Predicting the Quality of Fused Long Wave Infrared and Visible Light Images

David Eduardo Moreno-Villamarín, *Student Member, IEEE*,
Hernán Darío Benítez-Restrepo, *Senior Member, IEEE*,
and Alan Conrad Bovik, *Fellow, IEEE*

***Abstract*— The capability to automatically evaluate the quality of long wave infrared (LWIR) and visible light images has the potential to play an important role in determining and controlling the quality of a resulting fused LWIR-visible light image. Extensive work has been conducted on studying the statistics of natural LWIR and visible images. Nonetheless, there has been little work done on analyzing the statistics of fused LWIR and visible images and associated distortions. In this paper, we analyze five multi-resolution-based image fusion methods in regards to several common distortions, including blur, white noise, JPEG compression, and non-uniformity. We study the natural scene statistics of fused images and how they are affected by these kinds of distortions. Furthermore, we conducted a human study on the subjective quality of pristine and degraded fused LWIR-visible images. We used this new database to create an automatic opinion-distortion-unaware fused image quality model and analyzer algorithm. In the human study, 27 subjects evaluated 750 images over five sessions each. We also propose an opinion-aware fused image quality analyzer, whose relative predictions with respect to other state-of-the-art models correlate better with human perceptual evaluations than competing methods. An implementation of the proposed fused image quality measures can be found at https://github.com/ujemd/NSS-of-LWIR-and-Vissible-Images. Also, the new database can be found at http://bit.ly/2noZlbQ.**

***Index Terms*— NSS, LWIR, multi-resolution image fusion, fusion performance.**

I. Introduction

IN RECENT years, increasing levels of uncertain global security, along with the availability of cheap, intelligent digital cameras is encouraging interest in the development of video systems capable of detecting anomalies or events that may affect the economics and safety of human activities [1]. Popular outdoor video surveillance systems that rely on electro-optical sensors such as visible-light CCD cameras are often prone to failures due to ambient illumination changes and weather conditions [2], [3]. One way of improving performance is to use alternate modes of sensing, such as infrared sensing, either alone or in combination with visible light. Decreasing costs and increasing miniaturization have made infrared sensing an interesting element in surveillance system design [4]–[8].

Although Long Wave Infrared (LWIR) sensors can accurately capture useful video data in low-light and night-vision applications, the images obtained lack the color information and relative luminances of visible spectrum sensors. By contrast, RGB sensors do capture color and correct relative luminances, but are sensitive to illumination variations and lack the ability to capture revealing information available in the thermal bands [9]. Two main benefits of the joint use of thermal and visible sensors are: the complementary natures of the two modalities and the information redundancy captured by the sensors, which increases the reliability and robustness of a surveillance system. These advantages have motivated the computer vision community to study and investigate algorithms for fusing infrared and visible videos for surveillance applications [6].

Due to growing interest in LWIR and visible light image fusion, considerable efforts have been made to develop objective quality measures of fused images. The performance of different image fusion algorithms has been evaluated by image fusion quality metrics that are based on information theory [10]–[12], space and frequency based image features [13]–[16], image structural similarity [17]–[19], and models of human perception [20], [21].

Chen and Blum [21] investigated the performance of fusion metrics based on human vision system models, assuming the presence of several levels of additive white Gaussian noise (AWGN). Liu *et al.* [22] analyzed the impact of AWGN and blur on fused images. They found that the quality of fused images is degraded with decreases in the quality of the images being fused. When the AWGN level was severe, the fused images were all of almost the same quality, regardless of the fusion scheme used. These studies did not analyze important real distortions that often occur on LWIR sensors, such as "non uniformity' (NU) impairments and the "halo effect."

Manuscript received October 12, 2016; revised February 21, 2017 and April 1, 2017; accepted April 3, 2017. Date of publication May 3, 2017; date of current version May 19, 2017. This work was supported in part by the COLCIENCIAS, in part by the Convocatoria para el apoyo a proyectos con Norteamérica 2014 Program, and in part by the Pontificia Universidad Javeriana-Cali thorough the Evaluation of video distortions on fused infrared and visible videos in surveillance applications Project. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Stefan Winkler. *(Corresponding author: David Eduardo Moreno-Villamarín.)*

D. E. Moreno-Villamarín and H. D. Benítez-Restrepo are with the Departmento de Electrónica y Ciencias de la Computación, Pontificia Universidad Javeriana, Cali 760031, Colombia (e-mail: david.moreno@ieee.org; benitez@ieee.org).

A. C. Bovik is with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: bovik@ece.utexas.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author. The material includes pdf. The total size of the pdf is 202 kB. Contact david.moreno@ieee.org for further questions about this work.

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.2017.2695898

NU manifests as an undesirable grid-like pattern on images obtained using focal plane arrays [23], while the "halo effect" appears around very hot or cold objects in imagery from uncalibrated ferroelectric BST sensors [24], causing regions surrounding bright objects to grow darker, and regions around dark objects to grow lighter [25]. Although extensive work has been conducted on studying the statistics of natural scenes captured in the visible light spectrum and their relationship to picture quality [26]–[30], and some studies have been done on the statistics of LWIR images [31], [32], very little work has been done on analyzing the statistics of fused LWIR and visible light images, and how those statistics might be affected by the presence of any of multiple possible impairments.

The objective of this work is to analyze how image distortions such as AWGN, blur, JPEG compression, and non-uniformity noise in LWIR and visible images affect the statistics (NSS) of fused LWIR-visible images. We deploy previous bandpass image statistical models proposed in [31], [33], and [34] as a starting point, and create 'opinion-distortion-unaware' (ODU) and 'opinion-aware' (OA) no-reference image quality prediction models using them. An important distinction between the model we develop here and BRISQUE [33], is that we deploy an additional set of features, including log-derivative and divisively normalized steerable pyramid coefficients that provide higher sensitivity to high frequency noise and that explicitly capture band-pass characteristics. An image quality model is ODU if it does not require training on databases of human judgments of distorted images and does not rely on training on, or tuning to specific distortions. By contrast, a model is OA if it has been trained on a database(s) of human rated distorted images and associated human subjective opinion scores. A deep comparison of the results obtained by our proposed OA and ODU models with those of state-of-the-art algorithms shows that our new model achieves highly competitive results.

In a previous study, we analyzed the effects of distortions on the NSS of images fused by three different methods [35]. We significantly extend that work by modeling the statistics of fused LWIR and visible light images, by analyzing these statistics on five popular multi-resolution fusion methods, by conducting a human study on the subjective quality of pristine and degraded fused LWIR-visible images, and by creating new and effective ODU and OA fused image quality analyzers.

The remainder of this paper is organized as follows: the following subsections outline the image databases, statistical models, and image fusion methods that we use. Section 2 describes the processing and feature models we employ. Section 3 presents the results of an extensive subjective human study that we carried out, and also details two NSS-based fused image quality models that we have developed, and a comparison of their performance against other models in regards to their ability to predict subjective scores. Sections 4 and 5 broadly discuss the results obtained along with suggestions for further work.

A. LWIR and Visible Image Sources

This study of multimodal image fusion uses three databases that we hereafter refer to as OSU [36], TNO [37], [38], and MORRIS [32]. The OSU database contains 80 visible light and LWIR image pairs from the Ohio State University campus. The TNO database contains 74 image pairs. The MORRIS database contains 14 indoor and outdoor image pairs on urban environments. Before processing the images, they were all linearly re-scaled to the range $[0, 1]$ to be able to apply the simulated distortions consistently. A few example images from these databases can be seen in Fig. 1.

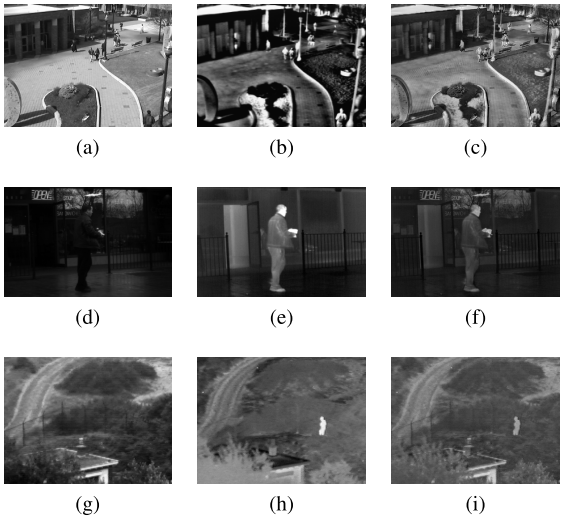


Fig. 1. Example images from the OSU (a-c), MORRIS (d-f) and TNO (g-i) databases. (a), (d), and (g) are visible light images, (b), (e), and (h) are LWIR images, and (c), (f), and (i) are images fused using a gradient pyramid.

Fig. 2. Examples of fused images after the following distortions were applied to the constituent visible light and LWIR images: (a) Additive white gaussian noise. (b) Non-Uniformity. (c) Blur. (d) JPEG compression. Images obtained from [32].

