the LIVE database [35] (containing 145 blurred images), the IVC database [36] (20 blurred images), the TID database [37] (96 blurred images), and the CSIQ database [34] (150 blurred images).

We compared our method against five sharpness estimators (ST [3], JNBM [11], CPBD [12], LPCM [22], and S_3 [26]), two blurriness estimators (MMZ [19] and MDWE [5]). These algorithms are selected because they have high reported performance and the implementation code is public available.

Table 3.2 shows the results of this evaluation. Both FISH and FISH_{bb} perform quite well on all four databases. In terms of CC and SROCC, FISH_{bb} outperforms other methods on the two largest databases (CSIQ and LIVE) and is competitive on the other two databases; FISH and MMZ are the two best methods on IVC. In terms of outlier analysis indicated by the OR and OD indices, FISH_{bb} also shows the best performance comparing to the other algorithms.

Table 3.2: Performances of various algorithms on no-reference quality assessment of blurred images; the two best results are highlighted; the last column shows the average weighted by number of images in each database.

	LIVE	**IVC**	**TID**	**CSIQ**	**Avg.**	**LIVE**	**CSIQ**
	SROCC					**OR**	
JNBM	0.787	0.666	0.714	0.762	0.755	68.97%	36.67%
CPBD	0.919	0.769	**0.854**	0.885	0.884	62.76%	37.33%
ST	0.702	0.406	0.516	0.705	0.645	76.55%	42.67%
MMZ	0.860	**0.971**	0.732	0.860	0.835	66.90%	31.33%
MDWE	0.804	0.685	0.717	0.770	0.765	64.83%	34.00%
LPCM	0.928	0.835	0.803	0.905	0.886	58.62%	31.33%
S_3	**0.944**	0.869	0.850	**0.906**	**0.904**	**53.10%**	32.67%
FISH	0.881	**0.932**	0.786	0.894	0.866	64.83%	**26.00%**
FISH$_{bb}$	**0.938**	0.919	0.841	**0.917**	**0.907**	**54.48%**	**24.00%**
	CC					**OD**	
JNBM	0.816	0.698	0.727	0.806	0.786	710.33	6.2824
CPBD	0.895	0.801	0.848	0.882	0.875	441.70	4.2688
ST	0.704	0.603	0.621	0.690	0.674	956.76	10.3675
MMZ	0.885	**0.956**	0.753	0.889	0.859	529.76	3.6740
MDWE	0.806	0.711	0.709	0.797	0.775	703.94	6.3801
LPCM	0.917	0.949	0.811	0.911	0.892	411.010	3.093
S_3	**0.943**	0.928	**0.877**	0.911	**0.914**	**285.97**	3.0307
FISH	0.904	**0.957**	0.816	**0.923**	0.893	470.83	**2.4168**
FISH$_{bb}$	**0.944**	0.941	**0.858**	**0.943**	**0.923**	**289.85**	**1.6476**

3.4.4 Local Sharpness Estimation

We compared the sharpness maps estimated by FISH_{bb} with ground-truth sharpness maps obtained from human subjects as reported in Ref. 26. In Figures 3.8, 3.9, and 3.10, we show the original images and the corresponding sharpness maps obtained from human subjects and estimated by S_3 and FISH_{bb}.

The S_3 algorithm was specifically designed to generate sharpness maps and as reported in Ref. [26], S_3 generally yields the best map predictions comparing to the other algorithms. As seen in Figures 3.8, 3.9, and 3.10, FISH_{bb} can yield maps that are well-correlated with ground-truth maps and quite competitive with S_3's maps. This latter assertion is quantified in Table 3.3, which shows the SROCC, CC (after non-linear regression), and RMSE between the ground-truth sharpness maps and the maps predicted via S_3 and FISH_{bb} for all six different images.

3.4.5 Monotonic Prediction of Blur Parameter

One criterion that a good blur quality metric has to meet is that it should predict well the blur parameters of the images. In case of the Gaussian blur images, the method should be able to predict the variance of the Gaussian filter. We use 29 reference images from the LIVE image database to generate a set of blur images using 15×15 Gaussian filter with different variances selected from $\{0.4, 0.8, 1.6, 2.0, 2.4, 2.8\}$. Each reference image has seven (7) blurred versions. Figure 3.11 shows the performance of some blur quality methods in predicting the blur parameters. All the quality indices are normalized to the range $[0, 1]$ for better comparison. We can see that our method is quite good in predicting the Spearman rank order of these parameters, the blur index is not widely spread at the same level of Gaussian filter variance.

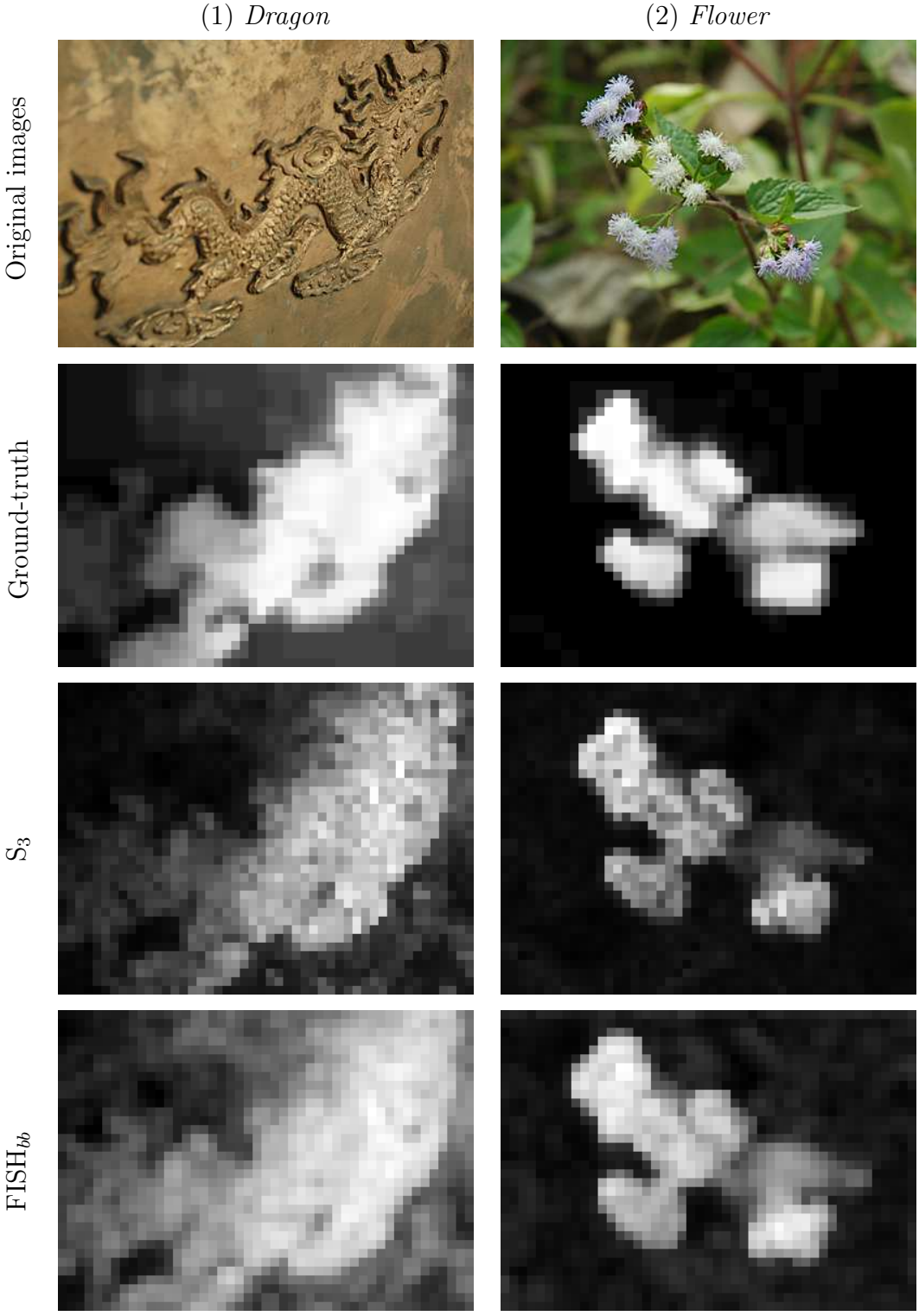


Figure 3.8: Local sharpness maps of the two images *Dragon* and *Flower* obtained from human observers, S_3 and FISH algorithms.

