Oriented Correlation Models of Distorted Natural Images With Application to Natural Stereopair Quality Evaluation

Che-Chun Su, Lawrence K. Cormack, and Alan C. Bovik, *Fellow, IEEE*

***Abstract*— In recent years, bandpass statistical models of natural, photographic images of the world have been used with great success to solve highly diverse problems involving image representation, image repair, image quality assessment (IQA), and image compression. One missing element has been a reliable and generic model of spatial image correlation that reflects the distributions of oriented and relatively oriented spatial structures. We have developed such a model for bandpass pristine images and have generalized it here to also capture the spatial correlation structure of bandpass distorted images. The model applies well to both luminance and depth images. As a demonstration of the usefulness of the generalized model, we develop a new no-reference stereoscopic/3D IQA framework, dubbed stereoscopic/3D blind image naturalness quality index, which utilizes both univariate and generalized bivariate natural scene statistics (NSS) models. We first validate the robustness and effectiveness of these novel bivariate and correlation NSS features extracted from distorted stereopairs, then demonstrate that they are predictive of distortion severity. Our experimental results show that the resulting 3D image quality predictor based in part on the new model outperforms state-of-the-art full- and no-reference 3D IQA algorithms on both symmetrically and asymmetrically distorted stereoscopic image pairs.**

***Index Terms*— Natural scene statistics (NSS), image quality assessment (IQA), 3D, stereoscopic image pair, bivariate model.**

I. Introduction

NATURAL scene statistics (NSS) have proven to be important ingredients towards both understanding the evolution of human vision systems (HVS) [3], [4] and solving diverse problems in image and video processing, e.g., representation, compression, denoising, and quality assessment [5]. In particular, much advancement has been made in predicting image and video perceptual quality by exploiting reliable NSS models of bandpass pristine and distorted images. Deviations from the regularity of natural statistics, when quantified appropriately, enable the design of algorithms capable of assessing the perceptual quality of an image without the need for any reference image. For example, Moorthy *et al.* [6], [7] deployed summary NSS features derived from luminance wavelet coefficients to develop a two-stage framework for 2D image quality assessment: distortion-identification followed by distortion-specific quality prediction models. In [8], Saad *et al.* proposed a pragmatic approach to no-reference 2D image quality assessment using a small number of features computed from an NSS model of block DCT coefficients. Recently, Mittal *et al.* proposed no-reference [9] and "completely blind" [10] 2D image quality assessment algorithms utilizing spatial-domain NSS models. These no-reference algorithms are able to deliver highly competitive performance relative to commonly used 2D full-reference quality metrics, e.g., MS-SSIM [11].

There has also been work conducted on exploring 3D natural scene statistics and their applications [12]–[14]. Potetz et al. [12] examined the relationships between luminance and range over multiple scales and applied their results to shape-from-shading problems. Liu et al. [13] explored the statistical relationships between luminance and disparity in the wavelet domain, and applied the derived models to a Bayesian stereo algorithm. Recently, Su *et al.* [14] proposed reliable statistical models for both marginal and conditional distributions of luminance/chrominance and disparity in natural images, and used these models to significantly improve a chromatic Bayesian stereo algorithm.

However, most natural scene statistical models proposed in the literature have been characterized only by univariate distributions of bandpass images, while higher-order dependencies between spatially adjacent bandpass responses to natural images are not yet well-understood. Some early work has been conducted on analyzing and modeling joint/bivariate relationships between sub-band natural image coefficients. For example, Portilla *et. al* [15], [16] proposed a Markov statistical descriptor of texture images using a set of parametric constraints on pairs of complex wavelet coefficients at adjacent spatial locations, orientations, and scales. In [17] and [18], the authors found that the coefficients of orthonormal wavelet decompositions of natural images are fairly well-decorrelated; however, they are not independent. Liu *et al.* [19] measured inter- and intra-scale dependencies between image wavelet coefficients using mutual information. In [20], Sendur *et al.* considered image wavelet coefficients and their parents (at adjacent coarser scale locations), and proposed a circularly symmetric bivariate distribution to model their dependencies.

Manuscript received March 27, 2014; revised October 6, 2014; accepted February 4, 2015. Date of publication March 4, 2015; date of current version March 23, 2015. This work was supported by the National Science Foundation under Grant IIS-0917175 and Grant IIS-1116656. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Adrian G. Bors.

C.-C. Su and A. C. Bovik are with the Center for Perceptual Systems, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: ccsu@utexas.edu; bovik@ece.utexas.edu).

L. K. Cormack is with the Center for Perceptual Systems, Department of Psychology, Institute for Neuroscience, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: cormack@utexas.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2015.2409558

Po *et al.* [21] applied a 2D contourlet transform to natural images, and examined both the marginal and joint distributions. They measured the dependencies between image contourlet coefficients using mutual information, and proposed a hidden Markov tree (HMT) image model with Gaussian mixtures that can capture inter-location, inter-scale, and inter-direction dependencies. Although these bivariate and joint statistics have been demonstrated to be able to provide improvements over univariate NSS models on practical image applications, e.g., image denoting [20], [22], texture synthesis [15], [16], and texture retrieval [21], none of them has offered a closed-form quantitative model of the bivariate correlations of bandpass natural images. If available, such a closed-form expression could be invaluable for analyzing spatial statistical image behavior and for formulating easily expressed and computed optimized solutions to a wide variety of image processing problems.

Specifically, there has been a missing element of reliable and generic modeling of the statistical distributions of oriented and relatively oriented spatial structures in natural images. Aiming to fill this gap, we developed in [1] a bivariate and a closed-form correlation NSS models of spatially adjacent pixels of bandpass pristine natural images, and demonstrated that these models also apply well to both natural spatial luminance/chrominance and spatial depth information [2]. The first contribution that we make here is a more general closed-form NSS model than the one in [2], that is able to capture the correlation structure of bandpass *distorted* images via a more versatile set of model parameters. Our second contribution is inspired by recent successful NSS-based 2D image and video quality assessment algorithms. We propose a no-reference stereoscopic/3D (S3D) image quality metric, dubbed Stereoscopic/3D BLind Image Naturalness Quality (S3D-BLINQ) Index, which deploys these novel NSS models to quantify and predict perceived distortions on stereoscopically viewed binocular/3D image pairs. S3D-BLINQ is the first quality assessment algorithm that utilizes bivariate and correlation NSS models, and it achieves superior performance over previous NSS-based IQA algorithms when predicting the perceptual quality of S3D images. In the context of image IQA, the most important reason for a bivariate model is that both images and distortions contain strong bivariate dependencies. Those of images create structure, while those of distortions degrade the structure, thus impairing the image appearance.

We proceed as follows. The new bivariate and correlation NSS models are introduced in Section II, while Section III reviews previous work on S3D image quality assessment. The proposed S3D-BLINQ Index framework is detailed in Section IV. Validation of the new NSS models on 3D data is also included in Section IV. Section V presents a performance evaluation of S3D-BLINQ Index. Finally, Section VI concludes with possible directions of future work.

II. Bivariate and Closed-Form Correlation NSS Models

It has been demonstrated that the distributions of divisive-normalized bandpass responses to natural, photographic images possess strong Gaussian-like regularities. Natural scene statistical models of this type have been widely deployed in numerous image and video applications with success [7], [22], [23]. However, less progress has been made on the development of accurate and general statistical models of the higher-order dependencies that exist between spatially neighboring bandpass image responses.

Towards developing such models of the bivariate, correlation statistics of oriented, bandpass images, we utilize the steerable pyramid decomposition of images, which is an over-complete wavelet transform that allows for increased orientation selectivity [24]. The wavelet transform model is motivated by the fact that its space-scale-orientation decomposition broadly resembles the band-pass filtering that occurs in area V1 of primary visual cortex [25], [26]. After applying the multi-scale, multi-orientation decomposition, the perceptually significant process of divisive normalization is applied to the image wavelet coefficients of all sub-bands [18]. The divisive normalization transform (DNT) used here is implemented as follows [27]:

$$S(x_i, y_i) = \frac{s(x_i, y_i)}{\sqrt{c_s + \mathbf{s_g}^\top \mathbf{s_g}}} = \frac{s(x_i, y_i)}{\sqrt{c_s + \sum_j g(x_j, y_j) s(x_j, y_j)^2}} \tag{1}$$

where (x_i, y_i) are spatial coordinates, s represents the sub-band wavelet coefficients, S represents the coefficients after DNT, and c_s is a semi-saturation constant. The sum occurs over neighborhood pixels indexed by j, where $\{g(x_j, y_j)\}$ is a Gaussian weighting function.

Previous work by others and ourselves [1], [28] showed that the empirical joint histograms of spatially adjacent sub-band coefficients of natural images can be well fitted by multivariate generalized Gaussian distribution (MGGD) models, which include both multivariate Gaussian and Laplace distributions as special cases. The probability density function of a multivariate generalized Gaussian distribution (MGGD) is defined as:

$$p(\mathbf{x}; \mathbf{M}, \alpha_b, \beta_b) = \frac{1}{|\mathbf{M}|^{\frac{1}{2}}} g_{\alpha_b, \beta_b}(\mathbf{x}^\top \mathbf{M}^{-1} \mathbf{x}) \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{M}$ is an $N \times N$ symmetric scatter matrix, α_b and β_b are the scale and shape parameters, respectively, and $g_{\alpha_b, \beta_b}(\cdot)$ is a density generator defined as:

$$g_{\alpha_b, \beta_b}(y) = \frac{\beta_b \Gamma(\frac{N}{2})}{(2^{\frac{1}{\beta_b}} \pi \alpha_b)^{\frac{N}{2}} \Gamma(\frac{N}{2\beta_b})} e^{-\frac{1}{2}(\frac{y}{\alpha_b})^{\beta_b}} \tag{3}$$

where $y \in \mathbb{R}^+$. Note that when $\beta_b = 0.5$, Eq. (2) becomes a multivariate Laplacian distribution, and when $\beta_b = 1$, Eq. (2) corresponds to a multivariate Gaussian distribution. Moreover, when $\beta_b \to \infty$, the MGGD converges to a multivariate uniform distribution. In our recent work, we have become interested in modeling the bivariate empirical histograms of horizontally adjacent sub-band coefficients of both 2D and 3D (cyclopean) images using the bivariate generalized Gaussian distribution (BGGD) with $N = 2$. The parameters of the BGGD can be obtained on the bandpass

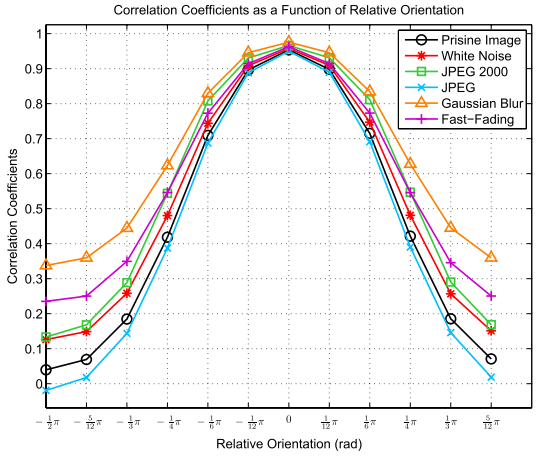

Fig. 1. The correlation coefficients between spatially adjacent sub-band responses as a function of relative orientation.

coefficients of images using the maximum likelihood estimator (MLE) algorithm described in [2].