B. Distortion Models

Several studies have characterized and modeled noise in the LWIR spectrum. Images obtained from focal plane arrays can present NU fixed pattern noise [23], which produces a grid-like pattern. In this work we deploy the spectral additive model of NU fixed pattern noise presented in [31] and [39].

The distortion level is controlled using the standard deviation parameter σ_{NU}, which scales the dynamic range of the NU noise. Other common types of distortions which could affect both LWIR and visible images are also considered here, such as AWGN, blur, and JPEG compression. Three distortion levels are used throughout the study for each distortion type, which were applied to the LWIR and visible images of the three databases. For AWGN and NU the standard deviation was varied as $\sigma_{AWGN} = \sigma_{NU} = \{0.0025, 0.01375, 0.025\}$; for blur, a Gaussian blur kernel of size 15×15 pixels with $\sigma_{blur} = \{1, 2, 3\}$ was used; and for JPEG compression, the quality was set to 100, 90 and 80 percent using the "`imwrite`" Matlab algorithm. Fig. 2 depicts several fused images obtained when

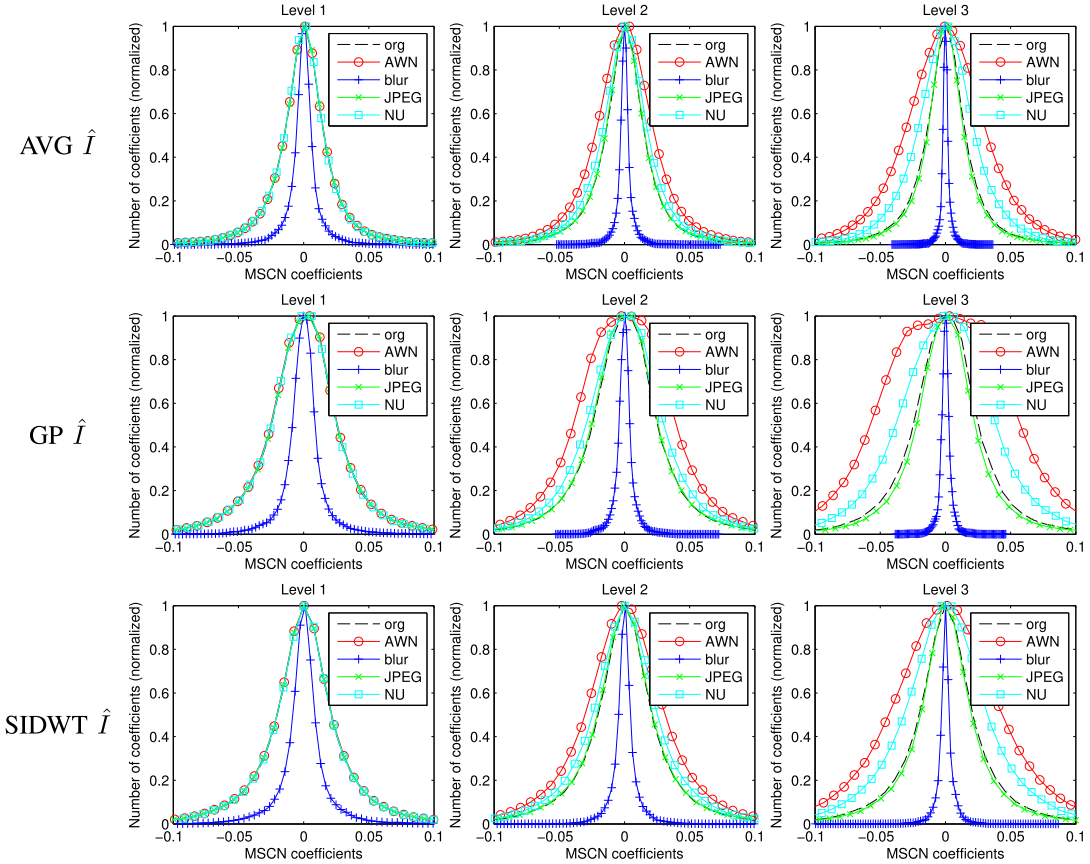


Fig. 3. Comparison of MSCN histograms of 154 ROIs from fused images (80 from the OSU database and 74 from the TNO database). The ROI sizes were 64×64 pixels. The figures show three distortion levels increasing from left to right for AWGN, blur, JPEG compression, and non uniformity (NU). In some cases the plots of the histograms of pristine images "org" are obscured by overlap of other curves. The terms AVG, GP, SIDWT refer to the fusion methods Average, Gradient Pyramid, and Shift Invariant Discrete Wavelet Transform, respectively.

both image sources were affected by the most severe distortion level.

C. Multi-Resolution Fusion Methods

In night vision applications, one of the most commonly used tools is multi-resolution image fusion (MIF), which aims to retain the main features from the source images [40]. This technique focuses on accessible multi-resolution feature representations and an image fusion rule to guide the combination of coefficients in the transform domain. How the fusion algorithm adapts to different object-to-background situations is still not well understood.

Liu *et al.* [22] used fusion performance models to evaluate six common multi-resolution fusion methods, of which we consider the following five: average (AVG), gradient pyramid (GP) [41], Laplacian pyramid (LP) [42], ratio of low-pass pyramid (RP) [43], and shift-invariant discrete wavelet transform with Haar wavelet (SIDWT) [44]. The decomposition level used in each of the algorithms was set to four, and the fusion rule used in each case was the maximum of the high-pass pair of channels and the average of the low-pass channels.

II. NSS of Fused LWIR and Visible Images

A. Processing Model

Prior research on non-reference IQA has determined that the most successful IQA measures are based on bandpass statistical image models [34], [45]. Hence, our approach deploys as processing models:

(i) Mean-Subtracted Contrast Normalized (MSCN) coefficients [33].

(ii) Four 'paired product' horizontal (H), vertical (V), and diagonal ($D1$ and $D2$) coefficients (or directional correlations) calculated as the products of adjoining MSCN coefficients [33].

(iii) MSCN coefficients supplemented by a set of log-derivative coefficients ($PD1 \ldots PD7$), which are intended to provide higher sensitivity to high-frequency noise [46].

(iv) Coefficients obtained from a steerable pyramid image decomposition are used to capture oriented band-pass characteristics [29].

In this section we illustrate the most representative histograms for three fusion methods (average, gradient pyramid, and shift-invariant discrete wavelet transform) and for each type of coefficients. The histograms of the MSCN coefficients of regions of interest (ROI) of fused LWIR and visible light images suffering from three levels of four kinds of applied distortion (as well as no distortion) are depicted in Fig. 3. A total of 154 ROIs from five scenes were selected by extracting center patches of size 64×64 from the OSU and TNO databases. The parameters of each noise type were as described earlier, in subsection I-B. Observe that blur distortion tended to

