

No-Reference Stereoscopic Image Quality Assessment

Appina Balsubramanyam

A Thesis Submitted to
Indian Institute of Technology Hyderabad
In Partial Fulfillment of the Requirements for
The Degree of Master of Technology



भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Electrical Engineering

July 2015

Declaration

I declare that this written submission represents my ideas in my own words, and where ideas or words of others have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.



(Signature)

Appina Balsubramanyam

(Appina Balsubramanyam)

EEEN14P10001


(Roll No.)

Approval Sheet

This Thesis entitled No-Reference Stereoscopic Image Quality Assessment by Appina Balsubramanyam is approved for the degree of Master of Technology from IIT Hyderabad



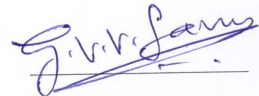
(Dr. Ketan P. Detroja) Examiner
Dept. of Electrical Eng
IITH



(Dr. K Sri Rama Murty) Examiner
Dept. of Electrical Eng
IITH



(Dr. Sumohana S. Channappayya) Adviser
Dept. of Electrical Eng
IITH



(Dr. G V V Sharma) Chairman
Dept. of Electrical Eng
IITH

Acknowledgements

I would like to thank my guide Dr. Sumohana Channappayya for his aspiring guidance and valuable suggestions and strong moral support. I would like to thank my labmates Sameeulla Khan MD, K Manasa and K.V.S.N.L. Manasa Priya for their valuable suggestions. I am also thankful to all my friends at IIT Hyderabad. Especially, L Praveen Kumar Reddy and P Charantej Reddy deserve a special mention for the enormous amount of help they have provided.

Abstract

We present two contributions in this work: i) a bivariate generalized Gaussian distribution (BGGD) model for the joint distribution of luminance and disparity subband coefficients of natural stereoscopic scenes, and ii) a no-reference (NR) stereo image quality assessment algorithm based on the BGGD model. We first empirically show that a BGGD accurately models the joint distribution of luminance and disparity subband coefficients. We then show that the model parameters form good discriminatory features for NR quality assessment. Additionally, we rely on the previously established result that luminance and disparity subband coefficients of natural stereo scenes are correlated, and show that correlation also forms a good feature for NR quality assessment. These features are computed for both the left and right luminance-disparity pairs in the stereo image and consolidated into one feature vector per stereo pair. This feature set and the stereo pair's difference mean opinion score (DMOS) (labels) are used for supervised learning with a support vector machine (SVM). Support vector regression is used to estimate the perceptual quality of a test stereo image pair. The performance of the algorithm is evaluated over popular databases and shown to be competitive with the state-of-the-art no-reference quality assessment algorithms. Further, the strength of the proposed algorithm is demonstrated by its consistently good performance over both symmetric and asymmetric distortion types. Our algorithm is called Stereo Quality Evaluator (StereoQUE).

Contents

Declaration	ii
Approval Sheet	iii
Acknowledgements	iv
Abstract	v
Nomenclature	vi
1 Introduction	1
1.1 Evolution of 3D multimedia systems:	1
1.1.1 Types of quality assessment methods and metrics:	2
1.2 Stereo Vision:	3
1.2.1 Disparity:	3
2 Literature	6
2.1 HVS inspired approaches	6
2.2 NSS-inspired NR-IQA Algorithms	9
2.2.1 Stereoscopic NSS Models	10
3 Metric development	12
3.1 Disparity calculation:	13
3.2 Stereable Pyramid decomposition:	14
3.3 Bivariate GGD Modeling & Correlation calculation:	14
3.4 Feature Consolidation:	18
3.5 Supervised Learning and Regression:	19
4 Results and Discussions:	20
5 Conclusion	24
References	25

Chapter 1

Introduction

1.1 Evolution of 3D multimedia systems:

The day by day advancements in the multimedia technology and devices, the consumer acceptance increases in tremendous way. Owing to these causes many research institutes and industries zones showing much attention in the development of this technology. This is especially true in the case of 3D or stereoscopic images and videos. The 3d imaging technology starts at the earlier time of 19th century but still it is a developing technology. With the ubiquitous usage of digital 3D content, so many movie industries and gaming zones are putting a large amount of budget on preparation of 3D content [1]. Now this technology is not ended up with the theater screen but it also extends to the mobile phones and broadcast televisions. This technology sends the movie lovers to a new world which feels like good entertainment. According to the statistics of MPAA [2], nearly one third of the general movie lovers at least attended a 3D movie in 2013. Due to this cause recently many theater screens are changed to 3D from 2D screen. Due to the involvement of several processing steps (sensors, compression, conversions, etc) involved on processing of multimedia content, quality degradation is happening to the end user. Based on present and future situation on the availability of digital 3D content, there is a need to improve the quality of experience (QOE) to the end user. The QOE of an observer depends on how perceptually feeling when he viewing that multimedia content. There are two types of methods to assess the quality of multimedia content. Those are Subjective assessment and objective assessment.

1.1.1 Types of quality assessment methods and metrics:

Quality Assessment Methods

In subjective assessment task, the human observers are giving their score of multimedia content based on how he is feeling when he perceive the multimedia content under certain conditions. Subjective assessment task is impractical and time consuming and cumbersome. But the subjective assessment scores are very important to develop or performance improvement of objective quality method. Because the subjective scores are best quality scores to mimic with HVS. So there is a need of good objective quality metric to predict the quality score of multimedia content without using any pre reference score. The pre reference scores are mean opinion scores (MOS) or difference mean opinion scores (DMOS). These objective experiments are subset of Psychophysical experiments because in these experiments we are trying to mimic with the Human Visual System (HVS). In the case of 2D multimedia content tremendous growth happening in the development of objective metric but in stereo or 3D it is very limited. There is need of good improvement in the developing of good objective metrics for 3D.

Quality assessment Metrics

There are three types of quality metrics are available to predict the quality score of given multimedia content. All these methods are working based on the availability of reference or pristine multimedia content. The Full reference image quality (FRQA) [3] model uses the entire information regarding to the reference multimedia content. In this metric, we compare the distorted and reference content based on our requirements, and it predict the quality score based on this comparison. In the Reduced reference image quality (RRQA) [4] model uses some fractional information about the reference or pristine content but in the case of No reference image quality (NRQA) [5, 6, 7] assessment does not require any information regarding to the reference or pristine content. Compare to all these methods NRQA is the robust method because of non-usage of reference content. Most of real time transmissions such as multimedia services in the context of wireless, telecom links did not contain any information regarding to the reference content. In this type of situations the NRQA is the robust method compare to the other quality metrics.

1.2 Stereo Vision:

Two frontal horizontally spaced eyes (HVS system or binocular visual system) always receive two slightly different images when we perceive one object or one scene. Our human brain will perform complex manipulations on these two images to create a 3D vision. Based on these complex manipulations we are getting some additional information regarding to the depth or disparity. This disparity or depth visualization create good difference in between the 2D and 3D perception. In the



(a) Left Stereo image.



(b) Right Stereo image.

Figure 1.1: Left and Right Stereoscopic views.

above Figure 1.1 shows the left and right views of stereoscopic image. It is clearly we can observe that both images are not same because the left image didn't contain the right most part of right image and the right image didn't contain the left most part of the left image.

1.2.1 Disparity:

Binocular disparity refers to the difference or deviation of image object location in the both views. The brain uses the disparity information to create the illusion of depth information.

The disparity of two retinal images on the actual retina depends on factors internal to the eye, especially the location of the nodal points, even if the cross section of the retina is a perfect circle. Disparity on retina conforms to binocular disparity when measured as degrees, while much different if measured as distance due to the complicated structure inside eye. In the Figure 1.2 shows the model to compute the disparity.