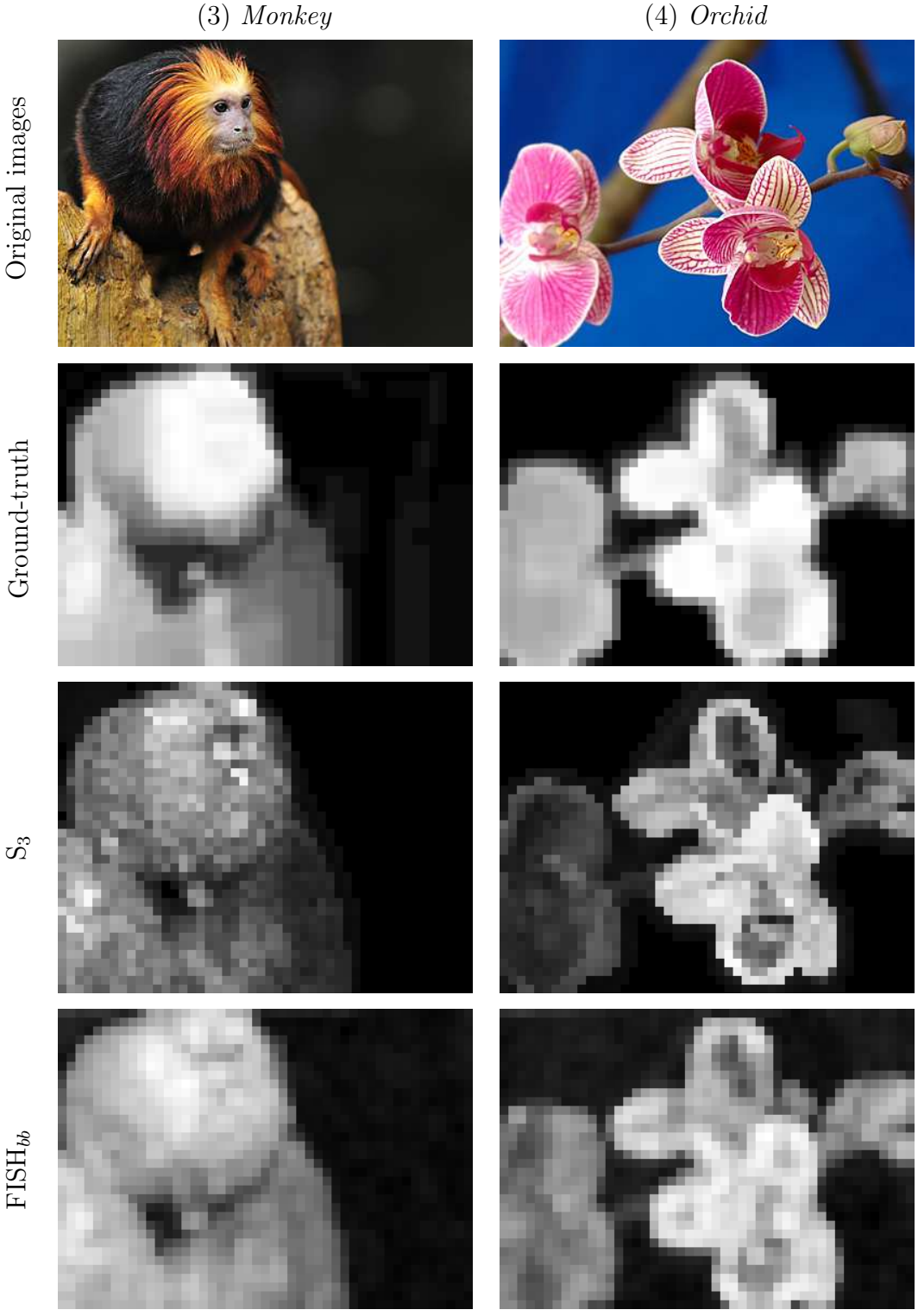


Figure 3.9: Local sharpness maps of the two images *Monkey* and *Orchid* obtained from human observers, S_3 and FISH algorithms.

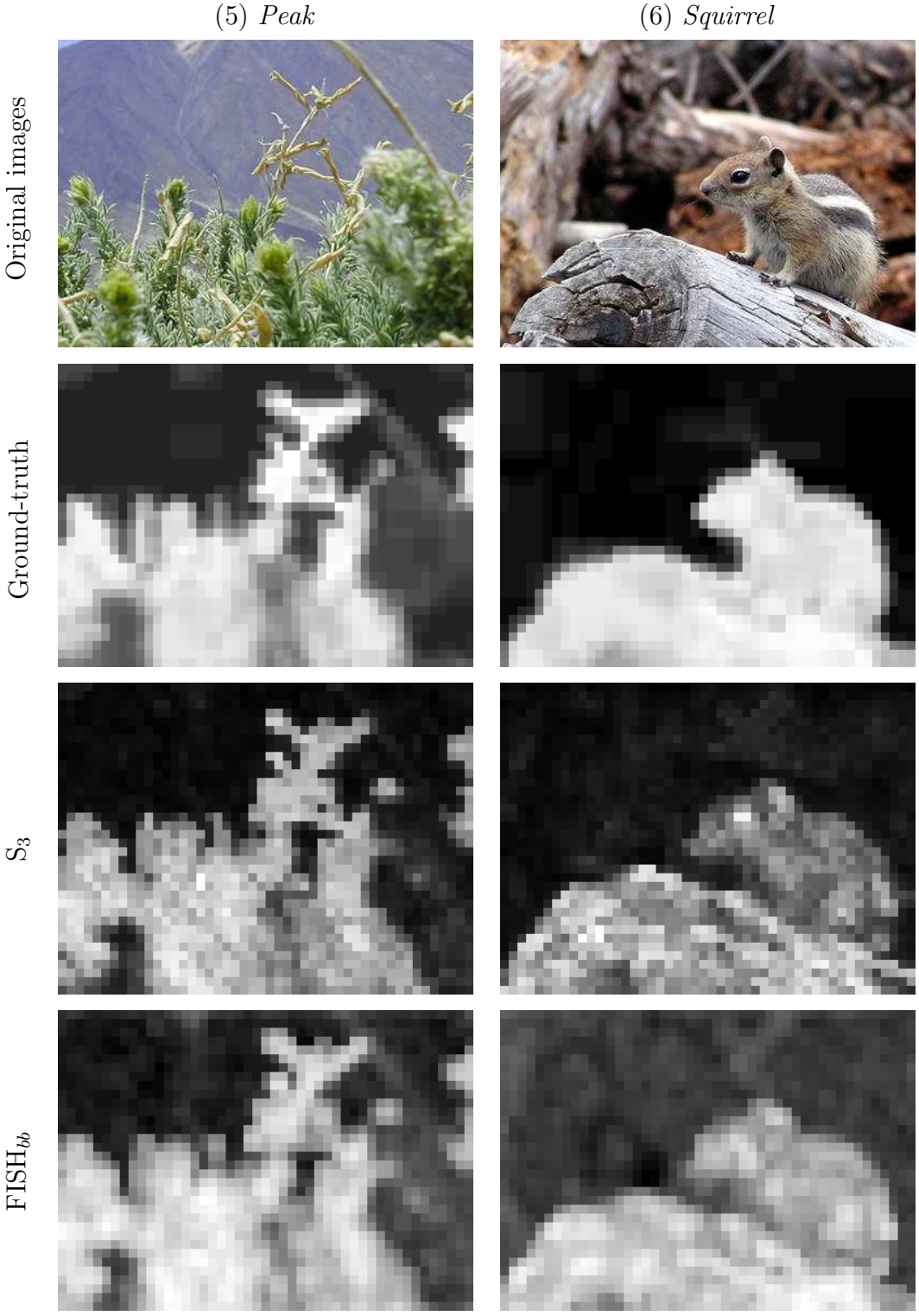


Figure 3.10: Local sharpness maps of the two images *Peak* and *Squirrel* obtained from human observers, S_3 and FISH algorithms.

Table 3.3: Local Sharpness Prediction Accuracy

	Image	JNB	CPBD	ST	MMZ	MARZ	S_3	$FISH_{bb}$
SROCC	*dragon*	0.589	0.646	0.866	0.892	0.606	**0.931**	0.923
	flower	0.706	0.756	0.643	0.547	0.631	0.712	**0.749**
	monkey	0.313	0.506	0.529	0.885	0.049	**0.916**	0.897
	orchid	0.439	0.796	0.759	0.841	0.409	**0.920**	0.910
	peak	0.202	0.052	0.860	0.881	0.490	0.901	**0.912**
	squirrel	0.835	0.829	0.704	0.801	0.684	0.794	**0.854**
	Average	0.514	0.598	0.727	0.808	0.478	0.862	**0.874**
CC	*dragon*	0.629	0.663	0.834	0.862	0.611	0.947	**0.950**
	flower	0.665	0.858	0.696	0.873	0.761	**0.936**	0.927
	monkey	0.428	0.495	0.517	0.940	0.276	0.944	**0.959**
	orchid	0.500	0.831	0.814	0.868	0.097	0.914	**0.929**
	peak	0.139	0.273	0.875	0.881	0.575	**0.928**	0.927
	squirrel	0.849	0.935	0.905	0.942	0.852	0.938	**0.941**
	Average	0.535	0.676	0.773	0.894	0.529	0.938	**0.941**
RMSE	*dragon*	0.629	0.663	0.834	0.862	0.611	26.414	**26.016**
	flower	0.665	0.858	0.696	0.873	0.761	**28.805**	31.191
	monkey	0.428	0.495	0.517	0.940	0.276	29.801	**25.906**
	orchid	0.500	0.831	0.814	0.868	0.097	39.671	**36.221**
	peak	0.139	0.273	0.875	0.881	0.575	**29.461**	29.662
	squirrel	0.849	0.935	0.905	0.942	0.852	**29.683**	31.256
	Average	0.535	0.676	0.773	0.894	0.529	30.639	**30.042**