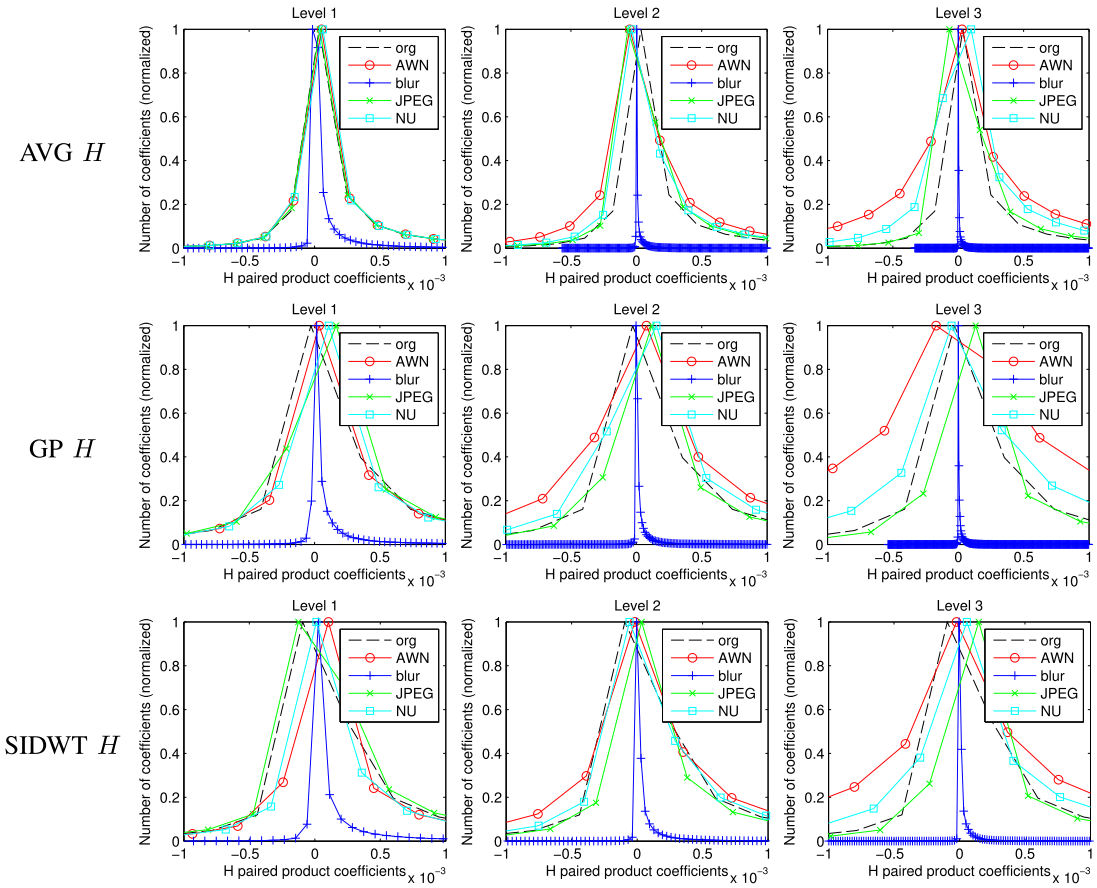


Fig. 4. Comparison of horizontal paired product histograms of 154 ROIs from fused images. The ROI sizes were 64×64 pixels. The figures show increasing distortion levels from left to right for AWGN, blur, JPEG compression, and NU distortions. In some cases the plots of undistorted "org" histograms are obscured by overlap of other curves.

produce thinner histograms, while AWGN and NU produced wider histograms. We also noticed how the Laplacian pyramid fusion method appeared to more severely affect the shapes of the fused image histograms (more examples can be found in the supplement at: http://bit.ly/2mi0CAy), and thereby the likely degree of the apparent naturalness of the resulting fused images.

Exemplar histograms of the paired product coefficients of fused images were generated following the same procedure as was used for the MSCN fused images. Fig. 4 depicts the horizontal (H) paired product histograms. A remarkable characteristic is the high sensitivity to blur distortions, which produces thinner histograms. In the case of the Gaussian pyramid fusion method, there is a noticeable sensitivity to AWGN, which leads to wider histograms, as it does for NU distortion, to a lesser degree. However, these histograms fail to effectively distinguish between JPEG compressed and pristine images. Histograms of V, $D1$, and $D2$ coefficients presented similar characteristics.

Using the same ROI extraction procedure on both pristine and distorted images, we computed the log-derivative coefficients, and plotted exemplar histograms of the $PD6$ coefficients in Fig. 5. These coefficients exhibit a higher sensitivity to blur than the other distortions. It is interesting to see that in the $PD6$ histograms, JPEG distortion produces increasingly thinner histograms as the image quality decreases.

In our case, the steerable pyramid decomposition was computed over six orientations, where each band is denoted d_α^θ, where α indicates the scale and $\theta \in \{0°, 30°, 60°, 90°, 120°, 150°\}$. Using the same pooled ROI extraction procedure, histograms produced from the $d_1^{30°}$ coefficients are plotted in Fig. 6, where the effect of AWGN is in general noticeable, yet not as apparent as it was with other types of studied coefficients. The effect of NU noise is minimal, contrary to blur distortion, where the widths of the histograms markedly decrease at higher distortion levels. However, when analyzing the distortion behavior in the horizontal and vertical subbands, d_1^0 and d_1^{90}, the NU becomes distinctively spread out. Please refer to the supplement at http://bit.ly/2mi0CAy for other examples of these histograms. In general, the various histograms are distinctively descriptive of the effects of the various types of distortions of fused images. By using closed form statistical models to parametrically fit these histograms, it is possible to extract distortion-sensitive features, as we show in the next section.

B. Feature Models

Prior work on statistical image modeling has led to the development of models of the empirical distributions of both

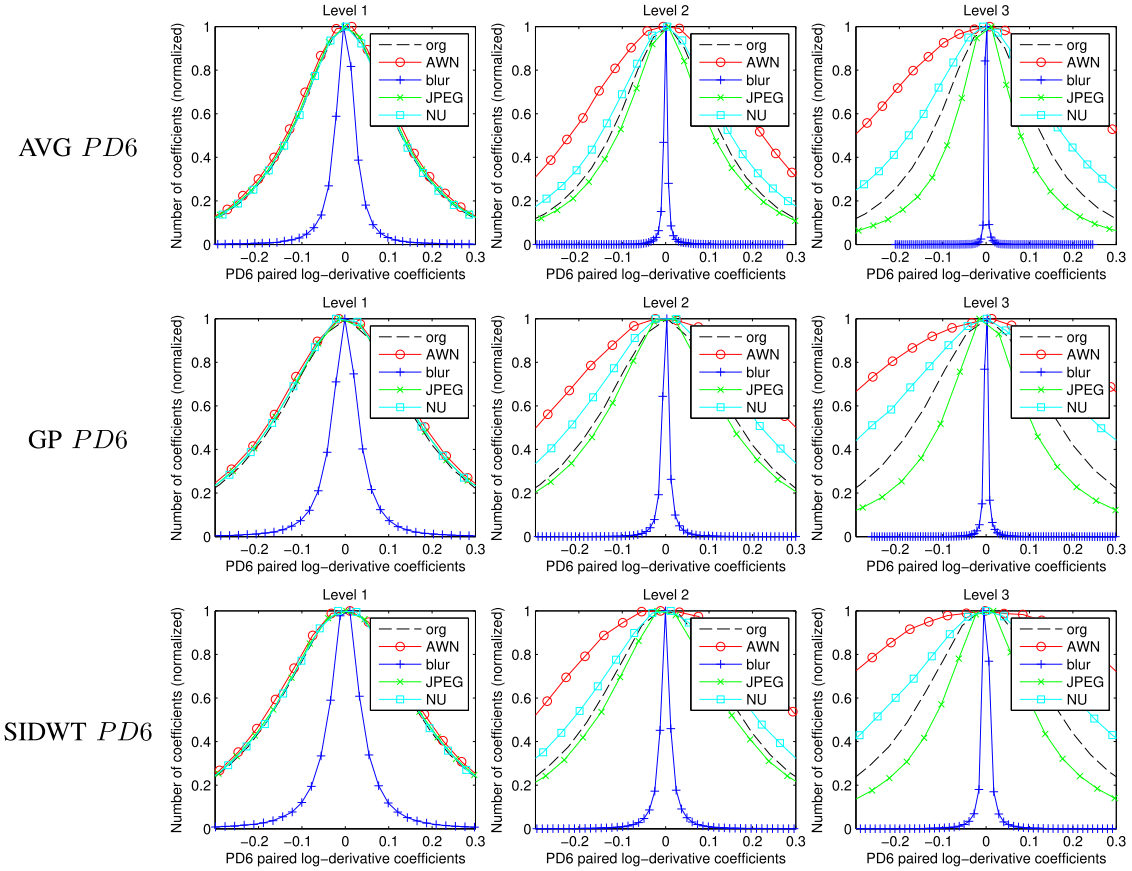


Fig. 5. Comparison of log-derivative histograms of 154 ROIs obtained from fused images. The ROIs sizes are 64×64 pixels. The figures show increasing distortion levels from left to right for AWGN, blur, JPEG compression, and NU. In some cases the plots of undistorted "org" histograms are obscured by overlap of other curves.

high-quality and distorted photographic visible light images, as well as of infrared pictures that have been subjected to bandpass processing followed by divisive normalization. These are both well-modeled as following a Generalized Gaussian Distribution (GGD). This is true of pictures processed by MSCN, paired log-derivative filters, and steerable pyramid filters [31], [33], although the fitting parameters will characteristically vary. The standard method is to fit the histogram of the bandpass coefficients to the GGD probability density function:

$$f(x;\alpha,\sigma) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}^{\alpha}\right)\right) \tag{1}$$

where

$$\beta = \sigma\sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}, \tag{2}$$

α is the shape parameter, σ is the spread parameter, and Γ is the Gamma function:

$$\Gamma(t) = \int_0^{\infty} x^{t-1}\exp^{-x}\,dx. \tag{3}$$

The products of spatially adjacent bandpass/normalized coefficients are well modeled as following an Asymmetric Gaussian Distribution (AGGD), with probability density function:

$$f(x;\upsilon,\sigma_l,\sigma_r) = \begin{cases} \dfrac{\upsilon}{(\beta_l+\beta_r)\Gamma(1/\upsilon)} \exp\left(-\left(\dfrac{-x}{\beta_l}^{\upsilon}\right)\right) & x<0 \\ \dfrac{\upsilon}{(\beta_l+\beta_r)\Gamma(1/\upsilon)} \exp\left(-\left(\dfrac{x}{\beta_l}^{\upsilon}\right)\right) & x\geq 0 \end{cases} \tag{4}$$

where

$$\beta_l = \sigma_l\sqrt{\frac{\Gamma(1/\upsilon)}{\Gamma(3/\upsilon)}} \tag{5}$$

and

$$\beta_r = \sigma_r\sqrt{\frac{\Gamma(1/\upsilon)}{\Gamma(3/\upsilon)}}. \tag{6}$$

Here υ is the shape, and σ_l and σ_r are the spread parameters of the left (negative) and right (positive) sides of the model density.