The full black circle is the point of fixation of both retinas. The blue object lies nearer to the observer. Therefore it has a "near" disparity d_n . Objects lying more far away (green) correspondingly have a "far" disparity d_f . Binocular disparity is the angle between two lines of projection in one eye (Mathematically, $d_n - d_f$, with sign, measured counterclockwise). One of which is the real projection from the object to the actual point of projection. The other one is the imaginary projection running

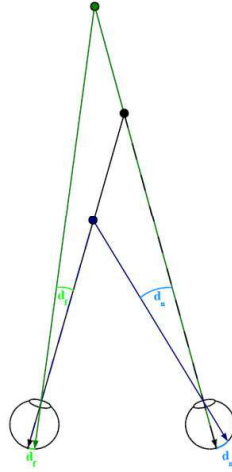


Figure 1.2: Disparity computation.

through the nodal point of the lens of the one eye to the point corresponding to the actual point of projection in the other eye. For simplicity reasons here both objects lie on the line of fixation for one eye such that the imaginary projection ends directly on the fovea of the other eye, but in general the fovea acts at most as a reference. Note that far disparities are smaller than near disparities for objects having the same distance from the fixation point.

Till now so many disparity computation algorithms are available to compute disparity map [8]. To compute disparity map we considered that the pixel positions are deviated only in vertical direction but not in horizontal position. The choosing of disparity algorithm we should consider different constraints based on computation complexity, time consuming, good visualization map and robustness. So that, to compute disparity map all algorithms are going to search in same horizontal position. We are hypothesizing that these algorithms are mimicing with the HVS because all these algorithms are showing very annoying errors (pixel errors) at occluded places.

The quality assessment of stereoscopic is much different from the 2D image quality assessment. The main difference between the 2D and stereo is the third dimension (disparity). The disparity plays a good role when we try to assess the quality of stereoscopic image. Chen et al. [9] studied the effect of binocular depth variation of stereoscopic image in the perceptual quality of a stereoscopic image. They finally concluded that the overall quality experience is going to decrease when the depth increases and also they showed that for 2D images the (QOE) will not effect when the depth variation happened.

In my thesis work I am going to present the No reference quality assessment algorithm for the stereoscopic images based on the Natural Scene Statistics (NSS) features. In this work totally relied

in the joint statistics of luminance and disparity. These statics are going to modeled with the help of Bivariate Generalized Gaussian Fitting (BGGD) fitting. It shows that good discrimination with the type and level of distortion. The correlation values between the luminance and the disparity coefficients served as one more feature for this algorithm.

Chapter 2

Literature

Stereoscopic image quality assessment receives very less attention compared to the 2D image quality assessment in all cases (FRQA,RRQA, NRQA). This could be attributed to several factors including but not limited to the smaller fraction of stereoscopic content compared to 2D content, an inclination to apply 2D metrics to each of the left and right stereo images, and paradoxically, the visual annoyance resulting from poorly designed stereo imaging systems. NR stereo quality assessment has received even lesser attention than FR and RR methods. We first review significant human visual system (HVS) inspired FR, RR and NR stereo image quality assessment algorithms. To place our proposed algorithm in perspective, we briefly review NSS-based 2D and stereo NR-IQA algorithms, followed by a review of natural stereo-scopic scene statistical models.

2.1 HVS inspired approaches

In the HVS inspired approaches, based on disparity they considered two types of approaches in quality assessment task. Campisis et al. [10] did a systematic study about the stereoscopic image quality assessment. In their study, they perform the subjective evaluation and proposed a FRQA metric based on the consideration of left and right views. They used 2D FR-IQA algorithms on the left and right views individually and combined these scores based on perceptually inspired approaches. Finally, concluded that the 2D-IQA algorithms on the individual views necessarily can not give a perfect score of a stereoscopic image set. Gorely et al. [11] also proposed a FRQA metric based on the Stereo Band Limited Contrast (SBLC) feature. The above two methods did not use any disparity map computation in their algorithms. Benoit et al. [12] are the first persons extensively using of disparity map in their metrics. They proposed a FRQA metric based on Local and global

luminance quality scores. The performance of this algorithm has been shown to be competitive over commonly occurring distortions. You et al. [13] also did the similar work based on the Local and global luminance quality scores. But they extensively showed on the LIVE database and Custom database. These two algorithms mainly concentrated on the disparity computation. Wang et al. [14] conducted a systematic study on symmetrically distorted stereoscopic images. In the symmetrically distorted means, both views are annoying with same level of distortion. They concluded that for symmetrically distorted views the overall perceptual quality score is approximately equal to the average of left and right views but in case of asymmetric it is not working. Bensalma and Larabi [15, 16, 17] proposed a series of FRQA approaches based on the modeling of binocular energy. In their method, to model the binocular energy they extensively used the simple cells and complex cells (in visual cortex) representation. To represent these cells they used the Bandlet wavelet transform and Complex wavelet transform. Ryu et al. [18] proposed a FR-IQA for stereoscopic image based on binocular energy modeling and the SSIM (i.e. luminance, contrast, structural similarity) parameters. They computed these parameters for each view finally they pooled these scores to get a single score. Li et al. [19] proposed a stereoscopic FR-IQA metric based on structure and texture decomposition. They measured gradient magnitude similarity and luminance and contrast similarity from the structure and texture decompositions. They compute these values for each image and finally combine these scores to come up with a single score for a stereo pair. Shen et al. [20] proposed a stereoscopic FR-IQA metric based on disparity map segmentation and SSIM values. In their metric, they computed the disparity maps for the pristine and distorted images and they computed the SSIM scores for these segmented disparity maps. HU et al. [21] proposed a stereoscopic full reference image quality metric based on distortion separation method. They used left and right views of stereoscopic view individually for distortion separation. Later they applied the singular value decomposition (SVD) and phase-amplitude description (PAD) methods. These scores are pooled to compute a quality score for the stereo image pair. Zhou et al. [22] proposed a RR-IQA metric based on the degradation level of water marking on symmetric distorted images. They embedded water marking on the extracted features, later at the receiver end they compared the degradation level of water marking on account of distortion.

Till now all these metrics are working based on individual views of stereoscopic view and disparity map. Later, so many researchers were proposed their metrics based on cyclopean model.

The cyclopean image paradigm has proven to be a successful approach for integrating depth information into a 2D image and subsequently relying on 2D IQA algorithms for quality assessment. Maalouf and Larabi [23] proposed a RR-IQA stereoscopic metric based on the cyclopean model of

stereoscopic image and color disparity maps. They construct individual cyclopean images and color disparity maps for the reference and distorted set and find the HVS-based thresholds. Later, they compare the sensitivity coefficients of reference and distorted images to predict the quality score. Chen et al. [24] proposed a FR-IQA metric for stereoscopic images based on the cyclopean paradigm. They considered the effect of binocular rivalry on the stereoscopic image that are asymmetrically distorted. They showed that the effect on quality score with and without the use of disparity information. Chen et al. [25] also proposed a NR-IQA metric for stereoscopic image based on the cyclopean model and NSS features. In this method, they again considered the effect of binocular rivalry on perceptual quality. In their approach, they used a combination of 2D features extracted from the cyclopean image and 3D features extracted from the uncertainty map produced by the stereo matching algorithm. Chetouani et al. [26] proposed a stereoscopic FR-IQA metric that is also based on the cyclopean image paradigm and 2D image quality metric (IQM) fusion. They extracted 2D features from the cyclopean image and used an artificial neural network to predict the quality score.