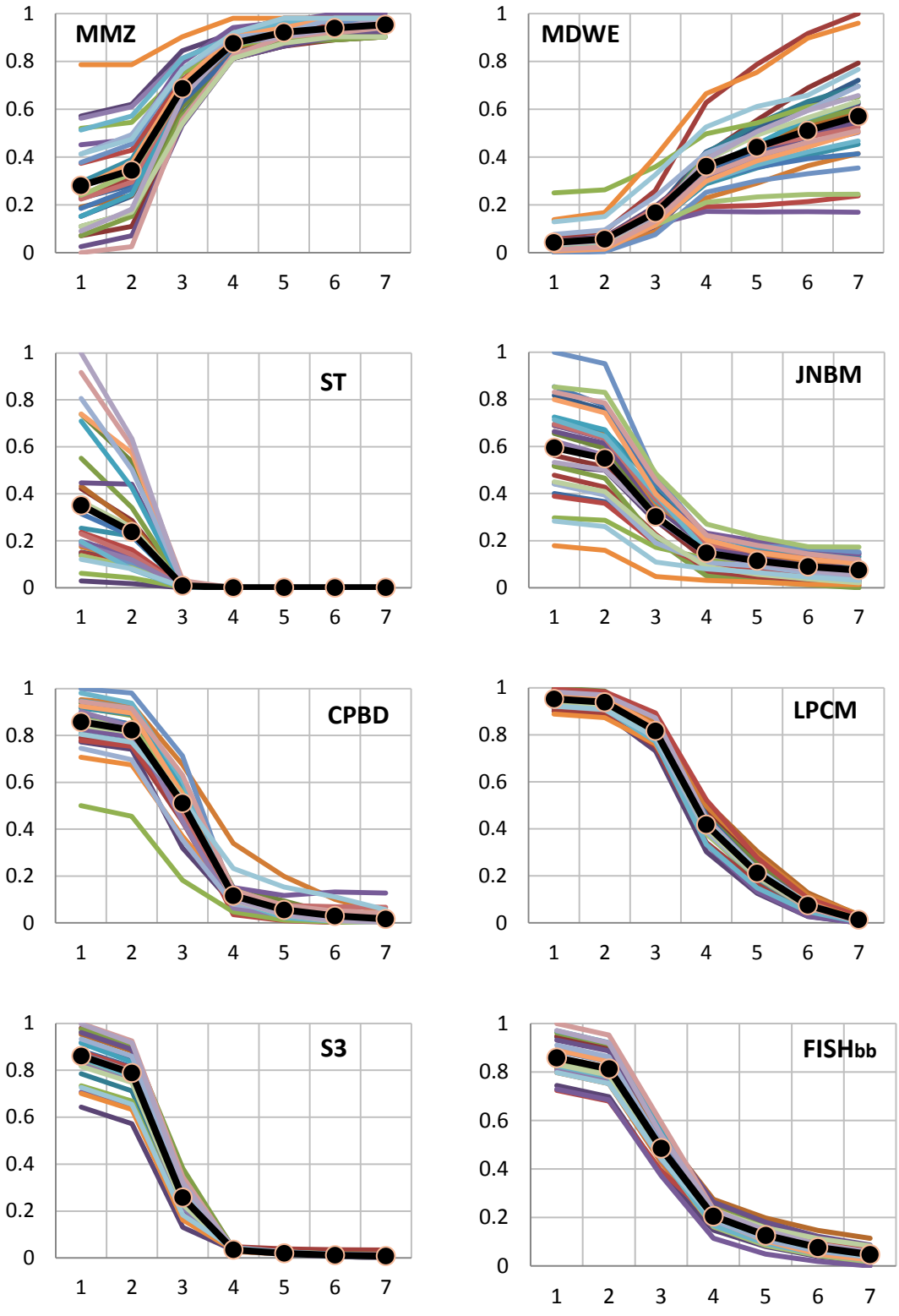


Figure 3.11: Performance of blur parameter prediction. All the quality indices are normalized to the range $[0, 1]$ for better comparison. The bolded lines represent the arithmetic mean of all the response lines.

Table 3.4: Run-time requirements (seconds/image) tested on different image sizes; the two fastest methods are highlighted.

	512 × 512	**1024 × 768**	**1280 × 960**	**1600 × 1200**
JNBM	1.854	5.563	8.779	14.812
CPBD	2.162	7.364	12.306	22.927
ST	**0.210**	**1.041**	**1.991**	**4.086**
MMZ	0.608	1.859	2.985	5.024
MDWE	0.914	3.712	7.225	14.273
LPCM	0.909	2.852	4.151	6.993
$\mathbf{S_3}$	29.154	64.522	122.640	142.841
FISH	**0.079**	**0.259**	**0.469**	**0.611**
$\mathbf{FISH}_{bb}$	1.309	4.018	6.291	10.126

3.4.6 Run-time vs. Image Size

To evaluate run-time performance, we applied FISH, FISH_{bb}, and other sharpness/blurriness estimators to images of size 512×512, 1024×768, 1280×960, and 1600×1200 pixels. Table 3.4 shows the average run-time of each algorithm in seconds, where the average was taken over 100 trials. This test was performed using a modern desktop computer (Intel Quad Core at 2.66 GHz, 12 GB RAM DDR2 at 6400 MHz, Windows 7 Pro 64-bit, Matlab 7.8). All of the methods were implemented in Matlab.

As shown in Table 3.4, FISH is the fastest algorithm for all image sizes, and FISH_{bb}, which is a block-based algorithm, is still significantly faster than the methods that yield competitive predictive performance (S_3, JNBM, CPBD; see Table 3.2). In terms of memory requirements, both FISH and FISH_{bb} have the same memory requirements as a standard DWT with only a negligible amount of additional memory needed for the output map (for FISH_{bb}) and other scalar variables.

3.5 Summary

This chapter presented a simple, yet effective algorithm (FISH) for estimating both global and local image sharpness. By decomposing the input image via a three-level separable DWT, FISH estimates sharpness based on a weighted geometric mean of the DWT subband energies. FISH can also operate in a block-based fashion (FISH_{bb}) by applying the same computation to groups of DWT coefficients to generate a local sharpness map. The proposed algorithm shows good performance in predicting the local sharpness as well as the quality of the blurred images. We also demonstrated the efficacy of FISH/FISH_{bb} on predicting the level of blur in the Gaussian blur images.

CHAPTER 4

APPLICATION OF SHARPNESS MEASURE TO JPEG2000 NO-REFERENCE IMAGE QUALITY ASSESSMENT

4.1 Introduction

The JPEG2000 compression standard has become popular in the field of image compression due to its high coding performance. Low-rate JPEG2000-compressed images exhibit blurring and ringing artifacts caused by the attenuation and quantization of the high-frequency components. Several algorithms have been developed to predict the quality of JPEG2000-compressed images by quantifying these artifacts in a manner that agrees with the visual quality judged by human subjects. The general approach involves either estimating the amount of blurring artifacts or estimating the perceived ringing artifacts, or a combination of the two estimations.

It is known that the destruction of sharp regions due to JPEG2000 encoding reduces visual quality, resulting in the blurring artifacts. Therefore, it is straightforward to apply a sharpness/blurriness estimator for JPEG2000-compressed images to estimate perceived blurring artifacts appear in such images. Many algorithms have been developed to estimate quality of JPEG2000-compressed images by estimating the amount of blurriness/sharpness in the image as discussed in Chapter 2. The application of the blurriness/sharpness algorithms to predict the quality of blurred images has shown competitive performance and promising results [38].

Nevertheless, the blurring artifacts in the JPEG2000-compressed images are different from the blurring artifacts in the blurred images, which are nearly uniform across the images. As shown in the first and second rows of Figure 4.1, besides the

blurring artifacts at some regions, a heavily distorted JPEG2000-compressed image still contains very sharp regions (the log and branches in the *logseaside* image (top row), the grass area in the *foxy* image (middle row)). These regions are well-captured by the respective FISH sharpness maps in the second column of Figure 4.1.

For comparison, the image *swarm* and its FISH sharpness map are shown in the third row of Figure 4.1. Due to the lack of high-frequency regions, this image is not as sharp as the other two images, but human subjects rate the image *swarm* slightly better than image *log_seaside* and image *foxy*. This is because of the the annoying high-frequency artifacts that are present in the latter images. Therefore, a sharpness estimator, which often estimates only the effect of the blurring artifacts in the image, is not adequate to predict quality of the JPEG2000-compressed images accurately.

One type of high-frequency artifact/distortion that appears in the JPEG2000-compressed image is the ringing artifact; these artifacts in the JPEG2000-compressed images make their containing regions sharper than the surrounding regions. The ringing artifacts are caused by the quantization of the high frequency coefficients resulting from a coding based on a wavelet transform; they often appear around the strong edges and locally produce haloes and/or rings in the image and make the image look sharper but worse in quality. Therefore, it is necessary to account for the effect of the ringing artifacts in the image in order to estimate quality of the JPEG2000-compressed image.

Tong *et al.* [39] presented a no-reference method for JPEG2000 images based on a principal components analysis (PCA). First, by dividing all edge points in JPEG2000 images into distorted and undistorted, local features are extracted at each of the detected edge points to indicate blurring and ringing. A model is then employed to map these local features to local distortion estimates through the probabilities of the edge points being distorted or undistorted. Quality is estimated based on the local distortion estimates.