Following [31], we estimate the GGD parameters (α, σ) and the AGGD parameters $(\upsilon, \sigma_l, \sigma_r)$ using the moment matching technique in [47]. For each coefficient product image, a mean parameter is also computed:

$$\eta = (\beta_r - \beta_l)\frac{\Gamma(2/\upsilon)}{\Gamma(1/\upsilon)}. \tag{7}$$

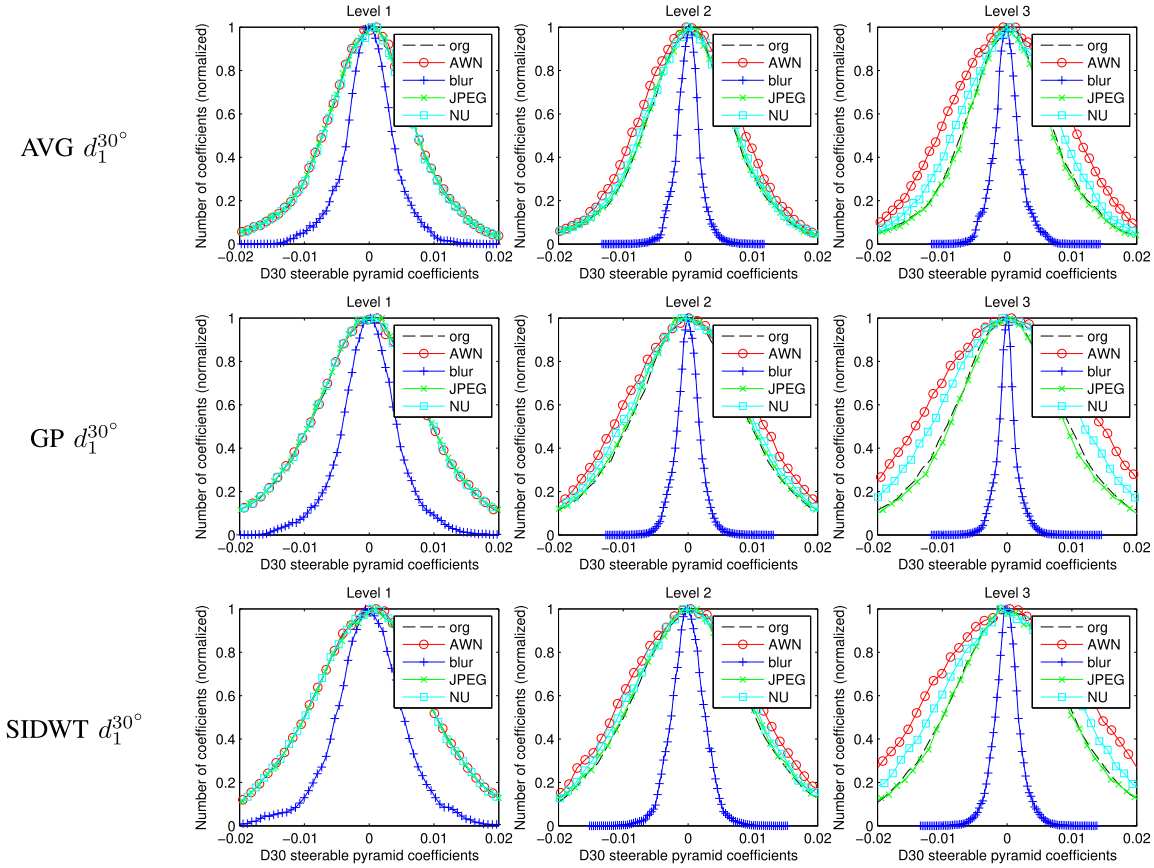


Fig. 6. Comparison of $d_1^{30^\circ}$ steerable pyramid histograms of 154 ROIs from fused images. The ROI sizes were 64×64 pixels. The figures show increasing levels of distortion from left to right for AWGN, blur, JPEG compression, and NU. In some cases the plots of the histograms of "org" are obscured by overlap of other curves.

TABLE I

FEATURE SUMMARY FOR MSCN (f), PAIRED PRODUCTS (pp), PAIRED LOG-DERIVATIVES (pd), AND STEERABLE PYRAMID COEFFICIENTS (sp)

Feature ID	Feature Description	Computation Procedure
$f1$ - $f2$	α and σ	GGD fit to $\tilde{I}$
$f3$ - $f4$	$\alpha_r - \alpha_l$ and $\sigma_r - \sigma_l$	GGD fit to right and left halves of $\tilde{I}$
$pp1$ - $pp4$	υ, η, σ_l, σ_r	AGGD fit to H
$pp5$ - $pp8$	υ, η, σ_l, σ_r	AGGD fit to V
$pp9$ - $pp12$	υ, η, σ_l, σ_r	AGGD fit to $D1$
$pp13$ - $pp16$	υ, η, σ_l, σ_r	AGGD fit to $D2$
$pd1$ - $pd2$	α and σ	GGD fit to $PD1$
$pd3$ - $pd4$	α and σ	GGD fit to $PD2$
$pd5$ - $pd6$	α and σ	GGD fit to $PD3$
$pd7$ - $pd8$	α and σ	GGD fit to $PD4$
$pd9$ - $pd10$	α and σ	GGD fit to $PD5$
$pd11$ - $pd12$	α and σ	GGD fit to $PD6$
$pd13$ - $pd14$	α and σ	GGD fit to $PD7$
$sp1$ - $sp2$	α and σ	GGD fit to $d_1^{0^\circ}$
$sp3$ - $sp4$	α and σ	GGD fit to $d_1^{30^\circ}$
$sp5$ - $sp6$	α and σ	GGD fit to $d_1^{60^\circ}$
$sp7$ - $sp8$	α and σ	GGD fit to $d_1^{90^\circ}$
$sp9$ - $sp10$	α and σ	GGD fit to $d_1^{120^\circ}$
$sp11$ - $sp12$	α and σ	GGD fit to $d_1^{150^\circ}$

Hence, four parameters (υ, σ_l, σ_r, and η) are extracted from the histograms of the adjacent products of MSCN coefficients. At a single scale of processing, we thereby obtain 46 features per image, as summarized in Table I. These features are all computed over three scales: the initial image scale, and other two scales reduced by factors of two and four along each dimension, yielding a total of 138.

As a way of visualizing the features and the way that they cluster in response to the presence of distortion, we projected an exemplar set onto a two-dimensional space using Principal Component Analysis (PCA). Fig. 7 depicts the two-dimensional PC space of features extracted from all of the fused images contained in the databases, and for each of the considered fusion algorithms. As may be seen, the positions and variances of the clusters generated by the average, gradient pyramid, and SIDWT fusion algorithms suggest that these approaches produce stable and consistent features. However, the features produced by the Laplacian and ratio pyramid fusion algorithms resulted in less consistent clusters.

Fig. 8 shows the same features for all the images and fusion algorithms plotted together. It may be observed that in Figures 8a and 8b, sub-clusters formed corresponding to each database, presumably due to differences between the LWIR sensor technologies. Fig. 8c plots features computed on pristine and distorted fused images, labeled according to the types of distortions. The OSU database LWIR images were captured using a ferro-electric thermal sensor that follows a non-linear function of intensity, which may affect the NSS features extracted from the images. Nonetheless, as shown in Fig.8c,

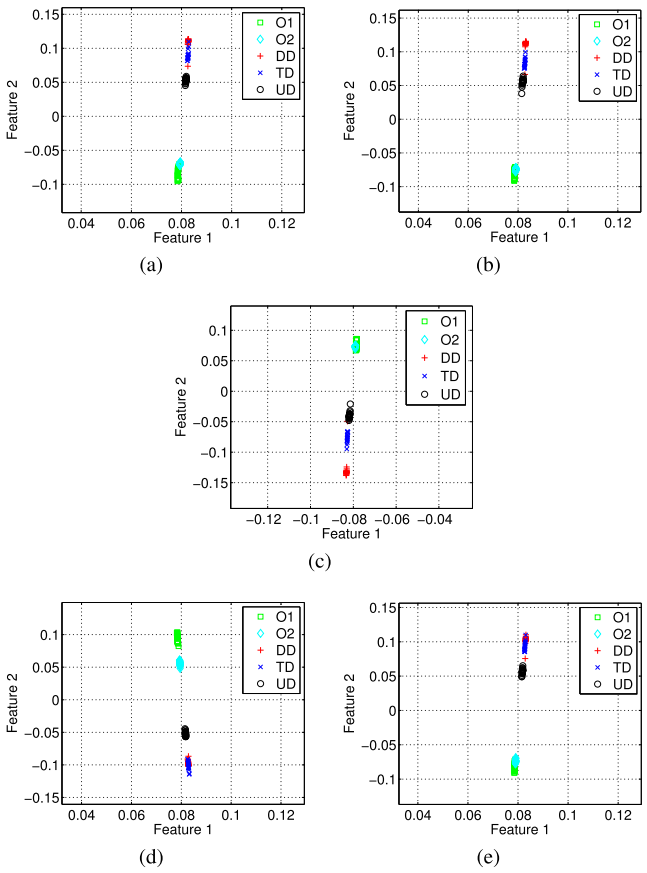


Fig. 7. Clustering of principal components of features extracted from images produced by different infrared/visible light fusion algorithms. (a) Average. (b) Gradient Pyramid. (c) Laplacian Pyramid. (d) Ratio Pyramid. (e) SIDWT with Haar wavelet. The labels O1 and O2 refer to two scenes from the OSU database, while the labels DD, TD, UD refer to three scenes from the TNO database.

features from distorted images still appear to cluster away from the pristine images.