Akther and Sezzad et al. proposed NRQA algorithm based on the local edge based and non edge based disparity information. Their work is based on the concept of high frequency components in the image because to get the disparity map the edges of an object is an important thing. Additionally they used the blockiness estimates for the distortion identification features. Mainly these people worked on the jpeg compressed artifacts. Saho et al. explored the NRQA metric for stereoscopic images based on the distortion specific features. They used different types of features to discriminate the type of the distortion i.e. features are varied with type of distortion. Later they used Support Vector Regression (SVR) to prepare the relation between the extracted features and the human observer scores. Fezza et al. [27] proposed a NRQA metric based on the blurriness maps and the local entropy values. In their method they calculated the blurriness maps for each stereo image. To calculate these blurriness maps they consider a thresholds based on the occluded region. They used the local entropy values as weights to pool these scores. Finally, they take the average between occluded region score and non-occluded region Score. The above methods are working based on the modelling of HVS. All these deterministic methods (FRQA, RRQA & NRQA) shows very good correlation with the HVS scores.

But recently many researchers are using the statistical modelling of features. These methods are showing even better performance compared to the above approach. In the following approach. I totally relied on the Natural Scene Statistical (NSS) modelling. The meaning of natural images meaning those are not processed with any computer application. This approach already used in the

context of 2D image and video assessment task. In their methods this approach showed very good performance with Dmos scores. So I inspired from these approaches to use the NSS modeling. Now we are reviewing the statistically inspired quality assessment methods.

2.2 NSS-inspired NR-IQA Algorithms

While deterministic (HVS-inspired) approaches to FR, NR, and RR IQA have been quite popular and successful, statistical approaches have also proven to be very successful. Our proposed algorithm is motivated by the success of these statistical approaches. We now review a small but significant subset of NSS-inspired 2D and stereo NR-IQA methods.

The generalized Gaussian density (GGD) and the Gaussian scale mixture (GSM) density are two models for 2D natural scene statistics [28, 29] that are very popular and have been widely employed in 2D IQA. Moorthy et al. [30] studied NSS feature (GSM and GGD model parameter) variations over natural and unnatural scenes. They proposed a NR-IQA metric for 2D images based on a two-stage frame work. With the help of NSS features they classified distorted image into different distortion classes and used these statistical features to predict the quality score of the distorted image. Mittal et al. [31, 32] proposed a 2D NR-IQA metric based on statistical modeling of natural scenes in the pixel domain. They considered mean subtracted contrast normalized luminance coefficients (MSCN coefficients) in the spatial domain and computed pairwise product of adjacent normalized luminance coefficients to get the distortion specific information. Asymmetric GGD fitting model was used to fit the MSCN coefficients in their method to discriminate the type of distortion. Saad et al. [6] proposed a 2D image NR-IQA metric based on NSS modeling of DCT coefficients, again using a univariate GGD model. Their method does not use any distorted specific features. They applied the sample Bayesian inference model to predict the quality score of an image. Sheikh et al. [33] proposed a 2D image NR-IQA metric based on the NSS modeling of JPEG compressed images. They showed that the deviation or variation of expected natural statistics of respective image is related to the perceptual quality. Chikkerur et al. [34] proposed a 2D video FR-IQA and RR-IQA quality metric based on the statistical modeling of natural visual characteristics and perceptual characteristics. Natural visual characteristics are divided into natural visual statistics and natural visual features. Similarly, perceptual characteristics are classified into either frequency or pixel-domain methods. Based on the classification scheme, they used MS-SSIM as a quality metric for natural visual statistics, VQAM as a quality metric for natural visual features and the MOVIE index

for perceptual quality characterization. Yeskethu et al. [35] proposed a NR-IQA stereo image and video quality metric based on the standard methods. They showed that VQM is a good metric to predict the overall quality of a video and PSNR, SSIM provides better in depth perception. Mittal et al. [36] proposed a NR stereoscopic quality of experience assessment metric based on the statistics of luminance images and disparity maps. A slew of statistics including mean, variance, skewness, kurtosis, differential disparity etc. are computed on the disparity maps. The mean, variance and kurtosis of spatial activity maps of the left and right images are also computed. These statistics compose the feature vector for the stereo pair. Dimensionality reduction techniques such as principal components analysis (PCA) and forward feature selection (FFS) are employed on a training images to reduce feature vector dimension. Finally, linear regression is used to estimate the perceived quality of experience. For video quality assessment, the same method is applied on a frame-by-frame basis. However, it is important to note that this approach, though based on statistical features, does not attempt to model and exploit the rich and unique signature of natural stereo scene statistics. In the following, we review statistical models of natural stereoscopic scenes.

2.2.1 Stereoscopic NSS Models

The statistical characteristics of natural stereoscopic scenes have been studied by a number of researchers. For stereoscopic images, the modeling is typically done on subband luminance and range/disparity coefficients – either jointly or marginally. Huang et al. [37] model the range statistics of the natural environment. To construct these statistics they used a laser range finder [38] to acquire the range maps of the natural environment. To describe the properties of range maps they consider the single pixel and derivative statistics. The discontinuities of range maps are explored with help of Haar wavelets. Liu et al. [39] construct the disparity information from the depth map which is created by the range finder when a fixed point described. To construct these maps, a spherical model of the eye structure is employed. Finally, they correlate the disparity distribution with stereopsis of HVS.

Potetz and Lee [40] and Liu et al [41, 42] explored the statistical relationship between the luminance and disparity maps in the spatial domain and in a multiscale subband domain respectively. They modeled conditional histograms of luminance and disparity/range using a univariate GGD [43] model. They also demonstrated considerable correlation between the luminance and range/disparity subband coefficients. Su et al. [44] conducted a study on the relationship between the chrominance and range components. They showed that the conditional distribution of luminance or chrominance

on the range gradients can be modeled with a Weibull distribution. Su et al. [45] proposed a stereoscopic NR-IQA metric based on the bivariate statistical correlation model (a bivariate GGD model) for spatially adjacent luminance subband coefficients. This model is applied to capture the joint statistics of subband coefficients of a cyclopean image. In addition to the bivariate statistics, they also rely on univariate NSS to form their feature vector. Support vector machine training is performed using these features and DMOS scores as labels. Support vector regression is used to predict the quality of a test stereo image pair. This algorithm is the current state-of-the-art NR stereo IQA technique and outperforms state-of-the-art FR stereo IQA algorithms as well.

Chapter 3

Metric development

Our proposed algorithm is closest in philosophy to the work by Su et al. [45]. As noted previously, their approach models adjacent luminance subband coefficients of cyclopean images using a bivariate GGD. However, they do not model the joint statistics of luminance and disparity/depth subbands. The flowchart of the proposed algorithm is shown in Fig. 3.1 and is described in the following subsections.

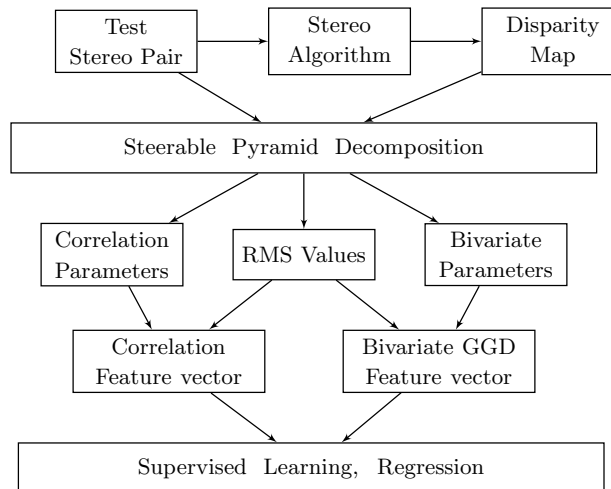


Figure 3.1: Flowchart of proposed algorithm.

To construct any NRQA metric, the available information to us is very minimum. Based on the available information we extracted some additional features which are very helpful to us. In our approach we used only the basic information of stereoscopic view i.e. left and right view. The proposed approach is starting with calculation of disparity map.

3.1 Disparity calculation:

The disparity evaluation is working on the basic principle of best matching between the left and right view in the same horizontal direction. There are several algorithms are used to evaluate the disparity map [8].