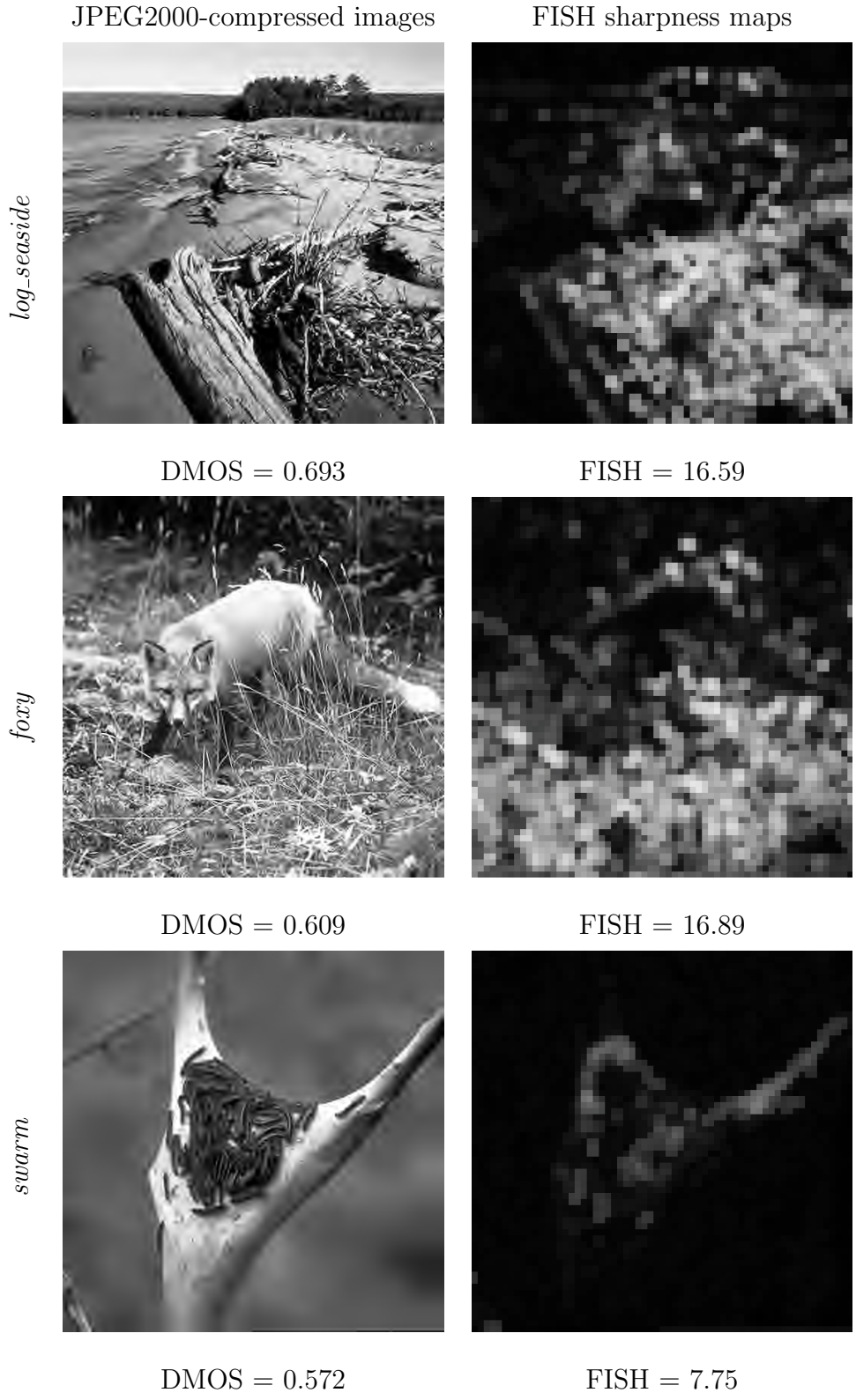


Figure 4.1: Examples of two heavily distorted JPEG2000-compressed images (*foxy* and *log_seaside*) but contain very sharp regions [(a) and (b)]. The image *swarm* and its sharpness map (c) is included for comparison.

The method proposed by Barland *et al.* in Ref. 40 involves two stages that estimate blurring and ringing artifact separately. The perceived blurring artifacts are estimated by the increase of using the spatial information and pixel activity in the whole image. The perceived ringing artifacts are estimated in the ringing regions, which are supposed to appear around strong edges. These two indices and the percentage of strong edges in the image are pooled to achieve the final quality index of the compressed image.

In Ref. 41, Liang *et al.* presented an algorithm that employs the statistical information on image gradient profiles. The algorithm measures blur artifacts by using a gradient profile sharpness histogram and an existing just noticeable difference (JND) threshold model [42]. The ringing artifacts are estimated by applying the perceived ringing metric proposed in Ref. 43 to the regions associated with the gradient profiles. A weighted Minkowski summation is employed to combine the perceived blurring and ringing estimates into the quality estimate of the image.

Instead of an attempt to quantify the predefined artifacts (blurring and ringing) in predicting the quality of a JPEG2000-compressed image, some researchers extract features from the pixels/regions in the image and study the changes of these features with respect to image quality. Sheikh *et al.* [44] presented a model which uses the statistics of natural scene features as reference information to assess quality of the distorted image. The authors reported that when JPEG2000 images are decomposed via a wavelet transform, the subband probabilities indicate the loss of visual quality. Image quality is estimated by computing features based on these probabilities from all wavelet subbands and by applying a nonlinear combination of the features.

The quality of the JPEG2000-compressed images can also be estimated by modeling the activities of pixels in the image. Sazzad *et al.* in Ref. 45 estimated image quality based on pixel distortions and edge information. Visual artifacts in the image cause pixel distortions, and these distortions are estimated by using local statistics of

partially overlapping 5×5 blocks. In addition, based on the assumption that human visual perception is very sensitive to edge information in images, two features of the edge information, zero-crossings rate and the histogram of vertical and horizontal histogram edge features, are extracted from the input image. These features are then combined via a training step to yield the final quality index of the image.

In Ref. 46, Zhang *et al.* proposed an algorithm that introduces a basic activity map of general pixels. An arbitrary pixel in the image can be classified as monotone-changing, zero-crossings, or inactive pixel depends on its activity with the adjacent pixels. The activity map is then weighted by image's structural content and pooled to yield an estimate of image quality.

Zhang *et al.* [47] proposed a method to estimate image quality based on either $1D$ or $2D$ kurtosis in the discrete cosine transform domain of general image blocks. The proposed measures were argued to be advantageous in terms of their parameter-free operation and their computational efficiency (they do not require edge/feature extraction). The image quality is estimated as the median of the kurtosis values computed for all blocks in the image.

Another approach to estimate quality of JPEG2000-compressed images is to apply a general no-reference image quality assessment. Modern no-reference image quality assessment methods such as BLIINDS-II [48], DIIVINE [49], BRISQUE [50] have been shown to have competitive performance on the subsets of JPEG2000-compressed image in specific image-quality databases. These methods often extract quality features from the images and apply a machine learning mechanism to train the model. Depending on the number of features, the method can take a significant amount of time to estimate image quality.

Specifically, in Ref. 48, Saad *et al.* presented the BLIINDS-II algorithms which estimate quality based on a generalized statistical model of local DCT coefficients. BLIINDS-II operates on each 17×17 image patch and extracts DCT-based contrast

and DCT-based structural features. DCT-based contrast is defined as the average of the ratio of the non-DC DCT coefficient magnitudes in the local patch normalized by the DC coefficient of that patch. The DCT-based structure is defined based on the kurtosis and anisotropy of each DCT patch. These features are combined to form the quality estimate.

In Ref. 49, Moorthy and Bovik presented the DIIVINE algorithm, which employs a steerable pyramid transform with two scales and six orientations. The features extracted in DIIVINE are based on statistical properties of the subband coefficients. A total of 88 features are extracted and used to estimate quality via the same two-stage classification/regression framework.

In Ref. 50, Mittal *et al.* presented the BRISQUE algorithm, a fast no-reference IQA algorithm which employs statistics measured in the spatial domain. BRISQUE operates on two image scales; for each scale, 18 statistical features are extracted. The 36 features are used to perform distortion identification and quality assessment via the aforementioned two-stage classification/regression framework.

Following the paradigm of estimating JPEG2000-compressed images based on the perceived blurring and ringing artifacts separately as in Refs. 40, 51, in this chapter, we present a no-reference quality assessment algorithm for JPEG2000-compressed images. The algorithm is named EDIQ (*ED*ge-based *I*mage *Q*uality) because it is based on the perceived artifacts in the edge/near-edge regions of the images. A part of this method has been presented at the SPIE conference [52]. The main assumption used by the EDIQ algorithm is that the quality of JPEG2000-compressed images can be estimated via the local perceived sharpness/blurriness and the local perceived ringing artifacts of the edge/near-edge regions. Local perceived blurring artifact is estimated via the FISH algorithm, and local perceived ringing artifact is estimated by the local variance of the Laplacian image. These local values are combined and collapsed into a quality estimate of the input JPEG2000-compressed image.

This chapter is organized as follows: In Section 4.2, we provide details of the EDIQ algorithm. Section 4.3 presents results of EDIQ on subsets of JPEG2000-compressed images from various image-quality databases. Summary and limitations are presented in Section 4.4.

4.2 Algorithm

The EDIQ algorithm estimates quality of a JPEG2000-compressed image by estimating perceived blurring artifact and perceived ringing artifact in the edge/near-edge regions of that image. Specifically, given an input image, EDIQ performs the following steps (shown in Figure 4.2) to estimate image quality:

1. Estimate local perceived blurring artifact by using FISH sharpness estimator.
2. Estimate local perceived ringing artifacts.
3. Determine edge/near-edge regions by using the Canny edge detection.
4. Combine and collapse maps to estimate image quality.

4.2.1 Estimate local blurring artifact

The blurring artifact in JPEG2000-compressed images is due to the attenuation of the high-frequency components in the image's frequency spectral. The more blur the image, the worse the quality of the image. As stated in Chapter 2, the blurriness and the sharpness can be used as antonym, a sharpness estimator can be used to estimate the amount of blurring in the image. Moreover, the FISH algorithm has shown competitive performance on predicting the quality of the blurred images; therefore, we employ the FISH sharpness estimator as presented in chapter 3, to examine local perceived blurring artifacts.