III. Quality Assessment of Fused LWIR and Visible Images

A. Subjective Study

Because such a resource was not already available, we conducted a human quality perceptual study, which we used to both create a trained opinion aware IQA model (later described in subsection III-C), and as a tool to assess how well fused image quality prediction models perform on fused LWIR and visible light images, as measured by how well they correlate with subjective judgments. Later, we report the experimental protocol of the study, and the method of processing of the opinion scores.

To avoid fatiguing the human subjects with too many images to evaluate, we selected a total of 25 pairs of pristine LWIR and visible images, including 11 pairs from the TNO database and 14 pairs from the MORRIS database. The images were processed using three types of distortion, three levels of each distortion and three fusion methods. The image pairs were first processed by three levels of each of three types of simulated distortion: additive white Gaussian noise and

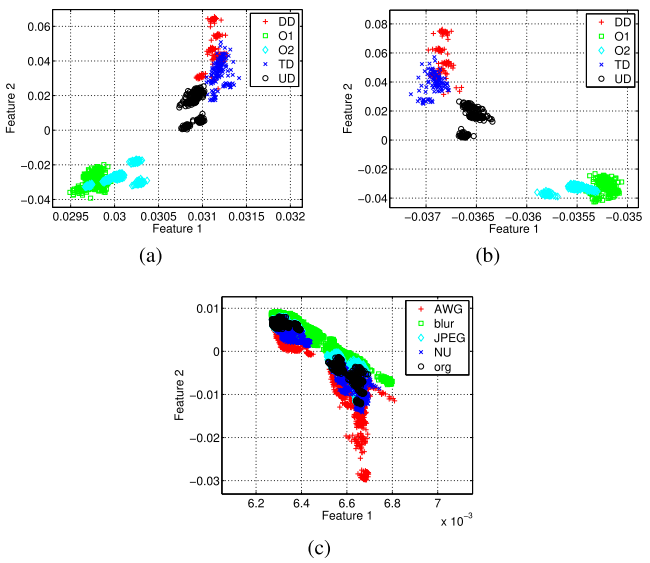


Fig. 8. A total of 138 features extracted from all of the images from the OSU and TNO databases are projected in a 2D space using PCA with a cumulative variance of 0.9973. (a) Features extracted only from pristine Visible light, LWIR, and fused images. (b) Features extracted only from fused images. (c) Features extracted from both pristine and distorted fused images. The labels O1 and O2 refer to pictures from the OSU database, while the terms DD, TD, UD refer to pictures from the TNO database. The terms AWG, blur, JPEG, NU refer to the image distortions, while org represents the pristine images.

blur were applied to both the LWIR and the visible light images, while non uniformity distortion was applied only on the LWIR images. For AWGN and NU, the distortion level was controlled using a standard deviation parameter $\sigma_{AWGN} = \sigma_{NU} = \{0.0025, 0.01375, 0.025\}$, while blur was applied using 15×15 Gaussian kernels having spread parameters $\sigma_{blur} = \{1, 2, 3\}$ pixels. We chose and applied the fusion methods that we judged to best preserve cluster stability and consistency: the average, the gradient pyramid, and the SIDWT, as shown in Fig. 7.

We conducted the study on 27 volunteers. Each person was asked to evaluate the images using a procedure we wrote using the Matlab Psychophysics Toolbox [48]. Each subject evaluated 150 single stimulus images over each of five testing sessions, yielding a total of 750 judged images apiece. The test procedure was conducted following the recommendations in [49], where the authors used a variant of the absolute category rating with hidden reference (ACR-HR) from ITU-T Rec. P.910, where each original image is included in the experiment but not identified as such. The screen resolution was 1024×768 and the stimulus images were displayed at their native resolution, at a viewing distance that varied between 45cm and 55cm. In each session, the images were randomly shuffled, then sequentially displayed for 7 seconds each as depicted in Fig. 9. Immediately following, the subject rated the image on a continuous sliding quality bar with Likert-like labels “Bad”, “Poor”, “Fair”, “Good”, and “Excellent”, as shown in Fig. 10. The recorded scores were sampled and converted to integer values on [0, 100]. In addition, we measured the illumination levels during the tests, which varied between 220 and 240 lux, ensuring that they did not change significantly between sessions.

Fig. 9. Example stimulus.

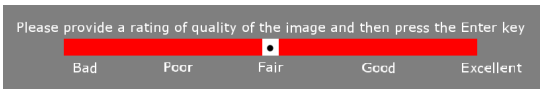
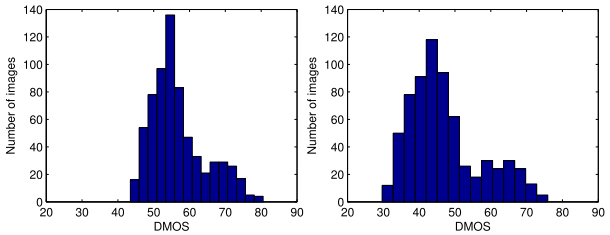


Fig. 10. Sliding quality bar.



(a) (b)

Fig. 11. Histograms of DMOS in 15 equally spaced bins for (a) scores obtained before subject rejection and (b) scores obtained after subject rejection.

The obtained subjective scores were then processed to discount individual preferences for images and differences in image content, as explained in [31] and [49]. First, we computed difference scores, defined as the difference between the score of each image and the score of its hidden reference, which were then converted to Z-scores and combined to build a matrix $[Z_{ij}]$ with elements corresponding to the Z-score given by subject i to image j. After obtaining the matrix, a subject rejection procedure was applied to discard unreliable scores, as specified in ITU-R BT 500.11 [49], [50].

Following this procedure, five outliers were found and removed, and the remaining Z-scores were linearly rescaled to $[0, 100]$. Finally, the Difference Mean Opinion Scores (DMOS) of each image were computed as the mean of the rescaled Z-scores for the remaining 22 subjects [31], [49].

Histograms of the DMOS are shown in Fig. 11, which indicate a fairly broad distribution of the DMOS. Scores before subject rejection fell within the range $[45, 80]$, while scores after subject rejection fell within $[31, 74]$ yielding a wider range of visual quality. For DMOS obtained after subject rejection, it should be noted that most of the subject evaluations were distributed over about half of the quality range. In the following section, we study the performance obtained when fusing pristine images by comparing them to Mean Opinion Scores (MOS).

B. Opinion-Distortion-Unaware Image Quality Analyzer

We have developed a 'completely blind' (opinion and distortion unaware) model of the perceptual quality of a fused image. To compute it, 80 pristine image pairs from the OSU database were used, with sizes of 298 × 217 pixels. These were fused (using the three aforementioned fusion algorithms), yielding 240 fused pristine images. We then extracted 138 'quality-aware' NSS features from each image, from which our pristine fused picture model is obtained. This model was calculated by fitting the features to a multivariate Gaussian model, as was done in the NIQE [34] and feature-enriched IL-NIQE [51] models. The model consists of a mean vector μ and a covariance matrix Σ; using these, it is possible to evaluate the quality of an image by comparing the pristine model to a similarly constructed model of the degraded image. The quality prediction score is then calculated as the (modified) Mahalanobis distance between the previously constructed pristine feature model and the feature model of the distorted image:

$$Q_D(\mu_1, \mu_2, \Sigma_1, \Sigma_2) = \sqrt{(\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1} (\mu_1 - \mu_2)}, \quad (8)$$

where μ_1, μ_2 and Σ_1, Σ_2 are the mean vectors and covariance matrices of the models obtained using the standard maximum likelihood estimation procedure in [52].