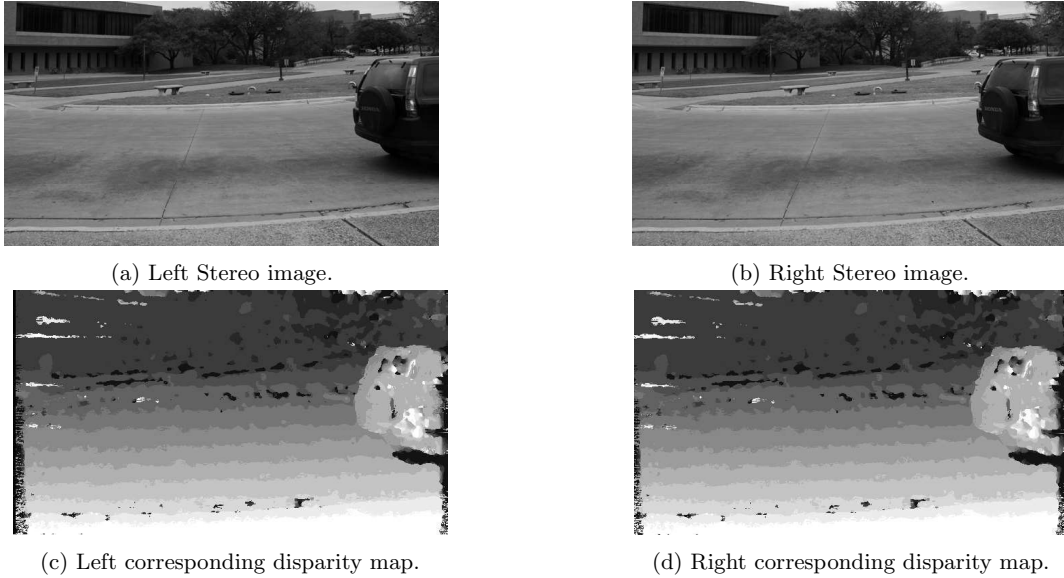


Figure 3.2: Left and Right Stereo images and corresponding disparity maps.

Based on the effect of time consuming and accuracy, complexity we could choose the SSIM based stereo matching algorithm. Even though high accuracy algorithms Belief propagation, adaptive window algorithm etc ... performed well in terms of disparity map because of time consuming and complexity we did not opted such algorithms and some less complex algorithms like SAD, SSD etc ... taking less time but those are not computing accurate disparity maps. So finally we relied on SSIM approach [25] with moderate speed and better disparity map condition. The SSIM stereo matching algorithm is working based on the best ssim score between two patches of left and right view. Here in our case we limited the pixel variation in disparity map upto 15 only. Two sets of disparity maps are computed per stereo pair one with the left image treated as reference and the other with the right image treated as reference. In the Figure 3.2 it shows the left and right stereo images and their corresponding disparity maps. The information from the disparity map is brighter pixels are corresponding to the nearer objects and black pixels corresponding to the far objects.

3.2 Steerable Pyramid decomposition:

The Steerable Pyramid is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications [46]. The steerable pyramid performs a polar-separable decomposition in the frequency domain, thus allowing independent representation of scale and orientation [47, 48]. More importantly, the representation is translation-invariant (i.e., the subbands are aliasing-free, or equivariant with respect to translation) and rotation-invariant (i.e., the subbands are steerable, or equivariant with respect to rotation). This can make a big difference in applications that involve representation of position or orientation of image structure.

At the outset we would like to note that our analysis is carried out on a multi-scale/multi-orientation subband decomposition of the stereo images and their disparity maps. This is motivated by the multiscale modeling of the human visual system. Specifically, we work with a steerable pyramid decomposition [47, 48] at 3 spatial scales and six orientations ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 150^\circ, 180^\circ$). The subband decomposition is performed on the logarithm of the luminance and disparity images [49].

3.3 Bivariate GGD Modeling & Correlation calculation:

The multivariate GGD distribution of a random vector $\mathbf{x} \in \mathbb{R}^N$ is given by

$$p(\mathbf{x}|\mathbf{M}, \alpha, \beta) = \frac{1}{|\mathbf{M}|^{\frac{1}{2}}} g_{\alpha, \beta}(\mathbf{x}^T \mathbf{M}^{-1} \mathbf{x}), \quad (3.1)$$

$$g_{\alpha, \beta}(y) = \frac{\beta \Gamma(\frac{N}{2})}{(2^{\frac{1}{\beta}} \Pi \alpha)^{\frac{N}{2}} \Gamma(\frac{N}{2})} e^{-\frac{1}{2}(\frac{y}{\alpha})^\beta}, \quad (3.2)$$

where \mathbf{M} is an $N \times N$ symmetric scatter matrix, α is the scale parameter, β the shape parameter and $g_{\alpha, \beta}(\cdot)$ is the density generator. Its heavy-tailed and completely parameterized form makes it particularly attractive in modeling natural scenes. We propose a bivariate GGD model ($N = 2$) for the joint distribution of luminance and disparity wavelet coefficients of natural stereoscopic scenes.

We hypothesize that a bivariate GGD can accurately model the *joint* distribution of luminance and disparity subband coefficients. Fig. 3.3 illustrates a bivariate GGD model [50] fit for the joint histogram of luminance and disparity subband coefficients of a representative pristine natural scene from the LIVE Phase-II [24] database at the first scale and 0° orientation. Also shown are the

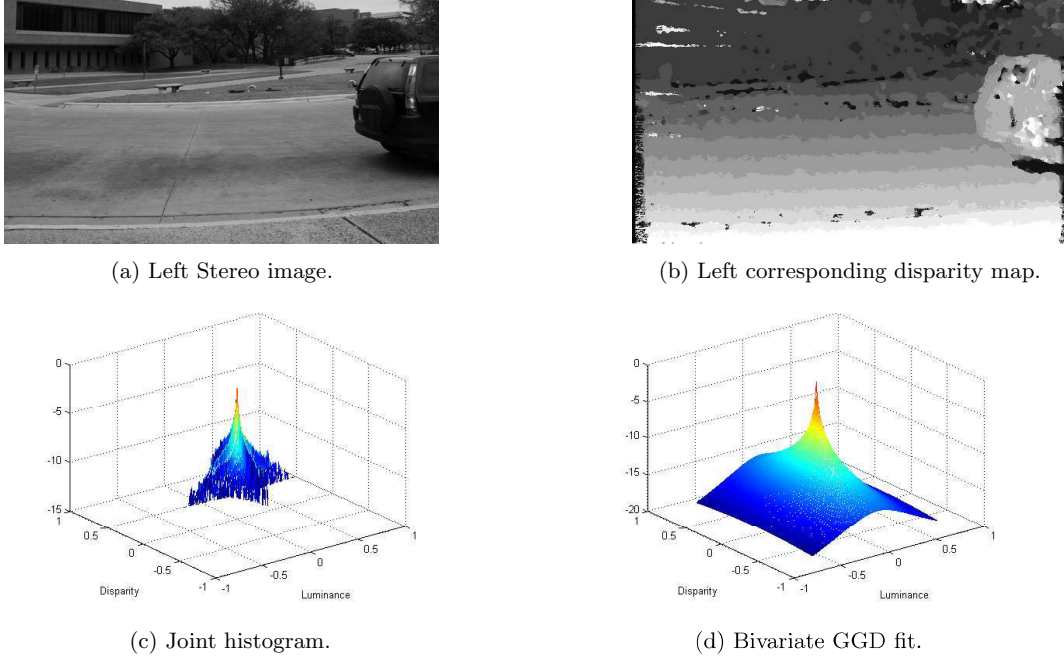
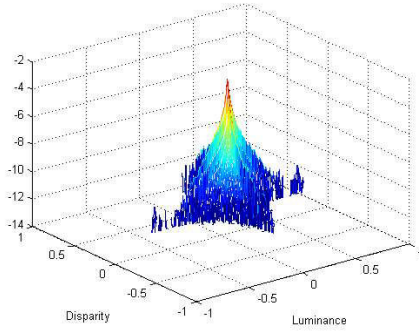


Figure 3.3: BGGD modeling. $\alpha = 9 \times 10^{-16}$, $\beta = 0.11$, $\chi = 5 \times 10^{-8}$.