By examining the fitted BGGD models of pristine bandpass images, i.e., not artificially subjected to, or containing any noticeable distortions, we have found orientation dependencies between spatially adjacent sub-band image coefficients [1]. In particular, when the spatial relationship of adjacent responses, e.g., horizontal, matches the sub-band orientation, e.g., $\frac{1}{2}\pi$, the joint distribution of the responses becomes peaky and extremely elliptical. On the other hand, when the spatial relationship and the sub-band orientation approach orthogonality, e.g., horizontal vs. 0 (rad), the joint distribution becomes nearly circular and more Gaussian-like.

We can seek to quantitatively capture this statistical dependency on relative orientation from an interesting, systematic and potentially useful perspective by directly modeling the correlation entries embedded in the scatter matrix $\mathbf{M}$ of the BGGD model. We define relative orientation to be the difference between the sub-band tuning orientation and the spatial orientation of adjacent responses. For example, if the sub-band tuning orientation is 0 (rad), and the pixels are horizontally adjacent, i.e., the spatial orientation is $\frac{1}{2}\pi$, meaning that they are orthogonal, then the corresponding relative orientation is equal to $0-\frac{1}{2}\pi = -\frac{1}{2}\pi$. We have observed that the correlation coefficient between horizontally adjacent bandpass responses reaches a maximum when the sub-band orientation is $\frac{1}{2}\pi$, whereas horizontal neighboring responses become nearly uncorrelated as the sub-band tuning orientation approaches 0 (rad) and π. In addition, the correlation takes symmetric values around sub-band orientation equal to $\frac{\pi}{2}$ and is periodic with respect to all possible relative orientations between the pairs of horizontally adjacent sub-band responses and their orientation tuning. In particular, we observed that the correlation coefficients, when plotted as a function of relative orientation, have a roughly sinusoidal shape but with narrowed lobes. We also observed that when images are distorted by commonly occurring impairments, such as blur, noise, compression or those found in publicly available image quality databases, the correlation plots as a function of relative orientation also take roughly sinusoidal shapes but with different degrees of lobe narrowness. Figure 1 illustrates an example of the correlation coefficient plots as a function of relative orientation for both pristine and impaired images afflicted by different types of distortions from the LIVE IQA Database [29]. We can clearly see that all curves possess sinusoidal-like shapes, but with different degrees of lobe narrowness depending on the distortion type. This strongly suggested to us that a successful model of the sub-band correlations could enable us to capture the effects of distortions on sub-band correlations.

We have found that the periodic relative orientation dependency of the correlation coefficients is well modeled by a closed-form exponentiated cosine function given by:

$$\rho = f(\theta_1, \theta_2) = A\left[\frac{1+\cos\left(2(\theta_2-\theta_1)\right)}{2}\right]^{\gamma} + c_e \quad (4)$$

$$= A\left[\cos\left(\theta_2-\theta_1\right)\right]^{2\gamma} + c_e \quad (5)$$

where ρ is the correlation coefficient between spatially adjacent bandpass responses, θ_1 and θ_2 represent spatial and sub-band tuning orientations, respectively, A is the amplitude, γ is the exponent, and c_e is the offset. Note that the correlation coefficient ρ is period-π in relative orientation and reaches maximum value when $\theta_2 - \theta_1 = k\pi, k \in \mathbb{Z}$. This three-parameter model is a more general form that extends the simple one-parameter correlation NSS model proposed in [1], which holds well for pristine images, but fails to capture the distinct characteristics of impaired images afflicted by different types of distortions, as shown in Figure 1. The model parameters are estimated via non-linear least squares using the Levenberg-Marquardt algorithm [30].

Since we developed this more general relative orientation correlation model in the context of our work on natural 3D statistics and their applications, we validate the generalized BGGD and exponentiated cosine models on perceptually relevant "cyclopean" images, formed from the left and right images of a stereopair, as discussed in Section IV-B. In particular, with a more versatile set of parameters, $[A, \gamma, c_e]^{\top}$, the general closed-form correlation model developed in this paper is able to describe and quantify both pristine and distorted S3D image pairs, as demonstrated later in Sections IV-F and V.

III. Previous Work on S3D IQA

As with other digital visual media [31], the quantity of S3D images and videos that are delivered by the cinema, television, and online entertainment industries on a daily basis for human consumption has been growing dramatically over the past few years. According to recent theatrical market statistics gathered by the Motion Picture Association of America (MPAA) [32], the proportion of cinema screens that are 3D has reached 35% worldwide, and approximate half of all moviegoers viewed at least one 3D movie in 2012. As Hollywood director James Cameron, who directed and produced Avatar, stated in an interview with BBC news in Aug. 2013 [33], "All forms of entertainment will eventually be 3D, because that's how we see the world." In fact, the wave of 3D has not been limited to the entertainment industry. Given the development of greatly improved acquisition and display technologies, S3D images and videos can provide natural and versatile visual presentations for numerous applications,

including robot navigation [34], remote education [35], anatomical exploration [36], therapeutic treatment [37], and so forth. As these large volumes of S3D data are making their way to consumers and other users, a variety of issues have arisen regarding efficient compression and reliable transmission of S3D content, especially when being transmitted over already-stressed wireless networks. At every stage of capture, compression, storage, and transmission, it is desirable to maximize the quality of the final visual experience, and in this regard, incorporating principles of the human perception of S3D quality is of importance [38]–[40].

As an interesting and important application of the new parametric correlation model [1], we develop an automatic no-reference (NR) S3D image quality model that is able to automatically predict the perceptual quality of distorted S3D images, without benefit of any reference signal, making it useful for practical applications. Models that attempt to solve the S3D IQA problem may be distinguished by whether they utilize computed or measured depth/disparity information from the stereoscopic pairs. Thus, the simplest S3D IQA models apply off-the-shelf 2D IQA algorithms to both left and right stereo images, then aggregate the two quality scores to form a final prediction of the quality of the fused stereopair. Both full-reference 2D models, e.g., PSNR [41], SSIM [42], and MS-SSIM [11], and 2D NR models, e.g., DIIVINE [7], BLIINDS [8], and BRISQUE [9], can be used in this way. Yasakethu *et al.* [43] applied a variety of 2D IQA algorithms to the left and right views independently, then averaged them to obtain S3D quality scores, achieving fairly good correlation with both perceived image and depth quality. Gorley *et al.* [44] reported a full-reference S3D IQA model that they found preferable to the PSNR for controlling practical S3D image compression rates. Recently, there has been increased emphasis on developing S3D IQA models that utilize the encoding of depth/disparity stimuli from the natural environment by modeling cortical neurons with disparity-tuned receptive fields [45]–[47]. Benoit *et al.* [48] predict the quality of S3D image pairs using the disparity information computed from off-the-shelf stereo algorithms [49], [50]. Recent studies have demonstrated the importance of depth/disparity for understanding perceptual S3D image quality. For example, Chen *et al.* [51] showed that when viewing S3D image pairs, subjects tend to agree on perceived image quality, but have more diverse opinions on their sensations of depth.

Although the depth/disparity information extracted from S3D image pairs does affect the perceptual quality of viewed stereoscopic images, the question of how best to exploit this information remains incompletely answered. You *et al.* [52] attempted to quantify the degradation of disparity information by applying 2D IQA algorithms on the disparity maps computed from both reference and distorted left-right image pairs [50]. Disparity information can also be used indirectly to bolster an S3D IQA algorithm. For example, Sazzad *et al.* [53] utilized disparity information to design an NR IQA algorithm to predict the quality of both symmetrically and asymmetrically JPEG-coded stereo image pairs.

However, the ultimate goal of an S3D IQA algorithm is to form predictions of the quality of the ultimate cyclopean image [54] formed within an observer's mind when a left-right image pair is stereoscopically presented. Towards this end, several recent researchers have attempted to evaluate perceptual quality by synthesizing an intermediate image that more-or-less agrees with cyclopean perception. Maalouf *et al.* [55] proposed a reduced-reference quality metric that compares the sensitivity coefficients [56] extracted from the two cyclopean images synthesized from the reference and distorted stereopairs. Chen *et al.* [57] proposed a full-reference S3D IQA algorithm exploiting a perceptually synthesized cyclopean image to account for binocular rivalry. In [58], the authors extended this framework to create a no-reference model using 2D and 3D natural scene statistical features extracted from S3D image pairs.

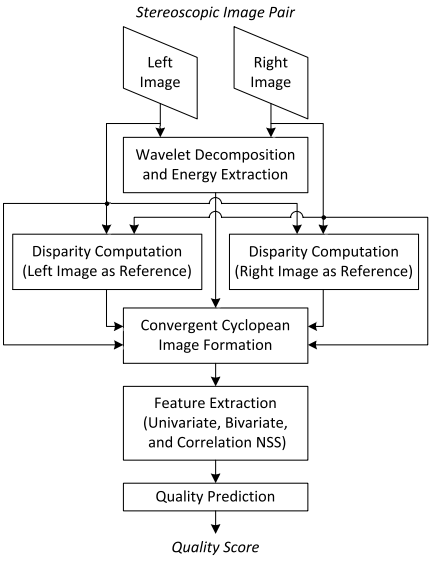

Fig. 2. S3D-BLINQ Index framework.

IV. Natural Stereopair Quality Index

Inspired by the success of 2D image/video quality assessment algorithms that use 2D natural scene statistics, we have developed a no-reference natural stereopair quality index (S3D-BLINQ Index), which achieves high correlations with human subjective judgments of S3D image quality using the novel bivariate and correlation NSS models explained in Section II, along with a symmetrically defined model of the cyclopean image, to extract robust, effective features for S3D image quality prediction.

A. Framework Overview

Figure 2 diagrams the processing flow of the proposed S3D-BLINQ Index framework. S3D-BLINQ Index first forms a convergent cyclopean image using disparity maps computed from both left- and right-view images as references. Next, both spatial-domain and wavelet-domain univariate NSS features, as well as the bivariate and correlation NSS features introduced in Section II, are extracted from the convergent cyclopean image. Finally, the perceptual quality of S3D images is predicted by mapping the extracted features to human opinion scores. Each step is detailed in the following sub-sections.

B. Convergent Cyclopean Image Formation

The synthesized cyclopean image model adopted in prior work was formed by fixing the left-view image, then warping the right-view image onto the corresponding left-view image coordinates using a disparity map computed by a stereo algorithm using the left-view image as reference [59]. Of course, the binocular human vision system does not form a cyclopean 3D precept using the left- or right-view image as reference; instead, it synergistically fuses the two views into an intermediate image on a coordinate frame defined relative to fixation [54]. A biased cyclopean image model may fail to capture certain parts of the 3D world accurately, e.g., near depth discontinuities, when the S3D image pair is asymmetrically distorted. To address this distinction, we deploy a more complete and hence perceptually relevant model of the convergent cyclopean image using a general parallel-viewing geometric model of practical S3D image display scenarios.