To verify the performance of our model, we compared the scores given to pristine fused images to the fusion quality model predictions shown in Table II, which were previously studied by Liu *et al.* in [22]. We also evaluated the performance of each of the individual feature groups (f, pp, pd, and sp) using the same approach, obtaining quality models similar to Q_D that we denote Q_f, Q_{pp}, Q_{pd}, and Q_{sp}. Previous studies have compared the resulting fusion performance predictions to subjective judgments of small sets of fused images [10], [13]. Here we analyze how well the predictions correlate to the subjective scores of pristine fused images. First, we computed the Z-score from each raw opinion

TABLE II
DESCRIPTION OF THE FUSION PERFORMANCE MODELS STUDIED IN [22]

Category	Model	Description
Information Theory	Q_{MI}	Normalized Mutual Information [10]
	Q_{TE}	Fusion Metric-Based on Tsallis Entropy [11]
	Q_{NCIE}	Nonlinear Correlation Information Entropy [12]
Image Features	Q_G	Gradient-Based Fusion Performance [13]
	Q_M	Image Fusion Metric-Based on a Multiscale Scheme [14]
	Q_{SF}	Image Fusion Metric-Based on Spatial Frequency [15]
	Q_P	Image Fusion Metric-Based on Phase Congruency [16]
Structural Similarity	Q_S	Piella's Metric [17]
	Q_C	Cvejic's Metric [19]
	Q_Y	Yang's Metric [18]
Human Perception	Q_{CV}	Chen-Varshney Metric [20]
	Q_{CB}	Chen-Blum Metric [21]

score s, then rescaled them to fill the range $[0, 1]$. Since DMOS can only be computed as the difference between the scores given to a pristine image and the scores given to a distorted image, we calculated Mean Opinion Scores (MOS) for each pristine fused image after removing outliers. In order to account for a possible non linear relationship between the quality predictions and MOS, the algorithm scores were passed through the following logistic function:

$$Q'_j = \beta_2 + \frac{\beta_1 - \beta_2}{1 + \exp^{-(Q_j - \beta_3/|\beta_4|)}} \tag{9}$$

where Q_j is the objective quality value for stimulus image j. Each β parameter was estimated via nonlinear least squares optimization using the Matlab function "`nlinfit`," to minimize the least squares error between MOS_j and the fitted scores Q_j. To facilitate numerical convergence, the quality predictions were first linearly rescaled before performing optimization. We chose the initial β parameters following the recommendation in [53]:

$$\beta_1 = \max(\text{MOS}) \tag{10}$$

$$\beta_2 = \min(\text{MOS}) \tag{11}$$

$$\beta_3 = \bar{Q} \tag{12}$$

$$\beta_4 = 1 \tag{13}$$

The new values Q'_j were used to compute the Spearman's Rank Correlation Coefficient (SRCC), Pearson's Linear Correlation Coefficient (LCC), and Root Mean Squared Error (RMSE) between samples. SRCC was deployed as a measure of non-linear monotonicity between the reference and predicted values, while LCC was used as a measure of linear correlation between actual and predicted values. RMSE evaluates the accuracy of the predictions. An effective objective quality measure would have SRCC and LCC closer to one and RMSE nearer to zero. The results from the evaluation are shown in Table III. Some of the resulting correlation coefficients yielded a negative sign, therefore we give their absolute values. In this comparison, we included the NR IQA algorithm IL-NIQE [51], whose pristine fused image features are extracted from the same set of images deployed for the models Q_D, Q_f, Q_{pp}, Q_{pd}, and Q_{sp}. The models that produced negative correlation coefficients include the spatial frequency metric Q_{SF} [17], the Chen-Blum metric Q_{CB} [21], IL-NIQE, and our models, which are based on NSS features. We observe that the Q_{pd} model provided the highest correlations with respect to subjective image quality judgments, followed by IL-NIQE and Q_D.

TABLE III

ABSOLUTE VALUES OF SRCC, LCC, AND RMSE BETWEEN MOS AND PREDICTED MOS PRODUCED BY VARIOUS QUALITY PREDICTION MODELS ON THE EVALUATED PRISTINE FUSED IMAGES

	Model	SRCC	LCC	RMSE
Models studied in [22]	Q_{MI}	0.561	0.557	0.054
	Q_{TE}	0.257	0.339	0.069
	Q_{NCIE}	0.586	0.581	0.073
	Q_G	0.398	0.491	0.068
	Q_M	0.214	0.301	0.076
	Q_{SF}	0.271	0.261	0.087
	Q_P	0.585	0.591	0.059
	Q_S	0.532	0.548	0.081
	Q_C	0.331	0.344	0.098
	Q_Y	0.525	0.536	0.084
	Q_{CV}	0.462	0.460	0.081
	Q_{CB}	0.302	0.328	0.095
	IL-NIQE	0.677	0.619	0.115
Feat. Groups	Q_D	0.650	0.618	0.115
	Q_{sp}	0.343	0.282	0.046
	Q_{pd}	**0.705**	**0.698**	**0.046**
	Q_{pp}	0.573	0.556	0.127
	Q_f	0.602	0.569	0.116

C. Opinion Aware Fused Image Quality Analyzer

As mentioned earlier, we also created an opinion aware (but otherwise blind) model by training on human subjective quality judgments of the images. To do this we employed a Support Vector Regression (SVR) algorithm to fit the NSS features to the DMOS, thereby obtaining a trained opinion aware quality model Q_{SVR}. This method has been previously applied to IQA using NSS-based features [31], [33]. We utilized the LIBSVM package [54] to implement an ϵ-SVR with a Radial Basis Function kernel, and found the best-fitting parameters C and γ using 5-fold cross-validation. The best-fitting parameters obtained were $C = 2^{11.5}$ and $\gamma = 2^{-5}$, which yielded a mean squared error $MSE = 7.59$. Our experiments were carried out by using the 675 distorted fused images and their corresponding DMOS from the subjective study.

In order to test our quality measure Q_{SVR}, we included BRISQUE in the set of quality measures to be evaluated with respect to their correlation to human judgments. BRISQUE was modified by using the same pristine fused image set and quality scores deployed in Q_{SVR}. Since Q_{SVR} and BRISQUE require a training procedure to calibrate, we divided the data from the subjective study into two random subsets, where 80% of the fused images and associated DMOS were used for training and 20% for testing, taking care not to overlap the train and test content. This was done to ensure that the results would not depend on features extracted from learned content, rather than from distortion. The predicted scores were then passed through the logistic non-linearity described in the previous section.

We repeated this process over 1000 iterations, computed SRCC, LCC, and RMSE for all models, and tabulated their median values in Table IV. The median of the SRCC, LCC, and RMSE values is not skewed much by extremely large or small values, thereby providing a robust figure of merit. For measures other than Q_{SVR} and BRISQUE, we used 80% of the data to estimate the β parameters, and the other 20% to validate the prediction of the logistic function. The results given by the validation were used to compute the correlation coefficients. Fig. 12 depicts a scatter plot of the predicted scores delivered by our quality model Q_{SVR} versus DMOS for all the images evaluated in the subjective study described in subsection III-A, along with the best-fitting logistic function. Observe that the model Q_{SVR} achieved the highest correlation against the human scores, followed by BRISQUE, while the other models yielded lower correlations. Notice that in this

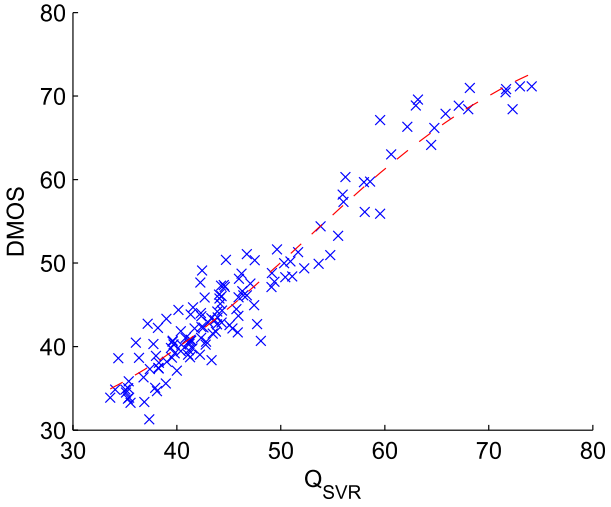


Fig. 12. Scatter plot of Q_{SVR} prediction scores versus DMOS for all images assessed in the subjective human study and the best fitting logistic function. Notice the near-linear relationship.

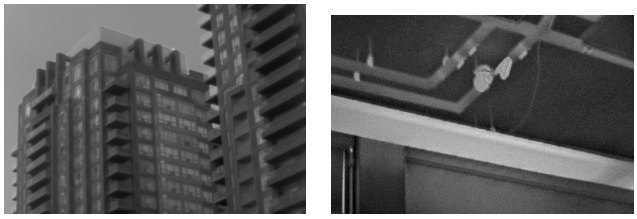
case, when comparing to scores given to distorted images, Q_M outperformed all of the fused image quality measures classified as Feature groups, further motivating the use of quality aware measures such as Q_{SVR} and BRISQUE. Even though the comparison between ODU and OA image quality measures may be regarded as biased because of the deployment of subjective scores in the OA approaches, it is important to note that the state-of-the-art measures of image fusion performance studied in [22] are ODU since, unlike Q_{SVR}, those fusion performance models do not involve a training process linked to subjective scores. Nonetheless, in Table IV, we included the OA image quality measure BRISQUE, which uses the same pristine fused image set and quality scores deployed in Q_{SVR}. Table III tabulates the SRCC, LCC, and RMSE between MOS and predicted MOS for pristine images, while Table IV presents the median SRCC, LCC, and RMSE between DMOS and predicted DMOS for distorted images.