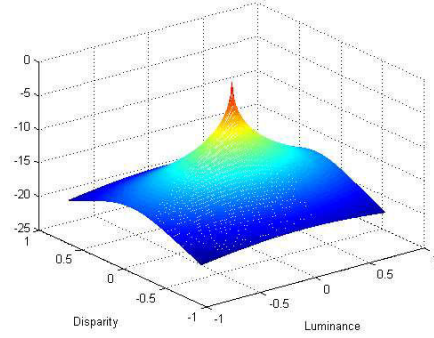
model parameters and the goodness of fit statistic χ . In Fig 3.4 shows the joint statistics variation on symmetric and asymmetric distortion of the same image in LIVE Phase-II database. These are also shown for first scale and 0° orientation. In our evaluation of the proposed BGGD model over the LIVE Phase-I, Phase-II, and the MICT [51] databases, the highest value for χ (indicating the worst fit) was found to be of the order of 10^{-7} . The low values of χ present excellent statistical support for our hypothesis.

We also found that the shape parameter β lies in the interval $[0.0014, 0.45]$, corresponding to heavy tailed distributions ($\beta = 1$ corresponds to the Gaussian distribution). This observation provides corroborative evidence for the findings of Liu et al. [49], and clearly highlights the non-Gaussian nature of the joint statistics. We now present the effect of distortion on the joint distribution of luminance and disparity subband coefficients. Fig. 3.5 shows the contour plots of the joint histogram of an undistorted image and several distorted versions of it (from the LIVE Phase-II database). These histograms have been computed at the first scale and 0° orientation. From the plots, it is clear that distortions affect the joint distribution.

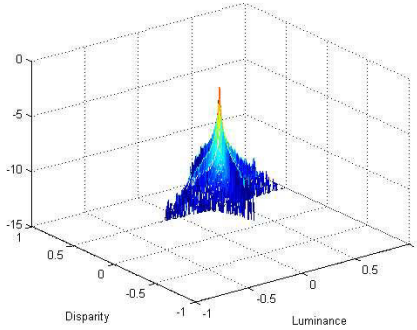
and therefore the best-fitting BGGD model parameters. Further, the change in the model parameter values relative to the reference is proportional to the amount of distortion. This observation is fundamental to our premise that the BGGD model parameters form good discriminatory features



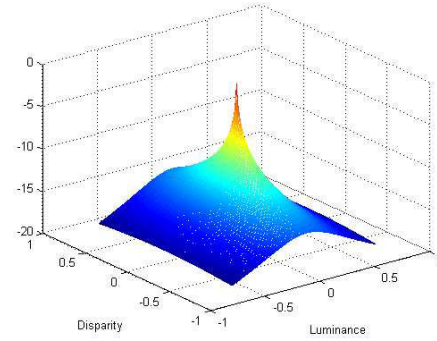
(a) Symmetrically distorted image joint histogram.



(b) BGGD fit for symmetrically distorted image.



(c) Asymmetrically distorted joint histogram.

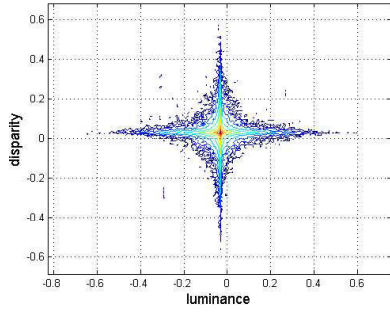


(d) BGGD fit for asymmetrically distorted image.

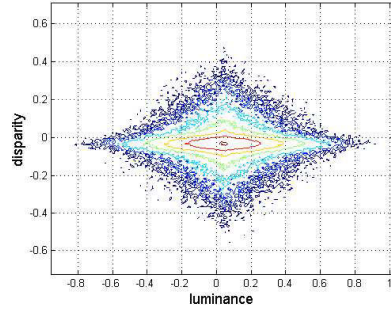
Figure 3.4: Illustration of BGGD variation over symmetric and asymmetric distortions.

for the quality assessment task. In addition to the BGGD model parameters, we rely on the result by Liu et al. [49] that there exists significant correlation between luminance and disparity subband coefficients.

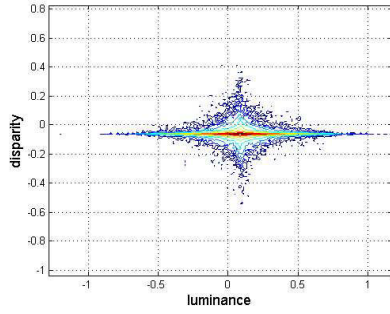
Fig. 3.6 shows the variation of correlation across the six orientations at the first scale for the same set of reference and distorted images used for the joint histogram computation. As with the joint histogram, Fig. 3.6 shows that correlation is also affected due to distortion in a manner that is proportional to the amount of distortion. Therefore, correlation also serves as a good feature for no-reference quality assessment. Our proposed NR stereo IQA algorithm is built on this statistically-inspired foundation and is presented in the following.



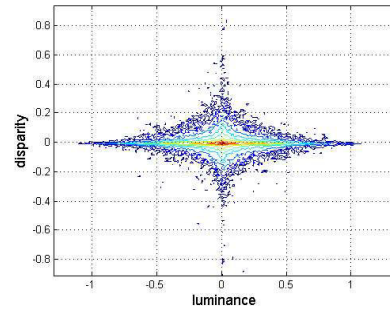
(a) Reference
 $(\alpha = 9.3 \times 10^{-16}, \beta = 0.11)$.



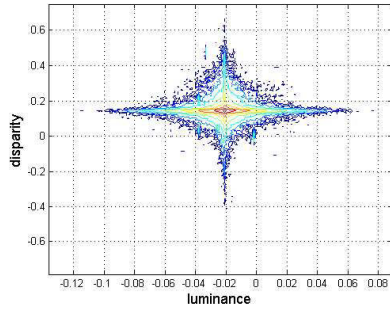
(b) AWGN
 $(\alpha = 1.1 \times 10^{-4}, \beta = 0.34)$.



(c) JPEG2000
 $(\alpha = 1.38 \times 10^{-10}, \beta = 0.15)$.



(d) JPEG
 $(\alpha = 4 \times 10^{-19}, \beta = 0.01)$.



(e) Gaussian blur
 $(\alpha = 1.5 \times 10^{-8}, \beta = 0.21)$.



(f) Fast fading
 $(\alpha = 2.23 \times 10^{-8}, \beta = 0.2)$.

Figure 3.5: Joint histogram contours for commonly occurring distortions.

Our analysis is carried out on oriented subbands of the luminance and disparity images. Specifically, we operate at 3 scales and 6 orientations resulting in 18 subbands. At each subband, the BGGD model parameters (α, β) and the correlation coefficient γ are computed. The features corresponding to the left image are subscripted by l and those corresponding to the right image are subscripted by r .

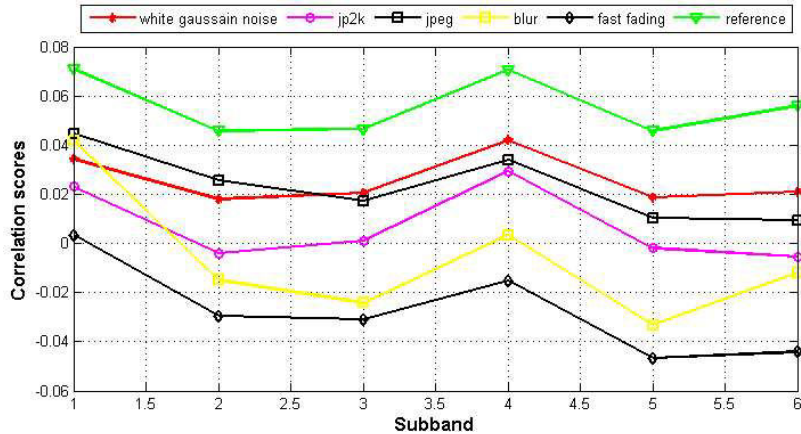


Figure 3.6: Correlation scores for LIVE Phase-II database.