A simple linear model proposed by Levelt [60], which remains the most widely-used cyclopean image model, explains the formation of the perceived cyclopean image I_C experienced when a stereoscopic stimulus is presented as a linear combination of neural representations of the stimuli I_L and I_R to the left and right eyes:

$$I_C = w_L \cdot I_L + w_R \cdot I_R \tag{6}$$

where w_L and w_R are the weighting coefficients on the corresponding stimuli with the constraint $w_L + w_R = 1$. Assuming the disparity map is computed using the left image as reference to match the right image, the synthesized cyclopean image may be generated as [55], [57], [58]:

$$\begin{aligned} I_C(x, y) = {} & w_L(x, y) \cdot I_L(x, y) \\ & + w_R(x - D_L(x, y), y) \cdot I_R(x - D_L(x, y), y) \end{aligned} \tag{7}$$

where (x, y) are spatial pixel coordinates, I_L and I_R are left and right image representations (e.g., luminance or bandpass luminance), respectively, w_L and w_R are the weighting maps, and D_L is the disparity map computed by matching elements of I_R to those in I_L, i.e., using I_L as reference. However, this synthesized cyclopean image may fail to capture certain characteristics that may affect the perception of an asymmetrically distorted stereopair. For example, consider two asymmetrically distorted stereopairs having the same content, one of them a pristine left-view image with a distorted right-view image, and the other a pristine right-view image with a similarly impaired left-view image. A synthesized cyclopean image arrived at using Eq. (7) will possibly generate very different results on these two asymmetrically distorted stereoscopic image pairs, whereas we would ordinarily expect the perception of these stereopairs to be similar. To address this possible bias in the synthesized cyclopean image, we propose a more perceptually relevant model that we call the convergent cyclopean image. This model has the virtue of, even for symmetrically distorted or undistorted images, providing a larger and collectively consistent set of constraints on this difficult, ill-posed problem.

Without loss of generality and towards practical applications, we adopt a simple parallel-viewing geometry to generate a convergent cyclopean image given a stereoscopic image pair, as illustrated in Fig. 3. In principle, a convergent cyclopean image may be formed as a linear combination of both the disparity-compensated left- and right-view images, yielding a coherent, symmetrically defined representation. The simplest approach is to model the convergent disparity as equal to half of the right-to-left or left-to-right disparities computed by a canonical stereo algorithm. Then, the convergent cyclopean image I_{CC} becomes:

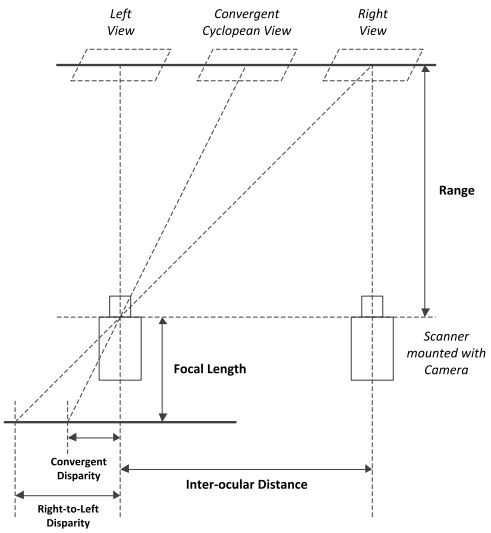

Fig. 3. Parallel-viewing geometry for generating the convergent cyclopean image from the left and right images.

$$\begin{aligned} I_{CC}(x, y) = {} & w_L(x + D'_R(x, y), y) \cdot I_L(x + D'_R(x, y), y) \\ & + w_R(x - D'_L(x, y), y) \cdot I_R(x - D'_L(x, y), y) \end{aligned} \tag{8}$$

where $D'_R(x, y) = \frac{D_R(x,y)}{2}$ and $D'_L(x, y) = \frac{D_L(x,y)}{2}$ are convergent disparity maps computed using the right and left images as references, respectively, and D_R and D_L are the canonical disparity maps computed using the right and left images as references, respectively.

The stimulus strengths, i.e., the weighting maps w_L and w_R in (8), are modeled as the sum of the energies of wavelet coefficients computed using a steerable pyramid, followed by a DNT taken across sub-bands as described in Section II. As a result, the convergent cyclopean image given a stereoscopic image pair is formed as

$$\begin{aligned} I_{CC}(x, y) = {} & \frac{E_L[x + D'_R(x, y), y]}{E_U(x, y)} \cdot I_L[x + D'_R(x, y), y] \\ & + \frac{E_R[x - D'_L(x, y), y]}{E_U(x, y)} \cdot I_R[x - D'_L(x, y), y] \end{aligned} \tag{9}$$

$$E_L(x, y) = \frac{\sum_{k=1}^{K} S^2_{L_k}(x, y)}{K} \tag{10}$$

$$E_R(x, y) = \frac{\sum_{k=1}^{K} S^2_{R_k}(x, y)}{K} \tag{11}$$

$$E_U(x, y) = E_L[x + D'_R(x, y), y] + E_R[x - D'_L(x, y), y] \tag{12}$$

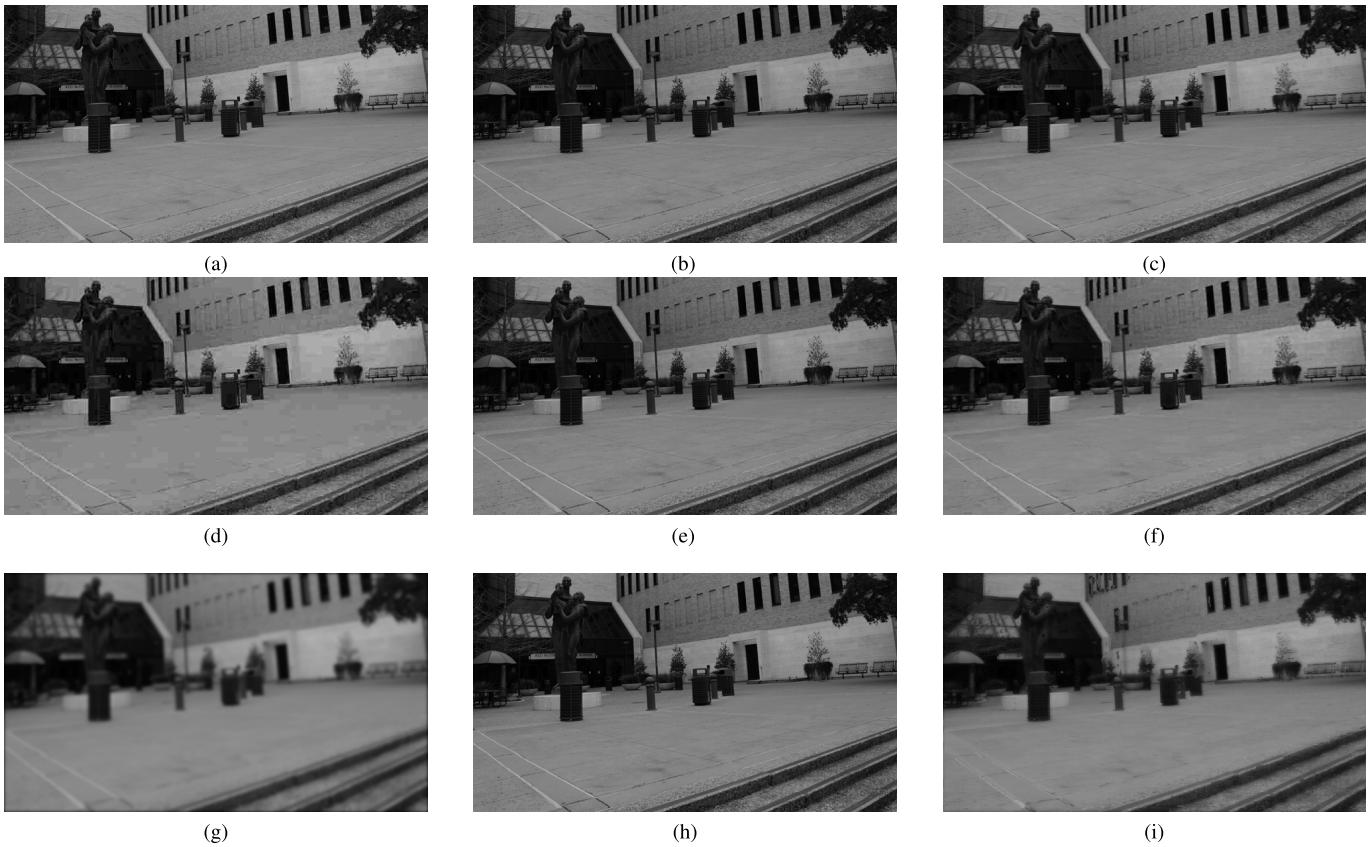

Fig. 4. Examples of convergent cyclopean images formed by pristine, asymmetrically JPEG-compressed, and asymmetrically Gaussian blurred stereopairs. (a) Pristine right view. (b) Pristine left view. (c) Pristine convergent cyclopean image. (d) JPEG-compressed right view. (e) Pristine left view. (f) Asymmetrically JPEG-compressed convergent cyclopean image. (g) Gaussian blurred right view. (h) Pristine left view. (i) Asymmetrically Gaussian blurred convergent cyclopean image.

where E_L and E_R are the left and right energy maps, S_{L_k} and S_{R_k} are the left and right sub-band coefficients at sub-band k, K is the number of sub-bands, and E_U serves to achieve unit-sum weighting as in (6).

Figure 4 shows some examples of convergent cyclopean images formed by pristine and asymmetrically distorted stereopairs. All left-view images are pristine, while the right-view image examples include pristine, JPEG-compressed, and Gaussian blurred images. When a human observes such stereopairs, the deep question arises whether, upon successfully free-fusing left- and right-view images, one would be able to construct a clear 3D percept by some process of distortion masking or whether the view might have an appearance of heightened distortion, perhaps owing to a facilitation of the asymmetric impairments. For example, blocky distortions on the a JPEG-compressed right-view image might be more or less apparent on 3D viewing, while Gaussian blur of one of the images may not reduce the sharpness of the overall 3D percept. Such questions have been explored deeply in the experiments reported in [51] and in more focused studies in the references cited there. In any case, the generated convergent cyclopean images render a means of capturing these perceptual effects.