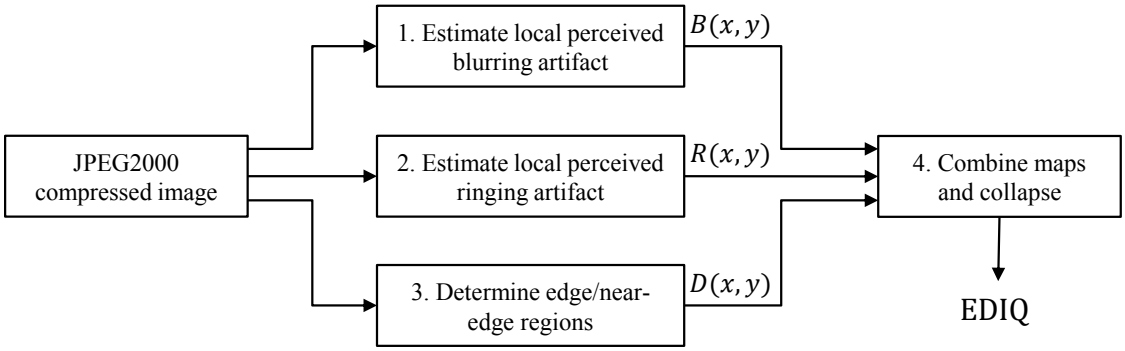


Figure 4.2: Block diagram of the EDIQ algorithm. The input JPEG2000-compressed image is used to generate a local perceived blurring map $B(x, y)$, a local perceived ringing map $R(x, y)$, and an edge/near-edge regions masking map $D(x, y)$. These maps are combined and collapsed into a single scalar that represents the perceived quality of the input image.

As described in chapter 3, the FISH algorithm is applied to every cluster of 16×16 DWT coefficients constructed with 50% overlap between neighboring blocks of DWT coefficients in each subband, yielding a local FISH sharpness map. The local perceived blurring artifact at location (x, y) of the input image, denoted by $B(x, y)$, is computed as the inverse of the local sharpness index at that location. In details, the blurring artifact at the location (x, y) is given by

$$B(x, y) = \frac{1}{FISH(x, y)}. \tag{4.1}$$

The greater the $B(x, y)$ value, the greater the perceived blurring artifact at the corresponding spatial region.

4.2.2 Estimate local perceived ringing artifact

As observed in Section 4.1, beside the blurring artifacts, the JPEG2000-compressed images also contain the ringing artifacts, which cause the oscillation near sharp regions. To estimate the perceived ringing artifacts, first we determine the ringing

regions in the image by determining the edge/near-edge regions. The perceived ringing artifacts are then estimated by computing local variance of the Laplacian image corresponding to the detected ringing regions. The details are described as follows.

The grayscale input image I is filtered by a Laplace filter kernel κ [53, 54], whose impulse response is given by :

$$\kappa = \frac{1}{4}\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Let K denote the filtered-image, Laplacian image, obtained by convolving the input image I with the filter kernel κ. The value at point (x, y) of the filtered-image K is given by

$$K(x,y) = \frac{1}{4}\left\{I(x-1,y) + I(x+1,y) + I(x,y-1) + I(x,y+1) - 4 \times I(x,y)\right\}. \tag{4.2}$$

As proposed in Refs. 40, 43, 55, the perceived ringing artifacts can be estimated by the local variance of the intensity in the input image. Here, we assume that the local perceived ringing artifact can be estimated by the local variance of the Laplacian image K within the detected ringing regions in a block-based fashion. Let $\mu(b)$ and $\sigma^2(b)$ denote the overall mean and variance of block b of size 16×16 in the Laplacian image K. These values are computed by

$$\mu(b) = \frac{1}{16}\sum_{x=1}^{16}\frac{1}{16}\sum_{y=1}^{16} b(x,y), \tag{4.3}$$

$$\sigma^2(b) = \frac{1}{16}\sum_{x=1}^{16}\frac{1}{16}\sum_{y=1}^{16}[b(x,y) - \mu(b)]^2. \tag{4.4}$$

The local perceived ringing artifacts at location of block b in the input image I, denoted by $R(b)$, is defined as

$$R(b) = \begin{cases} 1 & \text{if } \sigma(b) < 1 \\ \sqrt{\sigma(b)} & \text{otherwise.} \end{cases} \tag{4.5}$$

The greater the $R(b)$ value, the greater the perceived ringing artifact at location of block b. The process is performed on every block b of size 16×16 from the Laplacian image K with 50% overlap between neighboring blocks, yielding a map R that represents the local perceived ringing artifact of the image.

4.2.3 Determine edge/near-edge regions and estimate perceived quality

In JPEG2000-compressed images, the ringing artifact manifests itself in the form of spurious oscillations in reconstructed pixel values in the vicinity of edges and causes severe visual degradation [55]. We follow Refs. [40, 55] to estimate perceived quality of a JPEG2000-compressed image by using the perceived blurring and ringing artifacts in the edge/near-edge regions of the input image.

To determine the edge locations, we apply the Canny edge detection [56] with two thresholds to obtain a binary edge map E. The threshold values have been chosen empirically. The value at point (x, y) of the binary edge map E is given by

$$E(x,y) = \begin{cases} 1 & \text{if } E(x,y) \text{ is an edge point,} \\ 0 & \text{if } E(x,y) \text{ is not an edge point.} \end{cases} \tag{4.6}$$

Each edge point $E(x, y)$ in the binary edge map is expanded to a $d \times d$ square with the center located at that point to obtain a binary dilated map D. Here, $d = 2\rho+1$ is an odd positive integer to ensure the existence of the center point. We further study the robustness of our algorithm against different ρ values in Section 4.3.2. The binary dilated map D is later used as a masking map to separate the edge/near-edge regions and the non-edge regions. Specifically, the pixel value of the binary dilated map D at point (x, y) is given by:

$$\begin{cases} D(x+\Delta x, y+\Delta y) \underset{\{-\rho \le \Delta x, \Delta y \le \rho\}}{} = 1 & \text{if } E(x,y) = 1 \\ D(x,y) = 0 & \text{otherwise} \end{cases} \tag{4.7}$$

The regions in the input image corresponding to $D(x, y) = 1$ are considered the edge/near-edge regions, denoted by R_{ed}; the regions of the input image corresponding

to $D(x, y) = 0$ are considered the non-edge regions, denoted by R_{ne}. A demonstration of these regions in a JPEG2000-compressed image is shown in Figure 4.3. Figure 4.3 shows the input JPEG2000-compressed image *sparrow* (a), the binary edge map E (b), and the binary dilated map D (c), the edge/near-edge regions (d), and the non-edge regions (e).

4.2.4 Combine maps and estimate image quality

For the given input image, we have computed the perceived blurring map $B(x, y)$, the perceived ringing map $R(x, y)$. We now combine these two maps to estimate the perceived quality of the input image.

The local perceived artifacts at the location of point (x, y) is estimated as the product of the local perceived blurring and the local perceived ringing at that location.

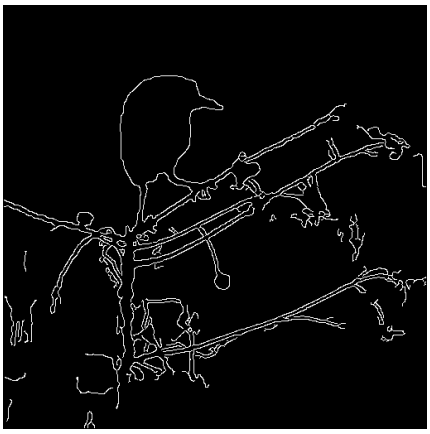
$$D(x, y) = B(x, y) \times R(x, y). \tag{4.8}$$

The quality of the input image is estimated by the root mean square of the inverse of the perceived artifacts in the edge/near-edge regions. Specifically, the quality index of the image, denoted by EDIQ, is given by

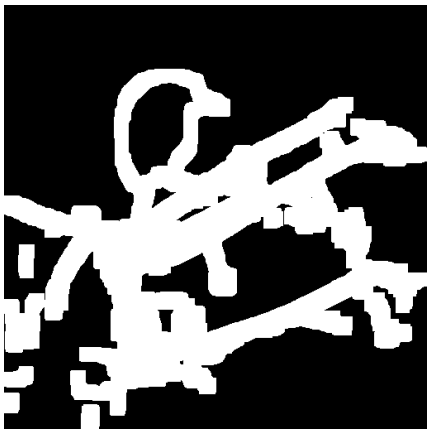
$$\text{EDIQ} = \sqrt{\frac{1}{N_b} \sum_{(x,y) \in R_{ed}} \left(\frac{1}{D(x, y)} \right)^2} \tag{4.9}$$

where N_b is the total number of points (x, y) in the edge/near-edge regions of the image. Here, EDIQ is a single scalar that represents the perceived quality estimate of the input JPEG2000-compressed image; the greater the EDIQ value, the better the quality of the input image.

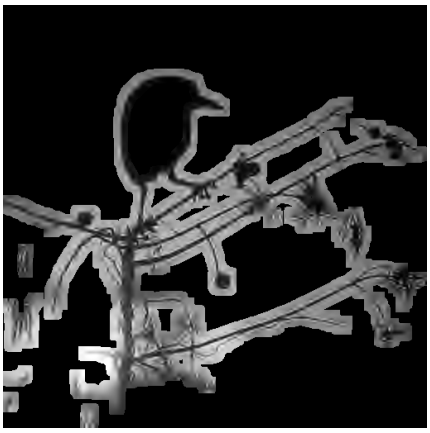
(a) JPEG2000-compressed



(b) Binary edge map



(c) Binary dilated map



(d) Edge/near-edge regions

(e) Non-edge regions

Figure 4.3: Illustration of a JPEG2000-compressed image *sparrow* with the separate edge/near-edge and non-edge regions. The figure shows the input image (a), the binary edge map (b), the binary dilated edge map (c), the edge/near-edge regions (d), and the non-edge regions (e) of the image.