We carried out a one-sample Kolmogorov-Smirnov test to establish whether the 1000 SRCC values from the predictions of 20 fused quality measures (20.000 SRCC values) come from a standard normal distribution (i.e. null hypothesis). This test rejected the null hypothesis at the 5% significance level. Hence, since nonparametric tests make no assumptions about the probability distributions of the variables, we conducted a Kruskal-Wallis test on each median value of SRCC between the DMOS and the quality measures (after nonlinear mapping), to evaluate whether the results presented in Table IV are statistically significant. Table V tabulates the results of the statistical significance test. The null hypothesis was that the median correlation for the (row) algorithm was equal to the median correlation for the (column) algorithm with a confidence of 95%. The alternate hypothesis was that the median correlation of the row was greater than or less than the median correlation of the column. From Table V, we conclude that Q_{SVR} produced highly competitive quality predictions on the tested fused pictures with statistical significance against all of the other quality algorithms tested.

TABLE IV

MEDIAN SRCC, LCC, AND RMSE BETWEEN DMOS AND PREDICTED DMOS OVER 1000 ITERATIONS

	Model	Median SRCC	Median LCC	Median RMSE
Models studied in [22]	Q_{MI}	0.185	0.241	9.785
	Q_{TE}	0.042	0.023	10.082
	Q_{NCIE}	0.194	0.270	9.713
	Q_G	0.216	0.252	9.754
	Q_M	0.585	0.798	6.085
	Q_{SF}	0.159	0.194	9.891
	Q_P	0.066	0.209	9.854
	Q_S	0.267	0.303	9.614
	Q_C	0.274	0.317	9.547
	Q_Y	0.172	0.278	9.681
	Q_{CV}	0.042	0.050	10.067
	Q_{CB}	0.070	0.079	10.048
	IL-NIQE	0.159	0.253	9.744
	BRISQUE	0.855	0.894	4.476
Feat. Groups	Q_D	0.264	0.405	10.032
	Q_{SP}	0.095	0.095	10.041
	Q_{PD}	0.224	0.259	9.742
	Q_{PP}	0.107	0.164	9.949
	Q_F	0.077	0.126	9.999
SVR	Q_{SVR}	**0.932**	**0.961**	**2.813**



(a) (b)

Fig. 13. Examples of best (a) and worst (b) rated images from the subjective study.

IV. RESULTS AND DISCUSSION

We found that fused LWIR-visible images created using multi-resolution fusion algorithms such as Average, Gradient Pyramid, Laplacian Pyramid, Ratio Pyramid, and SIDWT, possess statistical regularities when band-pass filtered and divisively normalized, and that these regularities can be modeled and used to characterize distortions and to predict fused image quality. As shown through the histogram analysis, some groups of NSS features are more predictive of some types of distortion than the rest: PD coefficients effectively responded to JPEG compression and blur distortions, while $d_1^{0^\circ}$ and $d_1^{90^\circ}$ were effective for measuring NU distortion. Furthermore, we developed both opinion-distortion-unaware ('completely blind') and opinion-aware image quality analyzers, which predict human quality evaluations of fused LWIR and visible images more reliably than other state-of-the-art models.

One limitation of our research was the limited availability of aligned LWIR and visible image pairs. The OSU database contains little image content diversity, making it unsuitable for inclusion in the subjective study. Other databases that we examined did not provide registered visible and infrared

TABLE V

STATISTICAL SIGNIFICANCE MATRIX OF SRCC BETWEEN DMOS AND PREDICTED QUALITY SCORES. A VALUE OF "1" INDICATES THAT THE PERFORMANCE OF THE MODEL IN THE ROW WAS STATISTICALLY BETTER THAN THAT OF THE MODEL IN THE COLUMN, "0" MEANS THAT IT IS STATISTICALLY WORSE, AND "-" MEANS THAT IT IS STATISTICALLY INDISTINGUISHABLE

	Q_{MI}	Q_{TE}	Q_{NCIE}	Q_G	Q_M	Q_{SF}	Q_P	Q_S	Q_C	Q_Y	Q_{CV}	Q_{CB}	IL-NIQE	BRISQUE	Q_D	Q_{SP}	Q_{PD}	Q_{PP}	Q_F	Q_{SVR}
Q_{MI}	-	1	0	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0
Q_{TE}	0	-	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0
Q_{NCIE}	1	1	-	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0
Q_G	1	1	1	-	0	1	1	0	0	1	1	1	1	0	1	1	-	1	1	0
Q_M	1	1	1	1	-	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0
Q_{SF}	0	1	0	0	0	-	1	0	0	0	1	1	-	0	1	1	0	1	1	0
Q_P	0	1	0	0	0	0	-	0	0	0	1	-	0	0	0	0	0	0	0	0
Q_S	1	1	1	1	0	1	1	-	0	1	1	1	1	0	1	1	1	1	1	0
Q_C	1	1	1	1	0	1	1	1	-	1	1	1	1	0	1	1	1	1	1	0
Q_Y	0	1	0	0	0	1	1	0	0	-	1	1	1	0	1	1	0	1	1	0
Q_{CV}	0	-	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0
Q_{CB}	0	1	0	0	0	0	-	0	0	0	1	-	0	0	0	0	0	0	0	0
IL-NIQE	0	1	0	0	0	-	1	0	0	0	1	1	-	0	1	1	0	1	1	0
BRISQUE	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	0
Q_D	0	1	0	0	0	0	1	0	0	0	1	1	0	0	-	1	0	-	1	0
Q_{SP}	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	-	0	0	1	0
Q_{PD}	1	1	1	-	0	1	1	0	0	1	1	1	1	0	1	1	-	1	1	0
Q_{PP}	0	1	0	0	0	0	1	0	0	0	1	1	0	0	-	1	0	-	1	0
Q_F	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	-	0
Q_{SVR}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-

images. Nonetheless, the results provided by our opinion-aware quality analyzer outperformed all the other fusion quality algorithms, while having the advantage of not needing source LWIR and visible light images like the models studied in [22] to compute a quality estimate. Q_{SVR}, as for any other OA method, requires training on human evaluations. Moreover, it is unable to provide a quality map of the image, where each pixel would represent an image quality value. In Fig. 13, we show examples of the best and worst images rated according to the DMOS obtained in the subjective study, with the best image having a DMOS of 29.659 and the worst image having a DMOS of 75.920.

Previous studies of fused image quality have not accounted for the presence of distortion in the source images, or even of LWIR-specific distortions. Moreover, in some cases the authors evaluated their proposed models using very limited sets of image pairs [10]–[13], [20]. Although the work in [22] assessed AWGN and blur distortions, our approach also considers the effects of NU, and proposes a fusion quality model that analyzes image degradation. To our knowledge, there has been no prior work on the analysis of NSS extracted from pristine and distorted fused LWIR and visible images. We believe that this work can serve as a solid starting point for further development of perceptual quality aware fusion algorithms.

V. CONCLUSION AND FUTURE WORK

NSS play an important role when analyzing distortions present in fused LWIR and visible light images, as they have previously proved useful in modeling degradations of visible and infrared pictures. We found that NSS are also potent descriptors of the quality of fused images affected by AWGN and NU. Therefore, we proposed ODU and OA fused image quality analyzers that outperform current fusion quality indexes, correlating better with human subjective evaluations. Although a broader spectrum of distortion types would have allowed deeper insights, it would have lengthened the duration of the human study to an unacceptable degree. Future studies might be able to use the proposed models to evaluate other distortions present in infrared images, and by using scene statistics of fused images measured on other types of image sensors. Furthermore, fused LWIR-visible videos used in surveillance applications are of great interest. These videos could be modeled and studied with the aid of spatio-temporal NSS to improve tracking algorithms.

ACKNOWLEDGMENT

The authors would like to thank Z. Liu for providing the source code of their fused image quality evaluation algorithms.

REFERENCES

[1] J. Lee and A. Bovik, "Video surveillance," in *The Essential Guide to Video Processing*. Amsterdam, The Netherlands: Elsevier, 2009, ch. 19, pp. 619–649.

[2] A. Benkhalil, S. Ipson, and W. Booth, "Real-time video surveillance system using a field programmable gate array," *Int. J. Imag. Syst. Technol.*, vol. 11, no. 2, pp. 130–137, 2000.

[3] G. L. Foresti, "A real-time system for video surveillance of unattended outdoor environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 6, pp. 697–704, Oct. 1998.

[4] F. Khodayar, S. Sojasi, and X. Maldague, "Infrared thermography and NDT: 2050 horizon," *Quant. Infr. Thermogr. J.*, vol. 13, no. 2, p. 246, 2016.