C. Spatial-Domain Univariate NSS Feature Extraction

It has been demonstrated that natural scene statistics (NSS) models provide powerful and robust tools for gauging human judgments of visual distortions on 2D images and videos [7], [8], [23]. Early on, Ruderman [61] showed that a simple non-linear operation of local mean subtraction and divisive variance normalization on natural image luminances results in a decorrelated, Gaussianized 'contrast' image. This spatial-domain NSS model has been extended in various ways and successfully deployed in no-reference 2D quality assessment algorithms [9], [10] which deliver highly competitive performance relative to top-performing full-reference metrics. We utilize a similar decomposition to extract spatial-domain univariate features from each convergent cyclopean image. First, the luminance of the convergent cyclopean image, I_{CC}, is transformed as:

$$\hat{I}_{CC}(x, y) = \frac{I_{CC}(x, y) - \mu(x, y)}{\sigma(x, y) + c} \tag{13}$$

where (x, y) are spatial pixel coordinates, μ and σ are locally weighted spatial means and standard deviations computed using a Gaussian window superimposed over the spatial neighborhood, and $c = 1$ is a constant that ensures stability. To capture a broader spectrum of distortion statistics than Gaussian on convergent cyclopean images, we use the univariate generalized Gaussian distribution (GGD) model to fit the empirical histograms of the contrast images $\hat{I}_{CC}$. The probability density function of a univariate GGD with zero mean is:

$$p(x; \alpha_u, \beta_u) = \frac{\beta_u}{2\alpha_u \Gamma(\frac{1}{\beta_u})} e^{-(\frac{|x|}{\alpha_u})^{\beta_u}} \tag{14}$$

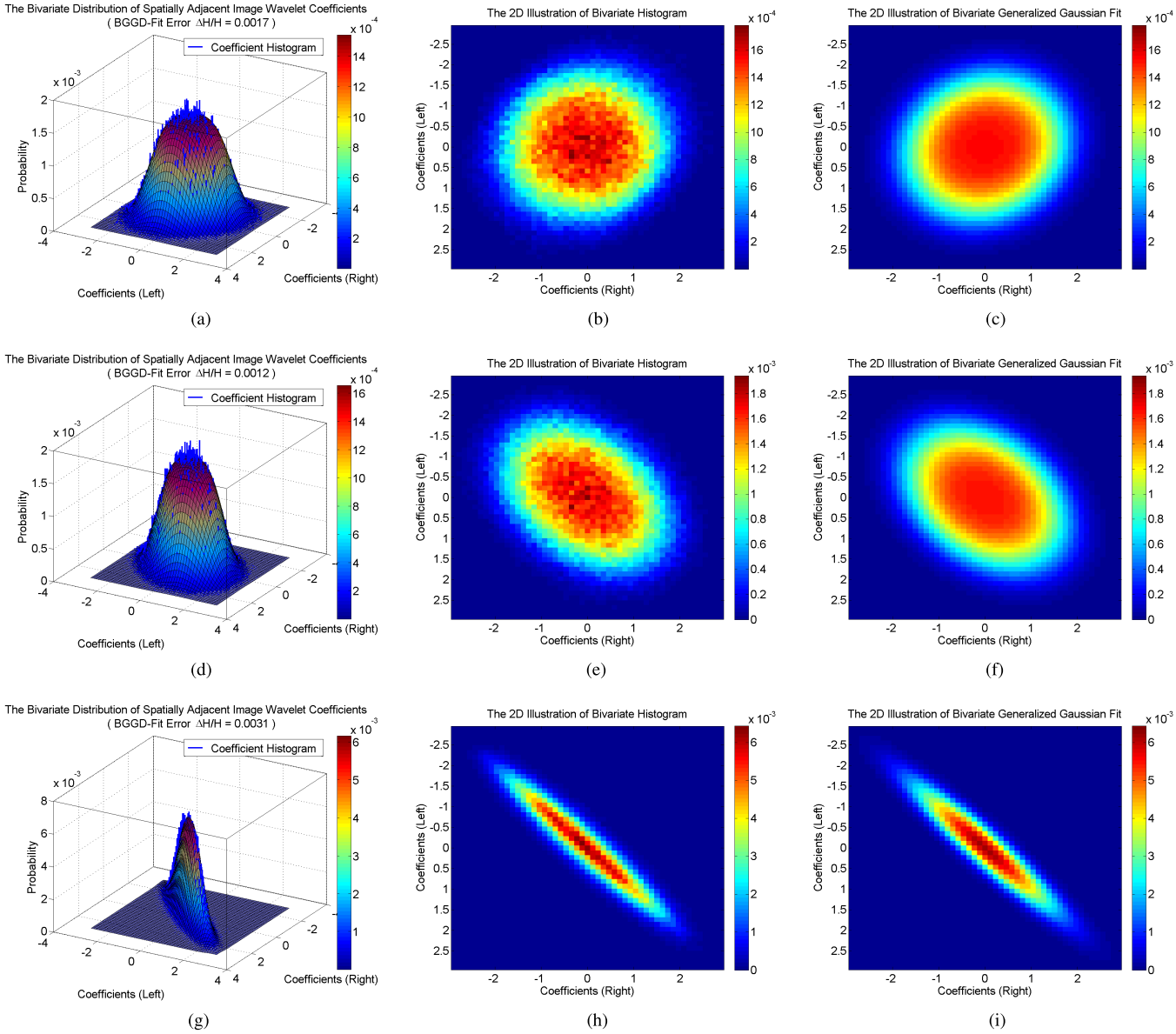

Fig. 5. Joint histograms of horizontally adjacent bandpass coefficients from a pristine convergent cyclopean image and corresponding BGGD fits at the finest scale along different sub-band tuning orientations. Top row: orientation $= 0$, middle row: orientation $= \frac{\pi}{4}$, and bottom row: orientation $= \frac{\pi}{2}$. (a) BGGD fit and histogram. (b) 2D illustration of histogram. (c) 2D illustration of BGGD fit. (d) BGGD fit and histogram. (e) 2D illustration of histogram. (f) 2D illustration of BGGD fit. (g) BGGD fit and histogram. (h) 2D illustration of histogram. (i) 2D illustration of BGGD fit.

where $\Gamma(\cdot)$ is the ordinary gamma function and α_u and β_u are scale and shape parameters, respectively.

We use the moment-matching based approach proposed in [62] to estimate the parameters of the univariate GGD fit. The two extracted univariate GGD parameters, $[\alpha_u, \beta_u]^\top$, are deployed as spatial-domain "quality-aware" features.

Since the normalizing operation (13) is isotropic, we also model the statistical relationships between neighboring pixels along different orientations using the very general univariate asymmetric generalized Gaussian distribution (AGGD) [9], [63]. Specifically, we fit the empirical histograms of pairwise products of adjacent (cardinal and diagonal) coefficients of the convergent cyclopean contrast image, $\hat{I}_{CC}$, using the multi-parameter univariate AGGD probability density function with zero mean:

$$p(x;\alpha_l,\alpha_r,\beta_a)=\begin{cases}\dfrac{\beta_a}{(\alpha_l+\alpha_r)\Gamma(\frac{1}{\beta_a})}e^{-(\frac{-x}{\alpha_l})^{\beta_a}}, & x<0\\ \dfrac{\beta_a}{(\alpha_l+\alpha_r)\Gamma(\frac{1}{\beta_a})}e^{-(\frac{x}{\alpha_r})^{\beta_a}}, & x\leq 0\end{cases} \tag{15}$$

where β_a is the shape parameter, and α_l and α_r are scale parameters that control the spread of the AGGD to the left and right of the origin. The parameters of the AGGD fits are also estimated using the moment-matching based approach in [62]. All three AGGD parameters, $[\alpha_l, \alpha_r, \beta_a]^\top$, extracted from each S3D convergent cyclopean contrast image are employed as spatial-domain quality features.

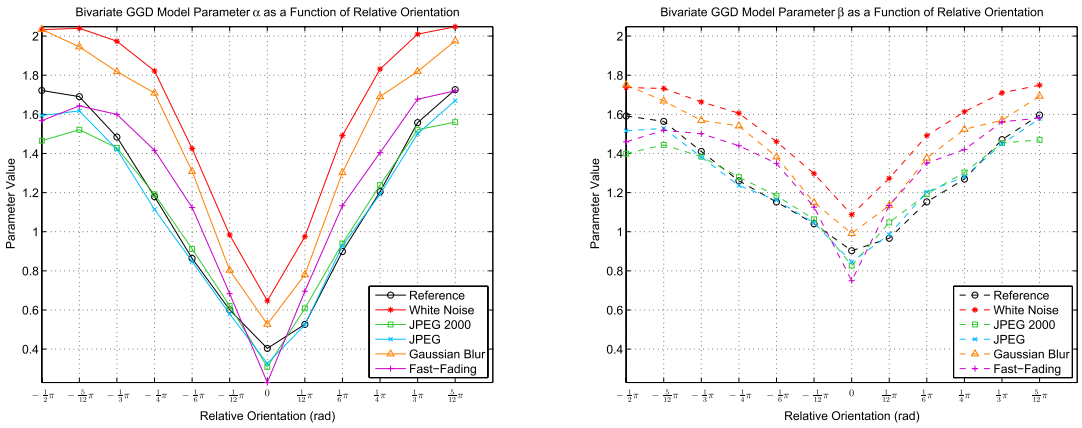

Fig. 6. Plots of the two BGGD model parameters as a function of relative orientation from pristine and distorted convergent cyclopean images. (a) scale parameter α_b. (b) shape parameter β_b.

D. Wavelet-Domain Univariate NSS Feature Extraction

Considerable work has focused on modeling the statistics of natural images using multi-scale, multi-orientation transforms, e.g., Gabor filters, wavelets, etc [4], [64]. Success has also been attained by utilizing transform-domain NSS models to create 2D image and video quality assessment models [7], [8], [23]. In these kinds of IQA models, perceptually relevant transform-domain features are computed via area V1-like band-pass filtering. Likewise, we process the convergent cyclopean image I_{CC} using the same steerable pyramid wavelet decomposition as described in Section II, followed by the divisive normalization transform (1). We again use the univariate GGD to fit the empirical histograms of these sub-band coefficients using (14). The two resulting GGD parameters from each sub-band, scale and shape, are included in the wavelet-domain feature set.

E. Bivariate Density and Correlation NSS Feature Extraction

We employ the novel bivariate density and correlation NSS models introduced in Section II to extract wavelet-domain features from the convergent cyclopean image I_{CC}, using the same steerable pyramid transform and DNT (1) as described in Section IV-D. First, we validate the efficacy of these novel NSS models by fitting the BGGD to the empirical histograms of spatially adjacent bandpass coefficients after DNT of convergent cyclopean images formed by undistorted stereopairs. Figure 5 shows the joint empirical histograms of horizontally adjacent bandpass coefficients and the corresponding BGGD fits at the finest scale over several different orientations $(0, \frac{\pi}{4}, \frac{\pi}{2})$. These bivariate histograms were obtained by first binning both responses at spatially adjacent locations. For example, given responses at locations (x, y) and $(x + 1, y)$ on a 2D grid, count the number of occurrences within each grid entry, then compute the height of each grid entry by normalizing its occurrence by the sum of occurrences from all entries. In the plots, blue bars represent the true histograms while colored meshes represent the fits. Clearly, the bivariate joint distributions of horizontally adjacent wavelet coefficients are well modeled by BGGD. The 2D figures, which are iso-probability contour maps of the joint distributions, also illustrate the high accuracy of the fits obtained using the BGGD models. The most important observation here is that the shapes and heights of the joint distributions both vary with sub-band orientation. This matches our early findings of BGGD fits on pristine 2D images [1], i.e., there exist much higher dependencies between spatially adjacent pixels after being decomposed by bandpass filters when the orientations are similar.