4.3 Results

4.3.1 Performance Evaluation

To evaluate performance of the EDIQ algorithm, we used the subsets of JPEG2000-compressed images from six image-quality databases: A57 [57] (9 JPEG2000-compressed images), IVC [36] (50 JPEG2000-compressed images), TID [37] (96 JPEG2000-compressed images), TOYAMA [58] (98 JPEG2000-compressed images), CSIQ [34] (150 JPEG2000-compressed images), and LIVE [35] (169 JPEG2000-compressed images).

We compared EDIQ with three JPEG2000-designed no-reference algorithms (NSS [44], SAZHV [45], K1FB [47]) and three general no-reference image quality algorithms (BLIINDS-II [48], DIVIINE [49], BRISQUE [50]) on these image subsets. BLIINDS-II, DIVIINE, and BRISQUE are chosen because they are currently the best-performing general no-reference image quality algorithms. The implementation code for NSS, BLIINDS-II, DIVIINE, and BRISQUE are publicly available while the other two algorithms, SAZHV [45] and K1FB [47], are implemented by the authors following the algorithm description from the corresponding papers.

Before evaluating performance of each algorithm on each subset of JPEG2000-compressed images, we also applied Equation 3.6 to the raw predicted scores. The performance of predicting subjective ratings of quality was measured in terms of the Spearman rank-order correlation coefficient (SROCC) for gauging prediction monotonicity; the Pearson linear correlation coefficient (CC) (following non-linear regression; for gauging prediction consistency; and the outlier ratio (OR) [33] and outlier distance (OD) [34] for outlier analysis.

Table 4.1 shows the performance of EDIQ and other six algorithms in terms of SROCC, CC, RMSE, OR, and OD. The results of the best-performing algorithm are bolded and the results of the second best-performing algorithm are italicized and bolded. The results listed as "trained" are from algorithms that are trained on the

LIVE image database and we do not make a comparison for those training data.

As seen in Table 4.1, the EDIQ algorithm yields high correlation with the subjective ratings of quality in all six subsets of images. For all databases, EDIQ is the best-performing algorithm or competitive with the best-performing algorithm. On the A57, IVC, TID, and CSIQ databases, EDIQ yields the best performance in predicting image quality. EDIQ is competitive with NSS on the TOYAMA database. On the LIVE database, EDIQ is better than K1FB while all the other algorithms are trained on this database.

The present of a training step in an algorithm may cause that algorithm suffering from over-fitting to the training database. For examples, NSS has a poor performance on the CSIQ database despite its good performance on the TOYAMA and LIVE databases; DIIVINE and BLIINDS-II do not perform well on the TOYAMA database; and SAZHV fails to predict quality of JPEG2000 images on the TID database. Only BRISQUE algorithm does not suffer from training and maintains competitive performance comparing to EDIQ.

Figure 4.4 shows the scatter-plots of the transformed EDIQ scores versus the subjective ratings of quality for all six subsets of images from different databases.

4.3.2 Performance with Different Dilated Size

To study the robustness of the EDIQ algorithm when the dilated size changes, we measured the performance of the EDIQ algorithm on all image databases by using different dilated size of the edge pixels from 9×9 to 25×25. Table 4.2 shows the resulting SROCC and CC of this evaluation.

As seen from Table 4.2, with all the dilated size, EDIQ performs reasonably well on predicting quality of JPEG2000-compressed images from all databases. The change of the dilated size does not affect the performance of the EDIQ algorithm across all databases, which shows that EDIQ is robust to the change of dilated size.

Table 4.1: Performances of EDIQ and other no-reference (NR) IQA algorithms on six different subsets of JPEG2000-compressed images.

		General NR IQAs			**JPEG2000-specific NR IQAs**			
		BLDS-II	DVNE	BRSQ	NSS	SAZHV	K1FB	EDIQ
SROCC	**A57**	0.347	0.475	***0.700***	0.414	0.633	0.475	**0.917**
	IVC	***0.911***	0.654	0.833	0.776	0.786	0.866	**0.954**
	TID	***0.911***	0.842	0.904	0.366	0.649	0.769	**0.919**
	TOYA	0.711	0.630	0.891	**0.894**	0.869	0.785	***0.899***
	CSIQ	***0.884***	0.830	0.867	0.568	0.824	0.811	**0.937**
	LIVE	trained	trained	trained	trained	trained	***0.907***	**0.932**
CC	**A57**	0.506	0.645	0.717	0.481	***0.814***	0.645	**0.920**
	IVC	***0.915***	0.709	0.840	0.770	0.776	0.885	**0.957**
	TID	***0.919***	0.876	0.906	0.345	0.657	0.757	**0.928**
	TOYA	0.729	0.670	***0.907***	***0.907***	0.871	0.792	**0.917**
	CSIQ	***0.912***	0.893	0.896	0.581	0.842	0.830	**0.950**
	LIVE	trained	trained	trained	trained	trained	***0.913***	**0.937**
RMSE	**A57**	0.194	0.172	0.157	0.197	***0.131***	0.172	**0.088**
	IVC	***0.521***	0.910	0.699	0.823	0.814	0.602	**0.373**
	TID	***0.756***	0.925	0.810	1.799	1.446	1.252	**0.715**
	TOYA	0.983	0.778	***0.479***	***0.479***	0.559	0.694	**0.453**
	CSIQ	***0.130***	0.142	0.140	0.257	0.171	0.176	**0.099**
	LIVE	trained	trained	trained	trained	trained	***10.271***	**8.822**
OR	**IVC**	***8.00%***	18.00%	18.00%	16.00%	20.00%	12.00%	**4.00%**
	TID	**64.58%**	82.29%	***70.83%***	90.63%	81.25%	77.08%	75.00%
	TOYA	15.31%	14.29%	**2.04%**	***5.10%***	7.14%	10.20%	6.12%
	CSIQ	48.67%	50.00%	***42.00%***	57.33%	48.67%	51.33%	**36.67%**
	LIVE	trained	trained	trained	trained	trained	***59.76%***	**57.40%**
OD	**IVC**	***1.264***	8.124	4.282	7.317	6.656	1.335	**0.221**
	TID	***37.822***	53.802	42.389	126.680	90.468	69.634	**36.342**
	TOYA	8.644	5.537	**0.759**	2.295	***2.174***	4.435	2.019
	CSIQ	***5.922***	7.722	6.099	18.838	9.944	10.851	**4.050**
	LIVE	trained	trained	trained	trained	trained	***661.709***	**548.222**

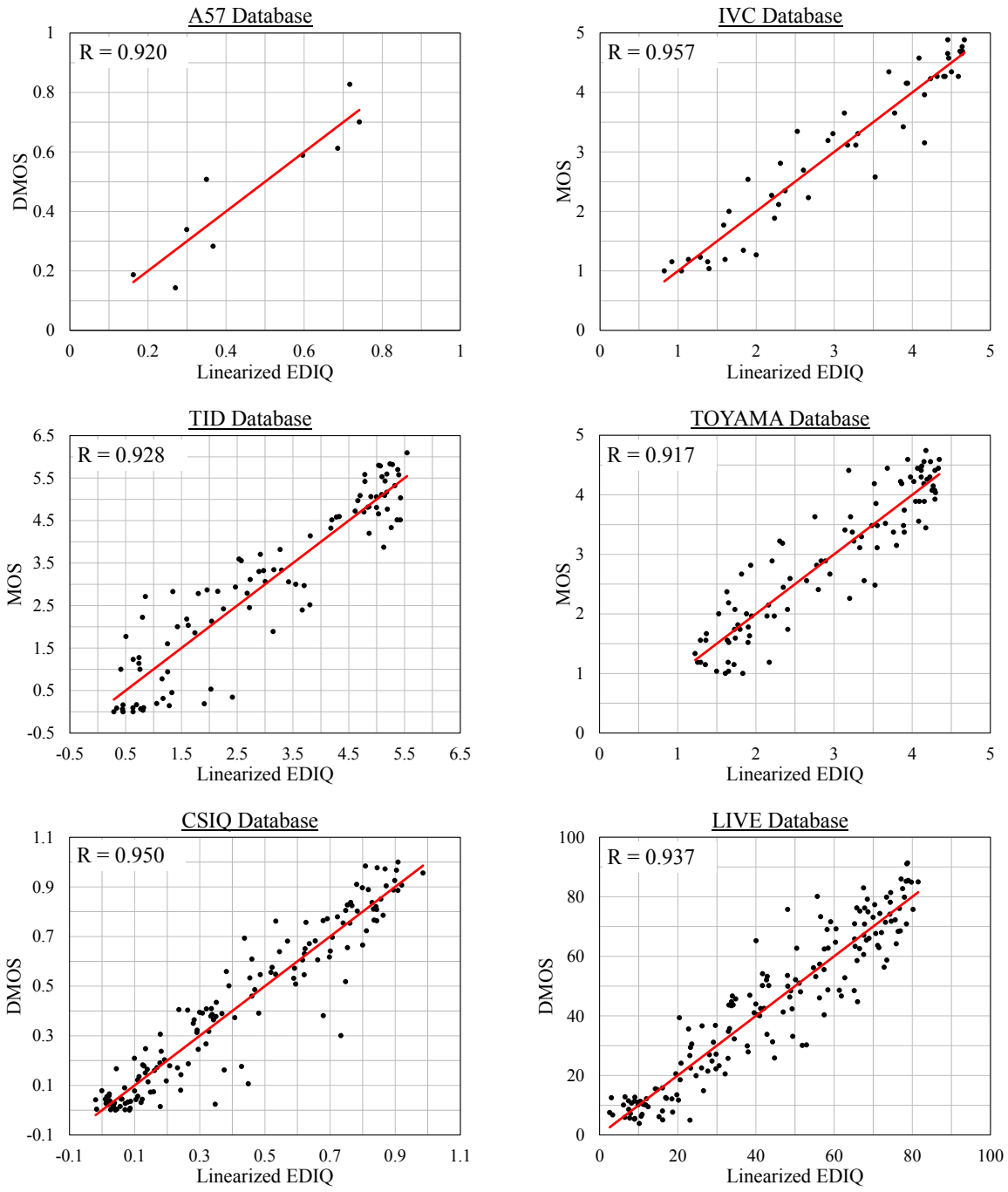


Figure 4.4: Scatter-plots of logistic-transformed scores predicted by EDIQ versus subjective scores on the six subsets of JPEG2000-compressed images. The R values denote correlation coefficient (CC) between the logistic-transformed scores and subjective ratings of quality.