This results in two feature vectors $\mathbf{f}_l = [\alpha_l^1 \dots \alpha_l^{18}; \beta_l^1 \dots \beta_l^{18}; \gamma_l^1 \dots \gamma_l^{18}]^T$, $\mathbf{f}_r = [\alpha_r^1 \dots \alpha_r^{18}; \beta_r^1 \dots \beta_r^{18}; \gamma_r^1 \dots \gamma_r^{18}]^T$, where the superscript corresponds to the subband index. Thus we end up with a pair of feature vectors $(\mathbf{f}_l, \mathbf{f}_r)$ containing a total of 108 elements per stereo image pair.

3.4 Feature Consolidation:

Once the feature vectors are extracted, the next step in our algorithm is supervised learning using DMOS scores as labels. While a supervised learning algorithm could be trained with the feature vector pair $(\mathbf{f}_l, \mathbf{f}_r)$, it is worthwhile exploring a reduction of the feature space dimensionality. Dimensionality reduction is further motivated by the fact that stereo images typically contain redundant information, and by implication, so would the pair $(\mathbf{f}_l, \mathbf{f}_r)$. One way to reduce dimensionality would be to consolidate the vector pair into a single vector. To achieve this goal, we propose the following consolidation strategy that is based on binocular rivalry [52] and the stimulus strength in the components of the stereo pair. Here in Table 3.1 showed the different set of weights that are used for combining \mathbf{f}_l and \mathbf{f}_r into a single feature vector \mathbf{f} per stereo image pair. Basically the third pooling combination produces very high correlation value but based on the convexity property we did not use that combination. So we relied on the fifth pooling strategy which produces the second highest correlation value. Those equations are shown below.

$$w_l^i = \frac{R_l^i R_{dl}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}, \quad (3.3)$$

$$w_r^i = \frac{R_r^i R_{dr}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}, \quad (3.4)$$

$$\mathbf{f} = \mathbf{w}_l \odot \mathbf{f}_l + \mathbf{w}_r \odot \mathbf{f}_r, \quad (3.5)$$

where w_l^i, w_r^i are the weights assigned to the i^{th} left and right subbands respectively, R_l^i, R_r^i are the RMS values of the i^{th} left and right luminance subbands respectively, R_{dl}^i, R_{dr}^i are the RMS values of the i^{th} left and right disparity subbands respectively. The vectors $\mathbf{w}_l = [w_l^1, \dots, w_l^{18}]^T$, $\mathbf{w}_r = [w_r^1, \dots, w_r^{18}]^T$ and \odot denotes element-wise product. The weights are chosen to be proportional to the stimulus strength to the respective eye and is inspired by binocular rivalry where the dominant stimulus strongly influences overall perception. The RMS values of the luminance and disparity subbands are chosen to be representatives of the respective subband strengths. The final feature vector \mathbf{f} is a convex combination of \mathbf{f}_l and \mathbf{f}_r .

Table 3.1: SROCC on LIVE PHASE-II database for different weighting factors.

<i>LeftWeight</i> (w_l^i)	<i>RightWeight</i> (w_r^i)	<i>SROCC</i>
$\frac{R_l^i}{R_l^i + R_r^i}$	$\frac{R_l^i}{R_l^i + R_r^i}$	0.7086
$\frac{R_{dl}^i}{R_{dl}^i + R_{dr}^i}$	$\frac{R_{dr}^i}{R_{dl}^i + R_{dr}^i}$	0.4891
$\frac{R_l^i R_{dl}^i}{(R_l^i + R_r^i)(R_{dl}^i + R_{dr}^i)}$	$\frac{R_r^i R_{dr}^i}{(R_l^i + R_r^i)(R_{dl}^i + R_{dr}^i)}$	0.8986
$\frac{1}{2} * (\frac{R_l^i}{R_l^i + R_r^i} + \frac{R_{dl}^i}{R_{dl}^i + R_{dr}^i})$	$\frac{1}{2} * (\frac{R_r^i}{R_l^i + R_r^i} + \frac{R_{dr}^i}{R_{dl}^i + R_{dr}^i})$	0.8214
$\frac{R_l^i R_{dl}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}$	$\frac{R_r^i R_{dr}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}$	0.8886

3.5 Supervised Learning and Regression:

The final stage of our algorithm is supervised learning using the feature vector \mathbf{f} of a stereo pair and its DMOS score as the feature label. We use a support vector machine (SVM) for training and estimate the score of test inputs using support vector regression (SVR). In our experiments with kernel choices, a radial basis function (RBF) kernel was shown to give the best performance. We report our results for this kernel choice.

Chapter 4

Results and Discussions:

The performance of our algorithm was evaluated on LIVE Phase-I [53], Phase-II [24],[25] and the MICT [51] stereo image databases. Both the LIVE databases consist of five distortion categories- White Gaussian, JP2K, JPEG, Gaussian Blur, Fast Fading and subjective quality as DMOS associated with each of its distorted stereo pair. LIVE Phase-I contains 20 pristine images and 365 distorted images. Each distortion type have 4 levels of distortion strengths.

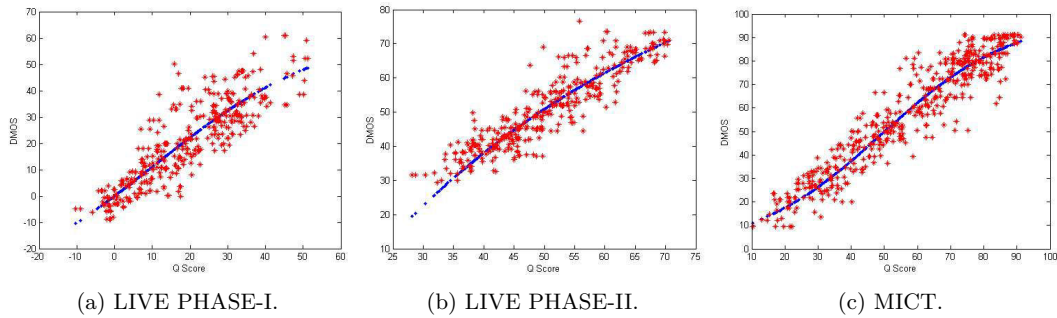


Figure 4.1: Scatter plots of StereoQUE versus DMOS for various stereo image databases.

LIVE Phase-II contains 8 reference images and 360 distorted images. Each distortion type have 9 levels of distortion strengths. The LIVE Phase-II images are also classified into symmetric and asymmetric distortions. The MICT database is composed of 13 reference images and JPEG distorted versions of them. For each database, 80% of the images were used to train the SVM and the rest were used for regression using the *libsvm* package [54]. For statistical consistency, we perform the training and testing 5000 times. We used random image assignment to the train and test sets for each iteration and ensured that the train and test sets did not overlap. We report results that are averaged over these 5000 iterations.

Table 4.1 and 4.2 show the performance of our algorithm StereoQUE on the LIVE Phase-I database.

Table 4.1: LCC(L) and SROCC(S) ON LIVE PHASE-I DATABASE.