To portray a clear picture of this relative orientation dependency, we plot the two BGGD model parameters α_b and β_b as a function of relative orientation on convergent cyclopean images afflicted by different types of distortion in Fig. 6. Clearly there is a strong relative orientation dependency of both parameters, each reaching minimum value when the spatial orientation matches the sub-band tuning orientation, i.e., $x_2 - x_1 = 0$. Thus, horizontally adjacent sub-band coefficients share the highest correlation when their tuning orientation is $\frac{\pi}{2}$, with the correlation declining away from $\frac{\pi}{2}$. These plots also show that different types of distortion cause different degrees of relative orientation dependency, which we shall also exploit to develop additional quality-aware correlation features. Specifically, the scale and shape parameters $[\alpha_b, \beta_b]^\top$ are deployed as bivariate NSS quality features.

As discussed in Section II, the relative orientation dependency of spatially adjacent bandpass responses of images is reflected by a systematic behavior of the correlation coefficients. Figure 7 shows exponentiated cosine function fits to the correlation coefficient plots of horizontally adjacent wavelet coefficients, as a function of relative orientation, on convergent cyclopean images afflicted by several different types of distortion. The appearances of the exponentiated cosine model fits are quite close to the orientation-dependent correlation coefficient curves, in agreement with the low mean squared errors (MSE), where the model parameters were estimated with non-linear least squares using the Levenberg-Marquardt algorithm [30]. Note that the model varies with distortion type. We also conducted a multivariate statistical hypothesis test as we describe in the next subsection that illustrates the discriminative power of the exponentiated

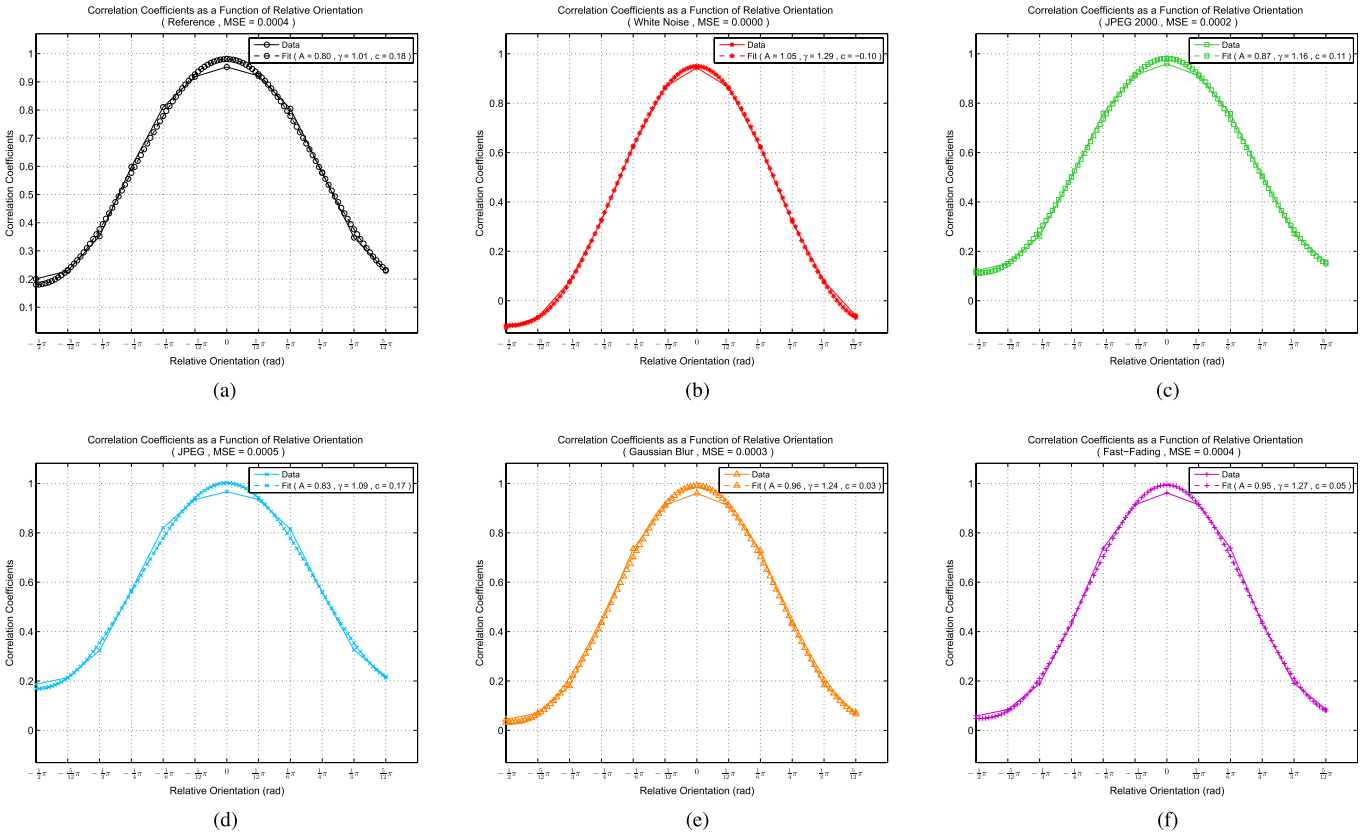

Fig. 7. Exponentiated cosine fits to the curves of correlation coefficients between spatially adjacent wavelet coefficients as a function of relative orientation for distorted convergent cyclopean images. (a) Pristine. (b) White Noise. (c) JPEG 2000. (d) JPEG. (e) Gaussian Blur. (f) Fast-Fading.

cosine model. Based on these amicable results, we include the exponentiated cosine model parameters, $[A, \gamma, c_e]^\top$, as distortion sensitive IQA features.

F. Validation of the Closed-Form Correlation NSS Model

The exponentiated cosine model of the correlation coefficients between spatially adjacent sub-band responses was shown to be a reliable model of natural photographic images in [1]. Here we validate the applicability of the more general exponentiated cosine model (5) for convergent cyclopean images formed from a diversity of distorted stereoscopic image pairs taken from the LIVE 3D Image Quality Database Phase II [65]. Specifically, we performed a statistical hypothesis test on the three parameters of the exponentiated cosine model across all computed convergent cyclopean images for each type of distortion, including the pristine convergent cyclopean images. First, we computed the orientation-dependent correlation coefficient curves at a particular scale for all N_D images afflicted by distortion type D in the database, and obtained corresponding exponentiated cosine fits for each distorted cyclopean image. Denote each exponentiated cosine fit by a vector $\mathbf{x} = [A, \gamma, c]^\top \in \mathbb{R}^3$, where A, γ, and c are the three model parameters, amplitude, exponent, and offset, respectively. For brevity, we tabulate the results only for the finest scale; however, we obtained similar results for other scales as well (see [66]). Next, we computed the mean model parameter vector across all convergent cyclopean images having distortion type D, denoted $\overline{\mathbf{x}}_D = \sum_{i=1}^{N_D} \mathbf{x}_{D_i}$, where $\mathbf{x}_{D_i} = [A_{D_i}, \gamma_{D_i}, c_{D_i}]^\top$. Then, we applied a two-sample multivariate t-test to determine whether the null hypothesis H_0, i.e., the two mean vectors $\overline{\mathbf{x}}_{D_1}$ and $\overline{\mathbf{x}}_{D_2}$ of two different distortion types D_1 and D_2 are equal, is supported. If the null hypothesis is supported, then the two exponentiated cosine models of distortion types D_1 and D_2 are statistically identical. However, if H_0 is rejected, we can conclude that distortion types D_1 and D_2 possess significantly different exponentiated cosine models. In particular, we computed Hotelling's two-sample T-squared statistic T^2, which generalizes the Student's two-sample t statistic:

$$T^2 = \frac{N_{D_1} N_{D_2}}{N_{D_1} + N_{D_2}} (\overline{\mathbf{x}}_{D_1} - \overline{\mathbf{x}}_{D_2})^\top \mathbf{S_p}^{-1} (\overline{\mathbf{x}}_{D_1} - \overline{\mathbf{x}}_{D_2}) \quad (16)$$

where $\mathbf{S_p}$ is an unbiased estimate of the pooled covariance matrix:

$$\mathbf{S_p} = \frac{\sum_{D_{1_i}=1}^{N_{D_1}} (\mathbf{x}_{D_{1_i}} - \overline{\mathbf{x}}_{D_1})(\mathbf{x}_{D_{1_i}} - \overline{\mathbf{x}}_{D_1})^\top}{N_{D_1} + N_{D_2} - 2} + \frac{\sum_{D_{2_i}=1}^{N_{D_2}} (\mathbf{x}_{D_{2_i}} - \overline{\mathbf{x}}_{D_2})(\mathbf{x}_{D_{2_i}} - \overline{\mathbf{x}}_{D_2})^\top}{N_{D_1} + N_{D_2} - 2} \quad (17)$$

Finally, T^2 can be related to the F-distribution as:

$$\frac{N_{D_1} + N_{D_2} - P - 1}{(N_{D_1} + N_{D_2} - 2)P} T^2 \sim F_{P, N_{D_1} + N_{D_2} - P - 1} \quad (18)$$

where P is the dimension of $x_{D_{1_i}}$ and $x_{D_{2_i}}$.

TABLE I

STATISTICAL HYPOTHESIS TEST RESULTS WITH COMPUTED p-VALUES. A VALUE OF '-1' IN THE PARENTHESIS INDICATES THAT THE NULL HYPOTHESIS IS REJECTED WHILE A VALUE OF '$+1$' INDICATES THAT IT IS SUPPORTED

	Pristine	WN	JP2K	JPEG	Blur	FF
Pristine	-	0 (-1)	9.12×10^{-08} (-1)	8.24×10^{-08} (-1)	1.11×10^{-16} (-1)	3.21×10^{-09} (-1)
WN	0 (-1)	-	0 (-1)	0 (-1)	0 (-1)	0 (-1)
JP2K	9.12×10^{-08} (-1)	0 (-1)	-	0 (-1)	3.05×10^{-13} (-1)	3.10×10^{-03} (-1)
JPEG	8.24×10^{-08} (-1)	0 (-1)	0 (-1)	-	4.74×10^{-11} (-1)	2.99×10^{-14} (-1)
Blur	1.11×10^{-16} (-1)	0 (-1)	3.05×10^{-13} (-1)	4.74×10^{-11} (-1)	-	3.92×10^{-06} (-1)
FF	3.21×10^{-09} (-1)	0 (-1)	3.10×10^{-03} (-1)	2.99×10^{-14} (-1)	3.92×10^{-06} (-1)	-

Therefore, we are able to compute the p-value of our null hypothesis test as:

$$p = 1 - C_{F_{P,N_{D_1}+N_{D_2}-P-1}}\left(\frac{N_{D_1}+N_{D_2}-P-1}{(N_{D_1}+N_{D_2}-2)P}T^2\right) \quad (19)$$

where $C_{F_{P,N_{D_1}+N_{D_2}-P-1}}$ represents the cumulative distribution function of the F-distribution. Note that in our test, $P = 3$ and $N_{D_1} = N_{D_2} = 72$. We repeated this null hypothesis test on all different pairs of distortion types, including the pristine, to examine the robustness of the exponentiated cosine model.