4.4 Summary and Limitations

In this chapter, an edge-based no-reference image quality algorithm for JPEG2000-compressed images is proposed. The algorithm, named EDIQ, estimates the perceived quality of given JPEG2000-compressed image by estimating the local perceived blurring and ringing artifacts within the edge/near-edge regions. Perceived blurring is estimated by using the FISH sharpness estimator and perceived ringing is estimated by the local variance of the Laplacian image. Testing on various subsets of image-quality databases demonstrated that the EDIQ algorithm is competitive with current state-of-the-art algorithms of no-reference image quality assessment for JPEG2000-compressed images.

However, the EDIQ algorithm is not without limitations. One limitation of the EDIQ algorithm is that EDIQ takes into account only the distortion within the edge/near-edge regions and ignores the distortion in the non-edge regions. The artifacts in the non-edge regions might have potential effects on the perceived quality estimate of the image and we are currently study this effect.

The second limitation of the EDIQ algorithm comes from the FISH sharpness estimator, which is used to estimate perceived blurring artifacts. The FISH algorithm considers the smooth regions (e.g: sky, non-texture area) as the blur regions and assigns the high perceived blurring to these regions. This limitation is partly overcome by considering only the edge/near-edge regions but it still has potential effect to the estimation of image quality.

As a part of the performance evaluation, we found that most of the JPEG2000-compressed images in the tested image-quality databases are compressed by using the baseline JPEG2000 standard, which is only one of many JPEG2000 compression standards. The performance of the EDIQ algorithm in predicting quality of images compressed by other standards remains an opening question; EDIQ might need to modified or enhanced in order to perform that task accurately.

Table 4.2: Performances of EDIQ on six different JPEG2000-compressed image subsets with respect to different dilated size ρ. The results show the robustness of the EDIQ algorithm against dilated size change.

ρ	4	5	6	7	8	9	10	11	12
$d = 2\rho + 1$	9	11	13	15	17	19	21	23	25
	SROCC								
A57	0.950	0.950	0.950	0.917	0.917	0.917	0.917	0.917	0.917
IVC	0.955	0.955	0.955	0.953	0.954	0.954	0.954	0.954	0.953
TID	0.912	0.913	0.916	0.917	0.919	0.918	0.919	0.918	0.919
TOYA	0.888	0.890	0.894	0.897	0.899	0.901	0.905	0.907	0.907
CSIQ	0.937	0.937	0.937	0.937	0.937	0.937	0.936	0.937	0.937
LIVE	0.929	0.930	0.931	0.931	0.932	0.933	0.932	0.932	0.932
	CC								
A57	0.935	0.932	0.928	0.924	0.920	0.913	0.910	0.905	0.899
IVC	0.958	0.958	0.957	0.958	0.957	0.957	0.957	0.957	0.956
TID	0.925	0.926	0.926	0.927	0.928	0.929	0.929	0.929	0.930
TOYA	0.908	0.910	0.913	0.915	0.917	0.919	0.919	0.920	0.921
CSIQ	0.950	0.950	0.950	0.950	0.950	0.949	0.949	0.949	0.949
LIVE	0.932	0.934	0.935	0.936	0.937	0.937	0.937	0.937	0.937

CHAPTER 5

CONCLUSIONS

In this thesis report, we have presented the following contents

5.1 Fast Image Sharpness Measure

This chapter presented a simple, yet effective algorithm (FISH) for estimating both global and local image sharpness. By decomposing the input image via a three-level separable DWT, FISH estimates sharpness based on a weighted geometric mean of the DWT subband energies. FISH can also operate in a block-based fashion (FISH_{bb}) by applying the same computation to groups of DWT coefficients to generate a local sharpness map. The proposed algorithm shows good performance in predicting the local sharpness as well as the quality of the blurred images. We also demonstrated the efficacy of FISH/FISH_{bb} on predicting the level of blur in the Gaussian blur images.

5.2 JPEG2000 No-Reference Image Quality Assessment

An edge-based no-reference image quality algorithm for JPEG2000-compressed images is proposed in chapter 4. The algorithm, named EDIQ, estimates the perceived quality of given JPEG2000-compressed image by estimating the local perceived blurring and ringing artifacts within the edge/near-edge regions. Perceived blurring is estimated by using the FISH sharpness estimator and perceived ringing is estimated by the local variance of the Laplacian filtered-image. Testing on JPEG2000-compressed image subsets of various image-quality databases demonstrated that the EDIQ algorithm is competitive with current state-of-the-art no-reference IQA algorithms in

predicting quality of the JPEG2000-compressed images.

5.3 Future Work

5.3.1 General no-reference image quality assessment

The EDIQ algorithm proposed in chapter 4 can offer benefits for the camera manufacturer, image compression, etc. However, many types of distortion appear in the digital images, such as blur, blocking, white noise, contrast, and so on, as is illustrated in Figure 5.1. Therefore, it is necessary and more efficient to have a general no-reference image quality assessment that can predict image quality with all kinds of distortions in a manner that agrees with subjective ratings of quality obtained from human subjects.

With the success of FISH as an image sharpness algorithm and EDIQ, a no-reference quality assessment for JPEG2000-compressed images, we will extend and improve the EDIQ algorithm to make it become a no-reference quality assessment for many kinds of distortions other than JPEG2000 compression artifact. The current strategy of separating images into regions with different effects to human can be applied to other distorted images.

A popular trend in the research of no-reference image quality assessment in particular and in signal processing in general is the application of machine learning to develop a well-performed method. In term of no-reference image quality assessment, because the reference image is absent, it is difficult for the computer to judge image quality with a high correlation to the human subjects' opinion. Application of machine learning can help train the algorithms to predict image quality in the manner that human subjects do and obtain better performance.

White noise - DMOS = 0.442

Blur - DMOS = 0.590

Contrast - DMOS = 0.458

Ringing - DMOS = 0.693

Blocking - DMOS = 0.639

Pink noise - DMOS = 0.673

Figure 5.1: Different types of distortion that appears in digital images and their associated subjective ratings of quality. The quality of the image can vary across the amount of distortion and across the distortion types.

5.3.2 JPEG2000-compressed images with different compression standards

In our proposed algorithm, we only test the algorithm on the images that are compressed using the baseline JPEG2000 compression. However, our algorithm has not been validated for some other JPEG2000 compression standards. One reason is that that kind of image-quality database is not available. The performance of the proposed algorithm in predicting quality of images compressed by other standards remains an opening question; we might need to modify or enhance our algorithm in order to perform that task.

5.3.3 Application to video quality assessment

A video is a sequence of frames and each frame is a normal image, therefore, it is possible to apply an IQA algorithm to every frame of the video to estimate quality of that video. The quality index of every frame can be collapsed by a pooling method (Minkowski, arithmetic mean, etc.) to yield a single scalar. This index can serve as the representative quality index of the video, or it can be used as a part of a video quality assessment algorithm.

REFERENCES

[1] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *IEEE Signal Processing Letters*, vol. 19, pp. 423 – 426, Jul 2012.

[2] G. C. Higgins and L. A. Jones, "The nature and evaluation of the sharpness of photographic images," *Journal of the Society of Motion Picture and Television Engineers*, vol. 58, no. 4, pp. 277 – 290, 1952.

[3] D. Shaked and I. Tastl, "Sharpness measure: towards automatic image enhancement," in *IEEE International Conference on Image Processing*, vol. 1 of *Proc. ICIP.2005*, pp. I–937 – I–940, Sep 2005.

[4] C. T. Vu and D. M. Chandler, "Main subject detection via adaptive feature refinement," *Journal of Electronic Imaging*, vol. 20, Mar 2011.

[5] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *IEEE International Conference on Image Processing*, vol. 3 of *Proc. ICIP*, pp. III–57 – III–60, 2002.

[6] E. Ong, W. Lin, Z. Lu, S. Yao, X. Yang, and L. Jiang, "No-reference jpeg-2000 image quality metric," in *International Conference on Multimedia and Expo*, vol. 1 of *Proc. ICME*, pp. I–545 – I–548, Jul 2003.

[7] J. Dijk, M. V. Ginkel, R. J. V. Asselt, L. J. V. Vliet, and P. W. Verbeek, "A new sharpness measure based on gaussian lines and edges," in *Computer Analysis of Images and Patterns*, vol. 2756 of *Lecture Notes in Computer Science*, pp. 149 – 156, Springer, 2003.

[8] Y.-C. Chung, J.-M. Wang, R. Bailey, S.-W. Chen, and S.-L. Chang, "A non-parametric blur measure based on edge analysis for image processing applications," in *IEEE Conference on Cybernetics and Intelligent Systems*, vol. 1 of *Proc. CCIS*, pp. 356 – 360, 2004.

[9] S. Wu, W. Lin, L. Jian, W. Xiong, and L. Chen, "An objective out-of-focus blur measurement," in *Fifth International Conference on Information Communications and Signal Processing*, Proc. ICICS, pp. 334 – 338, 2005.