Algorithm	LIVE Phase-I Database											
	WN		JP2K		JPEG		BLUR		FF		ALL	
	L	S	L	S	L	S	L	S	L	S	L	S
Benoit [12]	0.925	0.930	0.939	0.910	0.640	0.603	0.948	0.931	0.747	0.699	0.88	0.902
You [13]	0.941	0.940	0.877	0.860	0.487	0.439	0.919	0.882	0.730	0.588	0.881	0.878
Gorley [11]	0.796	0.741	0.485	0.015	0.312	0.567	0.852	0.750	0.364	0.366	0.451	0.142
Chen [24]	0.942	0.948	0.912	0.888	0.603	0.530	0.942	0.925	0.776	0.707	0.917	0.916
Hewage [55]	0.895	0.940	0.904	0.856	0.530	0.500	0.798	0.690	0.699	0.545	0.830	0.814
Akther [51]	0.904	0.914	0.905	0.866	0.729	0.675	0.617	0.555	0.503	0.640	0.626	0.383
Chen [25]	0.917	0.919	0.907	0.863	0.695	0.617	0.917	0.878	0.735	0.652	0.895	0.897
<i>StereoQUE</i>	0.919	0.910	0.938	0.917	0.806	0.782	0.881	0.865	0.758	0.666	0.917	0.911

Table 4.2: RMSE ON LIVE PHASE-I DATABASE.

Algorithm	WN	JP2K	JPEG	BLUR	FF	ALL
Benoit [12]	6.307	4.426	5.022	4.571	8.257	7.061
You [13]	5.621	6.206	5.709	5.679	8.492	7.746
Gorley [11]	10.197	11.323	6.211	7.562	11.569	14.635
Chen [24]	5.581	5.320	5.216	4.822	7.387	6.533
Hewage [55]	7.405	5.530	5.543	8.748	9.226	9.139
Akther [51]	7.092	5.483	4.273	11.387	9.332	14.827
Chen [25]	6.433	5.402	4.523	5.898	8.322	7.247
<i>StereoQUE</i>	6.664	4.943	4.391	6.938	9.317	6.598

Table 4.3: SROCC(S) ON LCC(L) ON MICT DATABASE.

Algorithm	L	S
Benoit [12]	0.910	0.902
You [13]	0.864	0.857
Chen [24]	0.864	0.862
Akther [51]	0.765	0.785
Chen [25]	0.913	0.904
<i>S3D-BLINQ Index</i> [45]	0.933	0.917
<i>StereoQUE</i>	0.935	0.936

Table 4.3 lists performance on the MICT database on the same set of images as in Chen et al.

[24]. Also shown are the performance numbers of the state-of-the-art methods.

Table 4.4 and 4.5 show the performance of our algorithm StereoQUE on the LIVE Phase-II database.

Table 4.4: LCC(L) and SROCC(S) ON LIVE PHASE-I DATABASE.

Algorithm	LIVE Phase-I Database											
	WN		JP2K		JPEG		BLUR		FF		ALL	
	L	S	L	S	L	S	L	S	L	S	L	S
Benoit [12]	0.925	0.923	0.939	0.751	0.640	0.867	0.948	0.455	0.747	0.773	0.88	0.728
You[13]	0.941	0.909	0.877	0.894	0.487	0.795	0.919	0.813	0.730	0.891	0.881	0.786
Gorley[11]	0.796	0.875	0.485	0.110	0.312	0.027	0.852	0.770	0.364	0.601	0.451	0.146
Chen[24]	0.942	0.940	0.912	0.814	0.603	0.843	0.942	0.908	0.776	0.884	0.917	0.889
Hewage[55]	0.895	0.880	0.904	0.598	0.530	0.736	0.798	0.028	0.699	0.684	0.830	0.501
<i>Akther</i> [51]	0.904	0.714	0.905	0.724	0.729	0.649	0.617	0.682	0.503	0.559	0.626	0.543
<i>Chen</i> [25]	0.917	0.950	0.907	0.867	0.695	0.867	0.917	0.900	0.735	0.933	0.895	0.880
<i>S3DBLINQ</i> [45]	0.953	0.946	0.847	0.845	0.888	0.818	0.968	0.903	0.944	0.899	0.913	0.905
<i>StereoQUE</i>	0.920	0.932	0.867	0.864	0.829	0.839	0.878	0.846	0.836	0.860	0.840	0.888

Table 4.5: RMSE ON LIVE PHASE-II DATASET.

Algorithm	WN	JP2K	JPEG	BLUR	FF	ALL
Benoit [12]	4.028	6.096	3.787	11.763	6.894	7.490
You [13]	4.396	4.186	4.086	8.649	4.649	6.772
Gorley [11]	5.202	9.113	6.940	4.988	8.155	9.675
Chen [24]	3.368	5.562	3.865	3.747	4.966	4.987
Hewage [55]	10.713	7.343	4.976	12.346	7.667	9.364
<i>Akther</i> [51]	7.416	6.819	4.535	8.450	8.505	9.294
<i>Chen</i> [25]	3.513	4.298	3.342	4.725	4.180	5.102
<i>S3DBLINQ</i> [45]	3.547	5.482	4.169	4.453	4.199	4.657
<i>StereoQUE</i>	4.325	5.087	4.756	6.662	6.519	7.279

The Fig. 4.1 shows the scatter plots for LIVE phase-I, LIVE phase-II and MICT databases. FR and RR methods are in normal font while the NR methods are *italicised*. From these results it is clear that StereoQUE performs competitively not just with NR algorithms but also with FR and RR algorithms. The strength of StereoQUE is highlighted by its consistently good performance on both symmetric and asymmetric distortions (Table 4.6). The performance of the algorithm can be explained by our choice of features and the weighting/consolidation strategy. The consolidation strategy is particularly effective when assessing the quality of asymmetrically distorted images.

Table 4.6: SROCC ON SYMMETRIC AND ASYMMETRIC DISTORTIONS IN LIVE PHASE-II DATABASE.

Algorithm	SYMM	ASYMM
Benoit [12]	0.860	0.671
You [13]	0.914	0.701
Gorley [11]	0.383	0.056
Chen [24]	0.923	0.842
Hewage [55]	0.656	0.496
<i>Akther</i> [51]	0.420	0.517
<i>Chen</i> [25]	0.918	0.834
<i>S3D-BLINQ</i> [45]	0.937	0.849
<i>StereoQUE</i>	0.857	0.872

Chapter 5

Conclusion

We presented a BGGD model for natural stereo scene statistics and an NR stereo IQA algorithm dubbed StereoQUE based on the model. StereoQUE was evaluated on popular stereo image databases and shown to perform competitively with state-of-the-art methods. Further, it delivers consistently high performance on symmetric and asymmetric distortions thereby demonstrating the effectiveness of the approach. While we demonstrated the efficacy of the proposed BGGD model using a NR stereo IQA application, the model could be useful in several other applications such as stereo correspondence, stereo image restoration etc. In future, we intend to explore some of these applications.

References

- [1] R. Tenniswood, L. Safonova, and M. Drake. 3Ds Effect on a Films Box Office and Profitability 2010.
- [2] Motion Picture Association of America. Theatrical Market Statistics 2013 2013.
- [3] H. Sheikh, A. Bovik, and L. Cormack. No-reference quality assessment using natural scene statistics: JPEG2000. *Image Processing, IEEE Transactions on* 14, (2005) 1918 –1927.
- [4] R. Soundararajan and A. Bovik. RRED Indices: Reduced Reference Entropic Differencing for Image Quality Assessment. *Image Processing, IEEE Transactions on* 21, (2012) 517 –526.
- [5] A. Moorthy and A. Bovik. Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality. *Image Processing, IEEE Transactions on* 20, (2011) 3350 –3364.
- [6] M. A. Saad, A. C. Bovik, and C. Charrier. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *Image Processing, IEEE Transactions on* 21, (2012) 3339–3352.
- [7] A. Mittal, R. Soundararajan, and A. C. Bovik. Making a completely blind image quality analyzer. *Signal Processing Letters, IEEE* 20, (2013) 209–212.
- [8] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision* 47, (2002) 7–42.
- [9] M.-J. Chen. Visual perception and quality of distorted stereoscopic 3D images .
- [10] P. Campisi, P. Le Callet, and E. Marini. Stereoscopic images quality assessment. In Proceedings of 15th European Signal Processing Conference (EUSIPCO07). 2007 .
- [11] P. Gorley and N. Holliman. Stereoscopic image quality metrics and compression. In Electronic Imaging 2008. International Society for Optics and Photonics, 2008 680,305–680,305.
- [12] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau. Quality assessment of stereoscopic images. *EURASIP journal on image and video processing* 2008.
- [13] J. You, L. Xing, A. Perkis, and X. Wang. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis. In Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA. 2010 .