Table I shows the test results on all distortion type pairs, where each entry records the computed p-values along with whether the null hypothesis between the row and column distortion types was rejected or not. Note that since Hotelling's two-sample T-squared statistic T^2 commutes, the entries in Tables I are diagonally symmetric. Clearly, at the one-sided significance level $\alpha = 0.05$, all hypothesis tests were rejected with p-value $< \alpha$, indicating that every distortion type, including the pristine, can be characterized by a distinct exponentiated cosine model. These results not only support the validity of the proposed exponentiated cosine model of the relative orientation-dependent correlation coefficients between spatially adjacent sub-band responses across different types of distorted convergent cyclopean images, but also substantiate its relevance for developing quality-predictive features on S3D images.

G. Quality Prediction

The last step in the proposed S3D-BLINQ Index framework is to predict the quality of stereoscopic image pairs using the aforedescribed NSS features extracted from the corresponding convergent cyclopean images. A mapping is learned from the feature space to human subjective quality scores using a regression model. The proposed framework is generically amenable to the application of any kind of regressor. The implementation of S3D-BLINQ Index described here utilizes a support vector machine (SVM) regressor (SVR) [67] using multiple train-test sequences as described in the next section. SVR is generally noted for being able to handle high dimensional data [68], and has also been used to create a variety of 2D IQA models [7], [69]. We utilized the LIBSVM package [70] to implement the SVR model with a radial basis function (RBF) kernel, whose parameter was estimated by cross-validation [71].

V. EXPERIMENTAL RESULTS

A. Performance Evaluation of S3D-BLINQ Index

In the previous section, we motivated and developed a statistics-based S3D IQA framework, dubbed S3D-BLINQ Index, that incorporates old and new models of the univariate and bivariate statistics of natural photographic S3D images. We next evaluated the efficacy of the new S3D IQA model against state-of-the-art 2D and S3D IQA models on the LIVE 3D Image Quality Database Phase II [65], which consists of both symmetrically and asymmetrically distorted stereopairs. There are five different types of distortions in the LIVE 3D Image Quality Database Phase II: JPEG and JPEG2000 (JP2K) compression, additive white Gaussian noise (WN), Gaussian blur (Blur), and a Rayleigh fast-fading channel distortion (FF). The severities of each of the degradations vary significantly and with good overall perceptual separations between distortion levels. For full-reference algorithms, we used all of the available reference and distorted stereopairs, while for no-reference algorithms, we divided the whole database into 80% training and 20% testing subsets at each train-test iteration so that there was no overlap between training and testing image content. This train-test procedure was repeated 1000 times to ensure that there was no bias introduced due to the image content used for training. We report the median performance across all iterations as the final performance score.

We computed two correlation measures, Spearman's rank-order correlation coefficient (SROCC) and Pearson's linear correlation coefficient (LCC), along with the root-mean-squared error (RMSE) between the predicted quality scores and the recorded subjective opinion scores (DMOS) to evaluate the performance of the quality assessment models [72]. Since both LCC and RMSE are accuracy measures, all algorithm scores were passed through a logistic non-linear function to map to DMOS space before computing LCC and RMSE [72]. The SROCC, LCC, and RMSE values of the tested 2D and S3D IQA models evaluated on the LIVE 3D Image Quality Database Phase II are summarized and tabulated in Tables II–IV. We reported the performance for each distortion type, as well as an "overall" metric score computed using all test image pairs across different distortion types. Higher values of the two correlation measures, SROCC and LCC, and lower values of RMSE indicate better performance.

From Tables II and III, it can be seen that the highest attained performance by any "pure 2D" IQA model on the LIVE S3D image pairs reached about 0.8 correlation

TABLE II
COMPARISON (SROCC) OF DIFFERENT 2D AND 3D IMAGE QUALITY ASSESSMENT MODELS ON DIFFERENT DISTORTION TYPES IN THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

	Algorithm†	WN	JP2K	JPEG	Blur	FF	Overall
2D	PSNR	0.919	0.597	0.491	0.690	0.730	0.665
	SSIM [42]	0.922	0.704	0.678	0.838	0.834	0.792
	MS-SSIM [11]	0.946	0.798	0.847	0.801	0.833	0.777
	BRISQUE [9]	0.846	0.593	0.769	0.862	0.935	0.770
3D	Benoit [48]	0.923	0.751	0.867	0.455	0.773	0.728
	You [52]	0.909	0.894	0.795	0.813	0.891	0.786
	Yasakethu [43]	0.880	0.598	0.736	0.028	0.684	0.501
	Cyclopean MS-SSIM [57]	0.940	0.814	0.843	0.908	0.884	0.889
	Sazzad [53]	0.714	0.724	0.649	0.682	0.559	0.543
	Chen [58]	0.950	0.867	0.867	0.900	0.933	0.880
	S3D-BLINQ Index	0.946	0.845	0.818	0.903	0.899	0.905

† *Italics* indicate no-reference IQA models. Others are full-reference IQA models.

TABLE III
COMPARISON (LCC) OF DIFFERENT 2D AND 3D IMAGE QUALITY ASSESSMENT MODELS ON DIFFERENT DISTORTION TYPES IN THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

	Algorithm†	WN	JP2K	JPEG	Blur	FF	Overall
2D	PSNR	0.917	0.627	0.459	0.706	0.762	0.680
	SSIM [42]	0.928	0.723	0.650	0.848	0.858	0.802
	MS-SSIM [11]	0.950	0.820	0.856	0.798	0.842	0.783
	BRISQUE [9]	0.845	0.681	0.795	0.951	0.931	0.782
3D	Benoit [48]	0.926	0.784	0.853	0.535	0.807	0.748
	You [52]	0.912	0.905	0.830	0.784	0.915	0.800
	Yasakethu [43]	0.891	0.664	0.734	0.450	0.746	0.558
	Cyclopean MS-SSIM [57]	0.957	0.834	0.862	0.963	0.901	0.900
	Sazzad [53]	0.722	0.776	0.786	0.795	0.674	0.568
	Chen [58]	0.947	0.899	0.901	0.941	0.932	0.895
	S3D-BLINQ Index	0.953	0.847	0.888	0.968	0.944	0.913

† *Italics* indicate no-reference IQA models. Others are full-reference IQA models.

TABLE IV
COMPARISON (RMSE) OF DIFFERENT 2D AND 3D IMAGE QUALITY ASSESSMENT MODELS ON DIFFERENT DISTORTION TYPES IN THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

	Algorithm†	WN	JP2K	JPEG	Blur	FF	Overall
2D	PSNR	4.269	7.674	6.514	9.865	7.456	8.275
	SSIM [42]	3.988	6.783	5.572	7.370	5.910	6.741
	MS-SSIM [11]	3.334	5.621	3.792	8.397	6.212	7.025
	BRISQUE [9]	5.731	7.193	4.448	4.323	4.206	7.038
3D	Benoit [48]	4.028	6.096	3.787	11.763	6.894	7.490
	You [52]	4.396	4.186	4.086	8.649	4.649	6.772
	Yasakethu [43]	10.713	7.343	4.976	12.436	7.667	9.364
	Cyclopean MS-SSIM [57]	3.368	5.562	3.865	3.747	4.966	4.987
	Sazzad [53]	7.416	6.189	4.535	8.450	8.505	9.294
	Chen [58]	3.513	4.298	3.342	4.725	4.180	5.102
	S3D-BLINQ Index	3.547	5.482	4.169	4.453	4.199	4.657

† *Italics* indicate no-reference IQA models. Others are full-reference IQA models.

against the subjective opinion scores. Again, the quality scores predicted by these models were obtained by simply averaging the scores computed on the left- and right-view images. Among these 2D quality metrics, the full-reference SSIM index achieved the best performance.

The best S3D image quality prediction models that utilize 3D information were able to deliver a significantly higher 0.9 correlation level of performance. In particular, utilizing a synthesized cyclopean image boosts the performance of simple 2D IQA models, such as MS-SSIM, by more than 0.1 correlation level. By combining a synthesized cyclopean image with statistical models of disparity statistics, the no-reference S3D IQA model proposed by Chen *et al.* [58] was able to deliver performance comparable to the best full-reference models. However, as mentioned earlier, perceptual issues arise when forming cyclopean images with left-right (or right-left) bias from an S3D image pair. Since all of the asymmetrically distorted stereoscopic image pairs in the LIVE 3D Image Quality Database Phase II [65] were created using a pristine left-view image and a right-view image impaired by different types and degrees of distortions. Therefore, the two synthesized left- and right-view cyclopean images with bias may present differing perceptual characteristics, possibly resulting in biased performance of S3D IQA models utilizing cyclopean images, such as cyclopean MS-SSIM [57] and Chen [58].

TABLE V
COMPARISON OF DIFFERENT CYCLOPEAN MS-SSIM IMPLEMENTATIONS ON THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

Implementation	SROCC	LCC	RMSE
M_1	0.889	0.900	4.987
M_2	0.704	0.743	7.559
M_3	0.778	0.787	6.971

To further investigate how much bias these perceptually distinct cyclopean images can introduce, we examined three different implementations of cyclopean MS-SSIM using the two possible left- and right-view cyclopean images. Implementation M_1 computed the MS-SSIM score between the undistorted and distorted left-view cyclopean images generated using the disparity maps computed using the left-view images, which are always pristine, as references. This is the same implementation adopted in [57]. Implementation M_2 computed the MS-SSIM score from right-view cyclopean images with disparity maps computed using the right-view images, which always contain some types of distortion, as references. In the last implementation M_3, we generated the final quality score by simply averaging the two above MS-SSIM scores. The last implementation could be used in practical scenarios because real-world stereoscopic image pairs can be impaired with either asymmetry. We tabulate the performance of these three different cyclopean MS-SSIM implementations in Table V. It can be seen that the performance drops dramatically for implementation M_2. The seemingly more natural implementation (M_3) also suffers with reduced performance not significantly different than the 2D MS-SSIM index. Regarding the no-reference S3D IQA model proposed by Chen *et al.* [58], since the features are extracted asymmetrically, their algorithm would require modification to be applied on arbitrary asymmetries. In the performance comparisons, the original implementations in [57] and [58] were used.

Another issue with the cyclopean image model is that the disparity maps computed from the distorted S3D image pairs invariably contain inaccuracies and errors. It has been previously shown that full-reference S3D IQA algorithms utilizing cyclopean image models suffer by degradation of the disparity maps computed from distorted image pairs [57]. To investigate how this disparity inaccuracy can affect the performance of the proposed no-reference S3D-BLINQ Index, for each distorted image pair, we used the disparity maps computed from its corresponding reference/pristine image pair

TABLE VI
COMPARISON OF THE PROPOSED S3D-BLINQ INDEX FRAMEWORK WITH THE CONVERGENT CYCLOPEAN IMAGE USING DIFFERENT DISPARITY MAPS ON THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

Disparity Map Source	SROCC	LCC	RMSE
Reference/Pristine Image Pair	0.888	0.898	4.907
Distorted Image Pair (*S3D-BLINQ Index*)	0.905	0.913	4.657

to form the convergent cyclopean image, then extracted the NSS features from the reference/pristine convergent cyclopean image, and finally evaluated the perceptual S3D image quality based on these features. Table VI shows the performance comparison using different disparity map sources. Unlike the full-reference case, the proposed no-reference S3D-BLINQ Index delivers better performance using the disparity maps computed from distorted image pairs, which would be the common scenario for practical applications. This result could be explained by the fact that human subjects also evaluate the quality of S3D image pairs by directly computing internal disparity representations from distorted image pairs, with no access to presumably more accurate disparity maps that could be acquired from reference/pristine image pairs, were they available. Therefore, we believe that the proposed convergent cyclopean image model can better alleviate the perceptual distinction introduced by conventional cyclopean image models.