[10] H. Liu, J. Wang, J. Redi, P. Le Callet, and I. Heynderickx, "An efficient no-reference metric for perceived blur," in *European Workshop on Visual Information Processing*, EUVIP.2011, pp. 174 – 179, Jul 2011.

[11] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Transactions on Image Processing*, vol. 18, pp. 717 – 728, Apr 2009.

[12] N. D. Narvekar and L. J. Karam, "A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection," in *International Workshop on Quality of Multimedia Experience*, QoMEx.2009, pp. 87 – 91, Jul 2009.

[13] S. Corchs, F. Gasparini, F. Marini, and R. Schettini, "A sharpness measure on automatically selected edge segments," in *Image Quality and System Performance IX*, vol. 8293 of *Proc. SPIE*, pp. 82930A – 82930A–8, 2012.

[14] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, "The blur effect: perception and estimation with a new no-reference perceptual blur metric," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6492 of *Proc. SPIE*, Mar 2007.

[15] E. Tsomko and H.-J. Kim, "Efficient method of detecting globally blurry or sharp images," in *Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, Proc. WIAMIS, pp. 171 – 174, 2008.

[16] L. Debing, C. Zhibo, M. Huadong, X. Feng, and G. Xiaodong, "No reference block based blur detection," in *International Workshop on Quality of Multimedia Experience*, Proc. QoMEx, pp. 75 – 80, 2009.

[17] C.-Y. Wee and R. Paramesran, "Image sharpness measure using eigenvalues," in *9th International Conference on Signal Processing*, Proc. ICSP, pp. 840 – 843, 2008.

[18] X. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise," in *International Workshop on Quality of Multimedia Experience*, Proc. QoMEx, pp. 64 – 69, 2009.

[19] X. Marichal, W. Ma, and H. Zhang, "Blur determination in the compressed domain using DCT information," in *IEEE International Conference on Image Processing*, vol. 2 of *Proc. ICIP.1991*, pp. 386 – 390, 1999.

[20] M. Kristan, J. Perse, M. Perse, and S. Kovacic, "A bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform," *Pattern Recognition Letters*, vol. 27, no. 13, pp. 1431 – 1439, 2006.

[21] N. Zhang, A. E. Vladar, M. T. Postek, and B. Larrabee, "A kurtosis-based statistical measure for two-dimensional processes and its application to image sharpness," *Section of Physical and Engineering Sciences of American Statistical Society*, pp. 4730 – 4736, 2003.

[22] R. Hassen, Z. Wang, and M. Salama, "No-reference image sharpness assessment based on local phase coherence measurement," in *IEEE International Conference*

on Acoustics Speech and Signal Processing, Proc. ICASSP.2010, pp. 2434 – 2437, Mar 2010.

[23] C. Hsin, J.-W. Jang, S.-J. Shin, and S.-H. Chen, "A no-reference objective image sharpness metric based on a filter bank of gaussian derivative wavelets," in *International Conference on Multimedia Technology*, Proc. ICMT, pp. 3362–3365, 2011.

[24] J. Caviedes and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis," in *International Conference on Image Processing*, vol. 3 of *Proc. ICIP*, pp. III–53 – III–56, 2002.

[25] M.-J. Chen and A. Bovik, "No-reference image blur assessment using multiscale gradient," *EURASIP Journal on Image and Video Processing*, vol. 2011, no. 1, pp. 1 – 11, 2011.

[26] C. T. Vu, T. D. Phan, and D. M. Chandler, "S_3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Transactions on Image Processing*, vol. 21, pp. 934 – 945, Mar 2012.

[27] D. J. Field and N. Brady, "Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes," *Vision Research*, vol. 37, no. 23, pp. 3367 – 3383, 1997.

[28] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259 – 268, Nov. 1992.

[29] C. Chen, W. Chen, and J. A. Bloom, "A universal reference-free blurriness measure," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7867 of *Proc. SPIE*, Jan. 2011.

[30] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd ed., 2008.

[31] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 45, pp. 485 – 560, Jun 1992.

[32] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment, phase II," tech. rep., Video Quality Expert Group, Aug 2003. http://www.vqeg.org.

[33] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 121 – 132, 2004.

[34] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, p. 011006, 2010.

[35] H. Sheikh, Z. Wang, L. Cormack, and A. Bovik, "Live image quality assessment database release 2. available online.," 2005.

[36] A. Ninassi, P. Callet, and F. Autrusseau, "Pseudo no reference image quality metric using perceptual data hiding," in *Human Vision and Electronic Imaging XI*, vol. 6057 of *Proc. SPIE*, pp. 146 – 157, Feb 2006.

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

[38] Computational Perception & Image Quality Lab Oklahoma State University, "S3: A spectral and spatial measure of local perceived sharpness in natural images. online supplement.," 2012. Available: http://vision.okstate.edu/s3.

[39] H. Tong, M. Li, H. J. Zhang, and C. Zhang, "No-reference quality assessment for JPEG2000 compressed images," in *IEEE International Conference on Image Processing*, vol. 5 of *Proc. ICIP.2004*, pp. 3539 – 3542, Oct 2004.

[40] R. Barland and A. Saadane, "Blind quality metric using a perceptual importance map for JPEG2000-compressed images," in *IEEE International Conference on Image Processing*, Proc. ICIP.2006, pp. 2941 – 2944, Oct 2006.

[41] L. Liang, S. Wang, J. Chen, S. Ma, D. Zhao, and W. Gao, "No-reference perceptual image quality metric using gradient profiles for JPEG2000," *Signal Processing: Image Communication*, vol. 25, no. 7, pp. 502 – 516, 2010. Special Issue on Image and Video Quality Assessment.

[42] X. K. Yang, W. S. Ling, Z. K. Lu, E. P. Ong, and S. S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing: Image Communication*, vol. 20, no. 7, pp. 662 – 680, 2005.

[43] H. Liu, N. Klomp, and I. Heynderickx, "A no-reference metric for perceived ringing artifacts in images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 4, pp. 529 – 539, 2010.

[44] H. R. Sheikh, A. C. Bovik, and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Transactions on Image Processing*, vol. 14, pp. 1918 – 1927, Nov 2005.

[45] Z. M. P. Sazzad, Y. Kawayoke, and Y. Horita, "No reference image quality assessment for JPEG2000 based on spatial features," *Signal Processing: Image Communication*, vol. 23, no. 4, pp. 257 – 268, 2008.

[46] J. Zhang and T. M. Le, "A new no-reference quality metric for JPEG2000 images," *IEEE Transactions on Consumer Electronics*, vol. 56, pp. 743 – 750, May 2010.

[47] J. Zhang, S. H. Ong, and T. M. Le, "Kurtosis-based no-reference quality assessment of JPEG2000 images," *Signal Processing: Image Communication*, vol. 26, no. 1, pp. 13 – 23, 2011.

[48] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.

[49] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, pp. 3350 – 3364, Dec 2011.

[50] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, pp. 4695 – 4708, Dec 2012.

[51] H. Tong, M. Li, H. Zhang, and C. Zhang, "Blur detection for digital images using wavelet transform," in *IEEE International Conference on Multimedia and Expo*, vol. 1 of *Proc. ICME.2004*, pp. 17 – 20, Jun 2004.

[52] P. V. Vu and D. M. Chandler, "A no-reference quality assessment algorithm for jpeg2000-compressed images based on local sharpness," in *Image Quality and System Performance X*, vol. 8653 of *Proc. SPIE*, pp. 865302–865302–8, 2013.

[53] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629 – 639, 1990.

[54] P. Bourdon, B. Augereau, C. Olivier, and C. Chatellier, "A pde-based method for ringing artifact removal on grayscale and color jpeg2000 images," in *International Conference on Multimedia and Expo*, vol. 2 of *Proc. ICME.2003*, pp. II–149 – 52, 2003.

[55] S. H. Oguz, Y.-H. Hu, and T. Q. Nguyen, "Image coding ringing artifact reduction using morphological post-filtering," in *IEEE Second Workshop on Multimedia Signal Processing*, Proc. MMSP.1998, pp. 628 – 633, 1998.

[56] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679 – 698, 1986.

[57] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Transactions on Image Processing*, vol. 16, pp. 2284 – 2298, Sep 2007.

[58] A. Ninassi, O. Le Meur, P. Le Callet, and D. Barba, "Which semi-local visual masking model for wavelet based image quality metric?," in *IEEE International Conference on Image Processing*, Proc. ICIP.2008, pp. 1180 – 1183, Oct 2008.

VITA

Phong Van Vu

Candidate for the Degree of

Master of Science

Thesis: ON THE USE OF IMAGE SHARPNESS MEASURE TO JPEG2000 NO-REFERENCE IMAGE QUALITY ASSESSMENT

Major Field: Electrical and Computer Engineering

Biographical:

Education:

Completed the requirements for the Master of Science in Electrical and Computer Engineering at Oklahoma State University, Stillwater, Oklahoma in December, 2013.

Completed the requirements for the Bachelor of Science in Electronics and Telecommunications at Posts and Telecommunications Institute of Technology, Hanoi, Vietnam in 2004.

Publications:

P. V. Vu and D. M. Chandler, "A No-Reference Quality Assessment Algorithm for JPEG2000-Compressed Images Based on Local Sharpness," *SPIE Image Quality and System Performance*, Burlingame, CA, USA, February 2013.

P. V. Vu and D. M. Chandler, "A Fast Wavelet-Based Algorithm for Global and Local Image Sharpness Estimation," *IEEE Signal Processing Letters*, vol. 19, no. 7, pp. 423-426, Jul 2012.