- [14] J. Wang and Z. Wang. PERCEPTUAL QUALITY OF ASYMMETRICALLY DISTORTED STEREOSCOPIC IMAGES: THE ROLE OF IMAGE DISTORTION TYPES .
- [15] R. Bensalma and M.-C. Larabi. Towards a perceptual quality metric for color stereo images. In Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, 2010 4037–4040.
- [16] R. Bensalma and C. Iarabi. A stereoscopic quality metric based on binocular perception. In Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on. IEEE, 2010 41–44.
- [17] R. Bensalma and M.-C. Larabi. A perceptual metric for stereoscopic image quality assessment based on the binocular energy. *Multidimensional Systems and Signal Processing* 24, (2013) 281–316.
- [18] S. Ryu, D. H. Kim, and K. Sohn. Stereoscopic image quality metric based on binocular perception model. In Image Processing (ICIP), 2012 19th IEEE International Conference on. IEEE, 2012 609–612.
- [19] K. Li, S. Wang, Q. Jiang, and F. Shao. Objective quality assessment for stereoscopic images based on structure-texture decomposition. *WSEAS Transactions on Computers* 13.
- [20] L. Shen, J. Yang, and Z. Zhang. Quality assessment of stereo images with stereo vision. In Image and Signal Processing, 2009. CISP'09. 2nd International Congress on. IEEE, 2009 1–4.
- [21] C. Hu, F. Shao, G. Jiang, M. Yu, F. Li, and Z. Peng. Quality assessment for stereoscopic images by distortion separation. *Journal of Software* 9, (2014) 37–43.
- [22] W. Zhou, G. Jiang, M. Yu, Z. Wang, Z. Peng, and F. Shao. Reduced reference stereoscopic image quality assessment using digital watermarking. *Computers & Electrical Engineering* 40, (2014) 104–116.
- [23] A. Maalouf and M.-C. Larabi. CYCLOP: A stereo color image quality assessment metric. In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011 1161–1164.
- [24] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, and A. C. Bovik. Full-reference quality assessment of stereopairs accounting for rivalry. *Signal Processing: Image Communication* 28, (2013) 1143–1155.
- [25] M.-J. Chen, L. K. Cormack, and A. C. Bovik. No-reference quality assessment of natural stereopairs. *Image Processing, IEEE Transactions on* 22, (2013) 3379–3391.
- [26] A. Chetouani. Full reference image quality metric for stereo images based on Cyclopean image computation and neural fusion. In Visual Communications and Image Processing Conference, 2014 IEEE. IEEE, 2014 109–112.
- [27] S. A. Fezza and M.-C. Larabi. No-reference perceptual blur metric for stereoscopic images. In 3D Imaging (IC3D), 2014 International Conference on. IEEE, 2014 1–8.
- [28] E. Y. Lam and J. W. Goodman. A mathematical analysis of the DCT coefficient distributions for images. *Image Processing, IEEE Transactions on* 9, (2000) 1661–1666.

- [29] M. J. Wainwright and E. P. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. In NIPS. Citeseer, 1999 855–861.
- [30] A. K. Moorthy and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *Image Processing, IEEE Transactions on* 20, (2011) 3350–3364.
- [31] A. Mittal, A. K. Moorthy, and A. C. Bovik. Blind/referenceless image spatial quality evaluator. In Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on. IEEE, 2011 723–727.
- [32] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *Image Processing, IEEE Transactions on* 21, (2012) 4695–4708.
- [33] H. R. Sheikh, A. C. Bovik, and L. Cormack. No-reference quality assessment using natural scene statistics: JPEG2000. *Image Processing, IEEE Transactions on* 14, (2005) 1918–1927.
- [34] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam. Objective video quality assessment methods: A classification, review, and performance comparison. *Broadcasting, IEEE Transactions on* 57, (2011) 165–182.
- [35] S. Yasakethu, C. T. Hewage, W. A. C. Fernando, and A. M. Kondo. Quality analysis for 3D video using 2D video quality models. *Consumer Electronics, IEEE Transactions on* 54, (2008) 1969–1976.
- [36] A. Mittal, A. K. Moorthy, J. Ghosh, and A. C. Bovik. Algorithmic assessment of 3D quality of experience for images and videos. In Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE. IEEE, 2011 338–343.
- [37] J. Huang, A. B. Lee, and D. B. Mumford. Statistics of range images. Institute of Electrical and Electronics Engineers, 2000 .
- [38] O. Wulf and B. Wagner. Fast 3D scanning methods for laser measurement systems. In International conference on control systems and computer science (CSCS14). 2003 2–5.
- [39] Y. Liu, A. C. Bovik, and L. K. Cormack. Disparity statistics in natural scenes. *Journal of Vision* 8, (2008) 19.
- [40] B. Potetz and T. S. Lee. Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *JOSA A* 20, (2003) 1292–1303.
- [41] Y. Liu, L. K. Cormack, and A. C. Bovik. Statistical modeling of 3-D natural scenes with application to bayesian stereopsis. *Image Processing, IEEE Transactions on* 20, (2011) 2515–2530.
- [42] Y. Liu, L. K. Cormack, and A. C. Bovik. Luminance, disparity, and range statistics in 3d natural scenes. In IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2009 72,401G–72,401G.
- [43] K. Kokkinakis and A. K. Nandi. Exponent parameter estimation for generalized Gaussian probability density functions with application to speech modeling. *Signal Processing* 85, (2005) 1852–1858.

- [44] C.-C. Su, A. C. Bovik, and L. K. Cormack. Natural scene statistics of color and range. In Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011 257–260.
- [45] C.-C. Su, L. Cormack, and A. Bovik. Oriented Correlation Models of Distorted Natural Images with Application to Natural Stereopair Quality Evaluation .
- [46] <http://www.cns.nyu.edu/~eero/steerpyr/>.
- [47] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In Image Processing, International Conference on, volume 3. IEEE Computer Society, 1995 3444–3444.
- [48] A. Karasaridis and E. Simoncelli. A filter design technique for steerable pyramid image transforms. In Acoustics, Speech, and Signal Processing, IEEE International Conference on, volume 4. IEEE, 1996 2387–2390.
- [49] Y. Liu, L. Cormack, and A. Bovik. Statistical Modeling of 3-D Natural Scenes With Application to Bayesian Stereopsis. *Image Processing, IEEE Transactions on* 20, (2011) 2515 –2530.
- [50] C.-C. Su, L. K. Cormack, and A. C. Bovik. Bivariate statistical modeling of color and range in natural scenes. In IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2014 90,141G–90,141G.
- [51] R. Akhter, Z. P. Sazzad, Y. Horita, and J. Baltes. No-reference stereoscopic image quality assessment. In IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2010 75,240T–75,240T.
- [52] W. J. Levelt. On binocular rivalry. Ph.D. thesis, Van Gorcum Assen 1965.
- [53] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik. Subjective evaluation of stereoscopic image quality. *Signal Processing: Image Communication* 28, (2013) 870–883.
- [54] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)* 2, (2011) 27.
- [55] C. T. Hewage and M. G. Martini. Reduced-reference quality metric for 3D depth map transmission. In 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2010. IEEE, 2010 1–4.