By synthesizing a more perceptually relevant and consistent convergent cyclopean image and by utilizing robust, effective bivariate and correlation natural image statistical models, S3D-BLINQ Index is able to achieve better than 0.9 correlation using both SROCC and LCC. It not only outperforms other state-of-the-art 2D and S3D IQA algorithms in terms of correlation monotonicity and accuracy, but also predicts the perceptual quality of stereoscopic image pairs with the lowest RMSE, as shown in Table IV.

Tables II – IV also detail the performance of each quality assessment algorithm on different types of distorted stereopairs. We can see that almost all 2D and S3D algorithms are able to predict quality scores that correlate well with human opinions for stereoscopic image pairs affected by the WN distortion. However, several quality metrics perform poorly when predicting the perceptual quality of stereopairs impaired by JPEG, JP2K, and Blur distortions. These poor performances may be explained as a result of binocular facilitation [51], [73] whereby distortions co-located with high depth variations are more easily found by human subjects. This observed effect is not yet well understood or properly modeled.

To examine the capability of different 2D and S3D IQA models when dealing with unequally distorted stereopairs, which may be more common in practice, we list in Table VII the performance of the same algorithms on both symmetrically and asymmetrically distorted stereoscopic image pairs in the LIVE 3D Image Quality Database Phase II [65]. It can be seen that most of the examined quality models are capable of predicting scores that correlate well with human judgments on symmetrically distorted stereopairs. However, almost all of them perform poorly on asymmetrically distorted stereopairs, except for those utilizing cyclopean images. Among these,

TABLE VII
COMPARISON (SROCC) OF DIFFERENT 2D AND 3D IMAGE QUALITY ASSESSMENT ALGORITHMS ON SYMMETRICALLY AND ASYMMETRICALLY DISTORTED STIMULI IN THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

	Algorithm†	Symmetric	Asymmetric	Overall
2D	PSNR	0.776	0.587	0.665
	SSIM [42]	0.828	0.733	0.792
	MS-SSIM [11]	0.912	0.684	0.777
	BRISQUE [9]	0.849	0.667	0.770
3D	Benoit [48]	0.860	0.671	0.728
	You [52]	0.914	0.701	0.786
	Yasakethu [43]	0.656	0.496	0.501
	Cyclopean MS-SSIM [57]	0.923	0.842	0.889
	Sazzad [53]	0.420	0.517	0.543
	Chen [58]	0.918	0.834	0.880
	S3D-BLINQ Index	0.937	0.849	0.905

† *Italics* indicate no-reference IQA models. Others are full-reference IQA models.

TABLE VIII
COMPARISON OF DIFFERENT 2D AND 3D IMAGE QUALITY ASSESSMENT MODELS ON THE TOYAMA DATABASE

	Algorithm†	SROCC	LCC	RMSE
2D	PSNR	0.586	0.554	0.971
	SSIM [42]	0.846	0.862	0.591
	MS-SSIM [11]	0.935	0.935	0.415
	BRISQUE [9]	0.828	0.830	0.573
3D	Benoit [48]	0.902	0.910	0.483
	You [52]	0.857	0.864	0.586
	Cyclopean MS-SSIM [57]	0.862	0.864	0.587
	Sazzad [53]	0.785	0.795	0.708
	Chen [58]	0.904	0.913	0.431
	S3D-BLINQ Index	0.917	0.933	0.360

† *Italics* indicate no-reference IQA models. Others are full-reference IQA models.

S3D-BLINQ Index afforded the best performance on both symmetrically and asymmetrically distortions, resulting in the best overall correlation numbers as well.

Finally, to further validate the robustness of S3D-BLINQ Index, we also evaluated its performance on another 3D image quality database, the Toyama Database [74], [75], which was created under significantly different conditions. The Toyama Database contains 10 reference stereoscopic image pairs, and 480 symmetrically/asymmetrically JPEG-coded stereopairs, all having a resolution of 640×480. A double stimulus impairment scale (DSIS) method was used to conduct the subjective experiment, where distorted and reference image pairs were displayed sequentially in random order. Subjects were asked to assess the annoyance they experienced when viewing the distorted image pair against the reference one using a five-level scale: very annoying $= 1$, annoying $= 2$, slightly annoying $= 3$, perceptible but not annoying $= 4$, and imperceptible $= 5$. A rather small 10-inch auto-stereoscopic display was used in their subjective experiment.

Table VIII summarizes the performance of the tested 2D and S3D IQA models on the Toyama Database in terms of SROCC, LCC, and RMSE against the recorded subjective mean opinion scores (MOS). Note that the reported performance of the tested no-reference IQA models were computed by a 10-fold, i.e., leave-one-out, cross-validation accuracy estimation procedure [71]. Except for PSNR, all 2D and

TABLE IX
COMPARISON (SROCC) OF THE PROPOSED S3D-BLINQ INDEX FRAMEWORK USING DIFFERENT FEATURE SETS ON SYMMETRICALLY AND ASYMMETRICALLY DISTORTED STIMULI IN THE LIVE 3D IMAGE QUALITY DATABASE PHASE II

Feature Set	Symmetric	Asymmetric	Overall
Spatial-Domain Univariate	0.911	0.808	0.873
Wavelet-Domain Univariate	0.852	0.815	0.854
Bivariate and Correlation	0.877	0.826	0.868
All (*S3D-BLINQ Index*)	0.937	0.849	0.905

S3D IQA models perform fairly well on the Toyama Database, while the full-reference MS-SSIM delivers the best performance. The S3D-BLINQ Index outperforms all of the other S3D IQA models, and is able to deliver performance that is competitive with MS-SSIM, achieving similar correlation levels and a lower RMSE.

B. Augmentation of the Bivariate and Correlation Models

The solid performance of S3D-BLINQ Index can be attributed to utilizing the perceptually relevant, convergent cyclopean image and robust and descriptive bivariate and correlation NSS models. Here we analyze the performance boost provided by the novel bivariate and correlation NSS features underlying the S3D-BLINQ Index learning process. Specifically, we incorporated three different sets of features extracted from the convergent cyclopean image, following the same framework as described in Section IV. The three feature sets include the spatial-domain univariate NSS features, the wavelet-domain univariate NSS features, and the bivariate and correlation NSS features. In essence, the univariate spatial- and wavelet-domain NSS features capture the marginal statistics of image bandpass responses at each sub-band, which are affected by different types of distortion. For example, WN always increases the variance of sub-band coefficients [7]. On the other hand, the bivariate NSS features model the joint statistics of spatially adjacent sub-band coefficients, which are of particular value since high-order image dependencies are disturbed by distortions. Most importantly, the NSS correlation features further capture the inter-orientation dependencies of image bandpass responses across sub-bands [21], thereby making it possible to quantify deviations caused by spatially correlated image impairments.

We tabulate the performance of these three different feature sets, as well as the combination of all, i.e., S3D-BLINQ Index, in Table IX. It can be seen that using only the spatial-domain univariate NSS features is able to achieve 0.9 level of correlation performance on symmetrically distorted stereopairs, while the wavelet-domain univariate NSS features improve performance on asymmetric distortions. The bivariate and correlation NSS features further augment performance on asymmetrically distorted stereopairs, resulting in an overall 0.9 SROCC score when combining all feature sets. This comparison also suggests that the best result can be attained by using the complete S3D-BLINQ Index feature set, while sufficient accuracy can be provided by using only the bivariate and correlation features to capture both symmetric and asymmetric distortions in practical applications.

VI. CONCLUSION

We generalized our new bivariate and correlation NSS models to capture the spatial oriented structure in bandpass distorted 2D and S3D images. These bivariate and correlation models are validated to be able to robustly and reliably quantify the statistical regularities embedded in spatially adjacent luminance pixels, and preliminarily yet systematically address one of the most important issues on NSS modeling of higher-order dependencies, which has not been well explored in literature. To demonstrate the efficacy of these new models, we deploy them to develop a new no-reference S3D IQA framework – the Stereoscopic/3D BLind Image Naturalness Quality (S3D-BLINQ) Index. Our contributions are twofold. First, we defined a novel and powerful set of quality-discriminative features by exploiting the new bivariate and correlation NSS models. Second, we proposed a convergent cyclopean image model to address bias encountered by earlier cyclopean image models.

We believe that both the new bivariate/correlation natural scene statistical models and the convergent cyclopean image model provide useful and robust tools for predicting perceptual S3D image quality. These novel models and tools can potentially be applied to solve various 3D vision and image/video processing problems, e.g., de-noising, super-resolution, shape-from-X, etc. Future work includes developing a "completely" blind S3D quality evaluator by appropriately modeling stereopair naturalness, and by incorporating perceptual measurements of visual discomfort into S3D quality of experience (QoE) models.

ACKNOWLEDGMENT

The authors thank the authors of [74], [75] for kindly providing access to the Toyama Database.

REFERENCES

[1] C.-C. Su, L. K. Cormack, and A. C. Bovik, "Closed-form correlation model of oriented bandpass natural images," *IEEE Signal Process. Lett.*, vol. 22, no. 1, pp. 21–25, Jan. 2015.
[2] C.-C. Su, L. K. Cormack, and A. C. Bovik, "Bivariate statistical modeling of color and range in natural scenes," *Proc. SPIE*, vol. 9014, p. 90141G, Feb. 2014.
[3] B. A. Olshausen and D. J. Field, "Natural image statistics and efficient coding," *Netw., Comput. Nerual Syst.*, vol. 7, no. 2, pp. 333–339, 1996.
[4] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1193–1216, Mar. 2001.
[5] Z. Wang and A. C. Bovik, "Reduced- and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Nov. 2011.
[6] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.
[7] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
[8] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[9] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
[10] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, vol. 2. Nov. 2003, pp. 1398–1402.

[12] B. Potetz and T. S. Lee, "Scaling laws in natural scenes and the inference of 3D shape," in *Advances in Neural Information Processing Systems*, vol. 18. Cambridge, MA, USA: MIT Press, 2006, pp. 1089–1096.

[13] Y. Liu, L. K. Cormack, and A. C. Bovik, "Statistical modeling of 3D natural scenes with application to Bayesian stereopsis," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2515–2530, Sep. 2011.

[14] C.-C. Su, L. K. Cormack, and A. C. Bovik, "Color and depth priors in natural images," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2259–2274, Jun. 2013.

[15] J. Portilla and E. P. Simoncelli, "Texture modeling and synthesis using joint statistics of complex wavelet coefficients," in *Proc. IEEE Workshop Statist. Comput. Theories Vis.*, vol. 12. Jun. 1999.

[16] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49–70, Oct. 2000.

[17] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," *Proc. SPIE*, vol. 3813, pp. 188–195, Oct. 1999.

[18] M. J. Wainwright, O. Schwartz, and E. P. Simoncelli, "Natural image statistics and divisive normalization: Modeling nonlinearities and adaptation in cortical neurons," in *Proc. Probab. Models Brain, Perception Neural Funct.*, Feb. 2002, pp. 203–222.

[19] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1647–1658, Nov. 2001.

[20] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[21] D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1610–1620, Jun. 2006.

[22] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[23] K. Seshadrinathan and A. C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 335–350, Feb. 2010.

[24] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3. Oct. 1995, pp. 444–447.

[25] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer. A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[26] B. A. Olshausen and D. J. Field, "How close are we to understanding V1?" *Neural Comput.*, vol. 17, no. 8, pp. 1665–1699, 2005.

[27] S. Lyu, "Dependency reduction with divisive normalization: Justification and effectiveness," *Neural Comput.*, vol. 23, no. 11, pp. 2942–2973, 2011.

[28] F. Pascal, L. Bombrun, J.-Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized Gaussian distributions," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5960–5971, Dec. 2013.

[29] H. R. Sheikh, Z. Wang, L. K. Cormack, and A. C. Bovik. *LIVE Image Quality Assessment Database*. [Online]. Available: http://live.ece.utexas.edu/research/quality/subjective.htm

[30] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[31] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proc. IEEE*, vol. 101, no. 9, pp. 2008–2024, Sep. 2013.

[32] Motion Picture Association of America (MPAA). (2012). *Theatrical Market Statistics*. [Online]. Available: http://www.mpaa.org/policy/industry

[33] BBC News—Technology. (Aug. 2013). *James Cameron: All Entertainment 'Inevitably 3D'*. [Online]. Available: http://www.bbc.co.uk/news/entertainment-arts-23790877

[34] J. Baltes, S. McCann, and J. Anderson, "Humanoid robots: Abarenbou and daodan," RoboCup, Bremen, Germany, Tech. Rep., 2006.

[35] A. M. William and D. L. Bailey, "Stereoscopic visualization of scientific and medical content," in *Proc. ACM SIGGRAPH Educ. Program*, 2006, Art. ID 26.

[36] C.-F. Westin, "Extracting brain connectivity from diffusion MRI [life sciences]," *IEEE Signal Process. Mag.*, vol. 24, no. 6, pp. 124–152, Nov. 2007.

[37] Y. A. W. de Kort and W. A. IJsselsteijn, "Reality check: The role of realism in stress reduction using media technology," *Cyberpsychol. Behavior*, vol. 9, no. 2, pp. 230–233, 2006.

[38] A. Puri, R. V. Kollarits, and B. G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *Signal Process., Image Commun.*, vol. 10, nos. 1–3, pp. 201–234, 1997.

[39] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 993–1008, Aug. 2003.

[40] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, document ITU-R Rec. BT.500-11, Sep. 2009.

[41] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[42] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[43] S. L. P. Yasakethu, C. T. E. R. Hewage, W. A. C. Fernando, and A. M. Kondoz, "Quality analysis for 3D video using 2D video quality models," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1969–1976, Nov. 2008.

[44] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," *Proc. SPIE*, vol. 6803, p. 680305, Feb. 2008.

[45] G. C. DeAngelis, I. Ohzawa, and R. D. Freeman, "Depth is encoded in the visual cortex by a specialized receptive field structure," *Nature*, vol. 352, no. 6331, pp. 156–159, 1991.

[46] D. J. Fleet, H. Wagner, and D. J. Heeger, "Neural encoding of binocular disparity: Energy models, position shifts and phase shifts," *Vis. Res.*, vol. 36, no. 12, pp. 1839–1857, 1996.

[47] B. G. Cumming, "An unexpected specialization for horizontal disparity in primate primary visual cortex," *Nature*, vol. 418, no. 6898, pp. 633–636, 2002.

[48] A. Benoit, P. L. Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP J. Image Video Process.*, vol. 2008, pp. 1–13, Jan. 2009.

[49] V. Kolmogorov and R. Zabih, "Multi-camera scene reconstruction via graph cuts," in *Proc. 7th Eur. Conf. Comput. Vis.*, vol. 2352. May 2002, pp. 82–96.

[50] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comput. Vis.*, vol. 70, no. 1, pp. 41–54, Oct. 2006.

[51] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "Distortion conspicuity on stereoscopically viewed 3D images may correlate to scene content and distortion type," *J. Soc. Inf. Display*, vol. 21, no. 11, pp. 491–503, Nov. 2013.

[52] 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. Int. Workshop Video Process. Quality Metrics*, vol. 9. 2010.

[53] Z. M. P. Sazzad, R. Akhter, J. Baltes, and Y. Horita, "Objective no-reference stereoscopic image quality prediction based on 2D image features and relative disparity," *Adv. Multimedia*, vol. 2012, Jan. 2012, Art. ID 8.

[54] B. Julesz, *Foundations of Cyclopean Perception*. Chicago, IL, USA: Univ. Chicago Press, 1971.

[55] A. Maalouf and M.-C. Larabi, "CYCLOP: A stereo color image quality assessment metric," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2011, pp. 1161–1164.

[56] S. J. Daly, "Visible differences predictor: An algorithm for the assessment of image fidelity," *Proc. SPIE*, vol. 1666, pp. 2–15, Feb. 1992.

[57] 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 Process., Image Commun.*, vol. 28, no. 9, pp. 1143–1155, Oct. 2013.

[58] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "No-reference quality assessment of natural stereopairs," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3379–3391, Sep. 2013.

[59] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, nos. 1–3, pp. 7–42, 2002.

[60] W. J. M. Levelt, *On Binocular Rivalry*. The Hague, The Netherlands: Mouton, 1968.

[61] D. L. Ruderman, "The statistics of natural images," *Netw., Comput. Neural Syst.*, vol. 5, no. 4, pp. 517–548, 1994.

[62] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 1, pp. 52–56, Feb. 1995.
[63] N.-E. Lasmar, Y. Stitou, and Y. Berthoumieu, "Multiscale skewed heavy tailed model for texture analysis," in *Proc. IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 2281–2284.
[64] D. J. Field, "Wavelets, vision and the statistics of natural scenes," *Philos. Trans. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 357, no. 1760, pp. 2527–2542, 1999.
[65] A. Moorthy, C.-C. Su, M.-J. Chen, A. Mittal, L. K. Cormack, and A. C. Bovik. *LIVE 3D Image Quality Database Phase I and Phase II*. [Online]. Available: http://live.ece.utexas.edu/research/quality/live_3dimage.html
[66] C.-C. Su, L. K. Cormack, and A. C. Bovik. *Validation of the Exponentiated Cosine Model*. [Online]. Available: http://live.ece.utexas.edu/research/3dnss/bicorr.html
[67] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, 2000.
[68] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
[69] M. Narwaria and W. Lin, "Objective image quality assessment based on support vector regression," *IEEE Trans. Neural Netw.*, vol. 21, no. 3, pp. 515–519, Mar. 2010.
[70] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. ID 27. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm
[71] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14th Int. Joint Conf. Artif. Intell.*, 1995, pp. 1137–1143.
[72] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
[73] M.-J. Chen, A. C. Bovik, and L. K. Cormack, "Study on distortion conspicuity in stereoscopically viewed 3D images," in *Proc. IEEE IVMSP Workshop*, Jun. 2011, pp. 24–29.
[74] Z. M. P. Sazzad, S. Yamanaka, Y. Kawayokeita, and Y. Horita, "Stereoscopic image quality prediction," in *Proc. Int. Workshop Quality Multimedia Exper.*, Jul. 2009, pp. 180–185.
[75] R. Akhter, Z. M. P. Sazzad, Y. Horita, and J. Baltes, "No-reference stereoscopic image quality assessment," *Proc. SPIE*, vol. 7524, p. 75240T, Jan. 2010.

Che-Chun Su received the B.S. degree in electrical engineering and the M.S. degree in communication engineering from National Taiwan University, Taipei, Taiwan, in 2004 and 2006, respectively, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, Austin, TX, USA, in 2014.

His research interests include computational image/video processing and computer vision, based on human visual perception and natural scene statistical modeling.

Lawrence K. Cormack is currently a Professor of Psychology and Neuroscience at the University of Texas at Austin, and is an Active Member of the Center for Perceptual Systems. He received his B.S. with Highest Honors in Psychology from the University of Florida in 1986 and his Ph.D. in Physiological Optics from the University of California, Berkeley, in 1992. Dr. Cormack's primary research interest is how the brain processes the motion of objects through the 3D environment. This topic is of interest because it is of utmost behavioral importance, it is not well understood, and yet it is a synthesis of two topics that are reasonably well understood (the binocular perception of depth, and the perception of 2D motion). He is also interested in natural scene statistics and their relation to the evolution of the visual system, particularly with regard to depth and motion processing.

Dr. Cormack teaches mostly graduate statistics.

Alan C. Bovik (F'96) holds the Cockrell Family Endowed Regents Chair in Engineering at The University of Texas at Austin, where he is Director of the Laboratory for Image and Video Engineering (LIVE). He is a faculty member in the Department of Electrical and Computer Engineering and the Institute for Neuroscience. His research interests include image and video processing, computational vision, and visual perception. He has published over 700 technical articles in these areas and holds several U.S. patents. His publications have been cited more than 40 000 times in the literature, his current H-index is over 70, and he is listed as a Highly-Cited Researcher by Thompson Reuters. His several books include the companion volumes *The Essential Guides to Image and Video Processing* (Academic Press, 2009).

He has received a number of major awards from the IEEE Signal Processing Society, including: the Society Award (2013); the Technical Achievement Award (2005); the Best Paper Award (2009); the *Signal Processing Magazine* Best Paper Award (2013); the Education Award (2007); the Meritorious Service Award (1998) and (co-author) the Young Author Best Paper Award (2013). He also was named recipient of the Honorary Member Award of the Society for Imaging Science and Technology for 2013, received the SPIE Technology Achievement Award for 2012, and was the IS&T/SPIE Imaging Scientist of the Year for 2011. He is also a recipient of the Hocott Award for Distinguished Engineering Research from the Cockrell School of Engineering at The University of Texas at Austin (2008) and the Distinguished Alumni Award from the University of Illinois at Champaign-Urbana (2008). He is a Fellow of the IEEE, the Optical Society of America (OSA), and the Society of Photo-Optical and Instrumentation Engineers (SPIE). He cofounded and was the longest-serving Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING (1996-2002), created and served as the first General Chair of the IEEE International Conference on Image Processing, held in Austin, Texas, in November, 1994, along with numerous other professional society activities, including: Board of Governors, IEEE Signal Processing Society, 1996-1998; Editorial Board, The PROCEEDINGS OF THE IEEE, 1998- 2004; and Series Editor for *Image, Video, and Multimedia Processing*, Morgan and Claypool Publishing Company, 2003-present. His was also the General Chair of the 2014 Texas Wireless Symposium, held in Austin in November of 2014.

Dr. Bovik is a registered Professional Engineer in the State of Texas and is a frequent consultant to legal, industrial and academic institutions.