[5] M. Freebody, "Consumers and cost are driving infrared imagers into new markets," *Photon. Spectra*, vol. 49, no. 4, pp. 40–44, Apr. 2015.

[6] A. Torabi, G. Massé, and G.-A. Bilodeau, "An iterative integrated framework for thermal–visible image registration, sensor fusion, and people tracking for video surveillance applications," *Comput. Vis. Image Understand.*, vol. 116, no. 2, pp. 210–221, 2012.

[7] J. Han and B. Bhanu, "Fusion of color and infrared video for moving human detection," *Pattern Recognit.*, vol. 40, no. 6, pp. 1771–1784, 2007.

[8] A. El Maadi and X. Maldague, "Outdoor infrared video surveillance: A novel dynamic technique for the subtraction of a changing background of IR images," *Infr. Phys. Technol.*, vol. 49, no. 3, pp. 261–265, 2007.

[9] E. P. Bennett, J. L. Mason, and L. McMillan, "Multispectral bilateral video fusion," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1185–1194, May 2007.

[10] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electron. Lett.*, vol. 38, no. 7, pp. 313–315, Mar. 2002.

[11] N. Cvejic, C. N. Canagarajah, and D. R. Bull, "Image fusion metric based on mutual information and Tsallis entropy," *Electron. Lett.*, vol. 42, no. 11, pp. 626–627, May 2006.

[12] Q. Wang, Y. Shen, and J. Jin, *Performance Evaluation of Image Fusion Techniques*. Amsterdam, The Netherlands: Elsevier, 2008, ch. 19, pp. 469–492.

[13] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, 2000.

[14] P.-W. Wang and B. Liu, "A novel image fusion metric based on multi-scale analysis," in *Proc. IEEE Int. Conf. Signal Process.*, Oct. 2008, pp. 965–968.

[15] Y. Zheng, E. A. Essock, B. C. Hansen, and A. M. Haun, "A new metric based on extended spatial frequency and its application to DWT based fusion algorithms," *Inf. Fusion*, vol. 8, no. 2, pp. 177–192, Apr. 2007.

[16] J. Zhao, R. Laganiere, and Z. Liu, "Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement," *Int. J. Innov. Comput., Inf. Control*, vol. 3, no. 6, pp. 1433–1447, Dec. 2007.

[17] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. Image Process.*, 2003, pp. III-173–III-176.

[18] C. Yang, J.-Q. Zhang, X.-R. Wang, and X. Liu, "A novel similarity based quality metric for image fusion," *Inf. Fusion*, vol. 9, no. 2, pp. 156–160, 2008.

[19] N. Cvejic, A. Loza, D. Bull, and N. Canagarajah, "A similarity metric for assessment of image fusion algorithms," *Int. J. Signal Process.*, vol. 2, no. 3, pp. 178–182, 2005.

[20] H. Chen and P. K. Varshney, "A human perception inspired quality metric for image fusion based on regional information," *Inf. Fusion*, vol. 8, no. 2, pp. 193–207, 2007.

[21] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, 2009.

[22] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

[23] N. Rajic, *Nondestructive Testing Handbook: Infrared and Thermal Testing*, vol. 3. Columbus, OH, USA: American Society for Nondestructive Testing, 2001, ch. 5.

[24] J. W. Davis and V. Sharma, "Background-subtraction in thermal imagery using contour saliency," *Int. J. Comput. Vis.*, vol. 71, no. 2, pp. 161–181, 2007.

[25] I. B. Schwartz, K. A. Snail, and J. R. Schott, "Infrared halo effects around ships," DTIC Document, Naval Res. Lab., Washington DC, USA, NRL Memorandum Rep. 5529, 1985.

[26] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 838–842, Jul. 2015.

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

[28] A. K. Moorthy and A. C. Bovik, "Visual quality assessment algorithms: What does the future hold?" *Multimedia Tools Appl.*, vol. 51, no. 2, pp. 675–696, 2011.

[29] 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.

[30] C.-C. Su, A. C. Bovik, and L. K. Cormack, "Natural scene statistics of color and range," in *Proc. 18th IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2011, pp. 257–260.

[31] T. R. Goodall, A. C. Bovik, and N. G. Paulter, Jr., "Tasking on natural statistics of infrared images," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 65–79, Jan. 2016.

[32] N. J. W. Morris, S. Avidan, W. Matusik, and H. Pfister, "Statistics of infrared images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–7.

[33] 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.

[34] 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.

[35] D. E. Moreno-Villamarín, H. D. Benítez-Restrepo, and A. C. Bovik, "Statistics of natural fused image distortions," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 1243–1247.

[36] J. W. Davis and M. A. Keck, "A two-stage template approach to person detection in thermal imagery," in *Proc. 7th IEEE Workshops Appl. Comput. Vis. (WACV/MOTION)*, Jan. 2005, pp. 364–369.

[37] A. Toet, J. K. IJspeert, A. M. Waxman, and M. Aguilar, "Fusion of visible and thermal imagery improves situational awareness," *Displays*, vol. 18, no. 2, pp. 85–95, 1997.

[38] A. Toet, M. Hogervorst, H. Lensen, K. Benoist, and R. de Rooy, "ATHENA: The combination of a brightness amplifier and thermal viewer with color," DTIC Document, TNO, Soesterberg, The Netherlands, Tech. Rep. TNO-DV 2007 A329, 2007.

[39] J. E. Pezoa and O. J. Medina, "Spectral model for fixed-pattern-noise in infrared focal-plane arrays," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Berlin, Germany: Springer, 2011, pp. 55–63.

[40] R. S. Blum and Z. Liu, Eds., *Multi-Sensor Image Fusion and Its Applications*. Boca Raton, FL, USA: CRC Press, 2005.

[41] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," *Proc. ICCV*, 1993, pp. 173–182.

[42] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA Eng.*, vol. 29, no. 6, pp. 33–41, 1984.

[43] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognit. Lett.*, vol. 9, no. 4, pp. 245–253, 1989.

[44] O. Rockinger and T. Fechner, "Pixel-level image fusion: The case of image sequences," *Proc. SPIE*, vol. 7, pp. 378–388, Jul. 1998.

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

[46] Y. Zhang and D. M. Chandler, "An algorithm for no-reference image quality assessment based on log-derivative statistics of natural scenes," *Proc. SPIE*, vol. 10, p. 86530J, Feb. 2013.

[47] 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.

[48] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[49] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1427–1441, Jun. 2010.

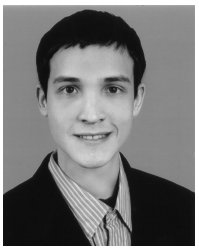
[50] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, document Rec. BT.500-11, ITU-R, International Telecommunication Union, 2002.

[51] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.

[52] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 4. New York, NY, USA: Springer, 2006.

[53] (2000). *Final Report From the Video Quality Experts Group on the Validation of Objective Quality Metrics for Video Quality Assessment, Phase I*. [Online]. Available: http://www.its.bldrdoc.gov/vqeg/projects/frtv-phase-i/frtv-phase-i.aspx

[54] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm



David Eduardo Moreno-Villamarín (S'16) received the B.S. degree (Hons.) in electronics engineering for his undergraduate project and academic performance from Pontificia Universidad Javeriana, Cali, Colombia, in 2017. His main research interests include computer vision, image and video processing, and image quality assessment.

Hernán Darío Benítez-Restrepo (S'05–SM'14) received the B.S. degree in electronics engineering from Pontificia Universidad Javeriana, Cali, Colombia, in 2002, and the Ph.D. degree in electrical engineering from Universidad del Valle, Cali, in 2008. Since 2008, he has been with the Department of Electronics and Computing, Pontificia Universidad Javeriana Sede Cali. Since 2010, he has been an Adjunct Professor with the Laboratory of Computer Vision and Systems, Université Laval, Québec City, Canada. In 2011, he received the Merit Scholarship for short-term research from the Ministére de l'Education, du Québec to pursue research on infrared vision at the Laboratory of Computer Vision and Systems, Université Laval. He has been the Chair of Colombia's IEEE Signal Processing Chapter, since 2012. He is member of the Scientific Editorial Board of the *Quantitative Infrared Thermography Journal* since 2014. His main research interests encompass image and video quality assessment, infrared vision, and digital signal processing. He is member of SPIE.

Alan Conrad Bovik (F'96) is currently the Cockrell Family Endowed Regents Chair of Engineering with The University of Texas at Austin, where he is the Director of the Laboratory for Image and Video Engineering. He is a Faculty Member of the Department of Electrical and Computer Engineering and the Institute for Neuroscience. He has authored over 750 technical articles in his research areas and holds several U.S. patents. He has authored or co-authored over 45,000 times in the literature, his current h-index of 75, and he is listed as a Highly-Cited Researcher by Thompson Reuters. His current research interests include image and video processing, computational vision, and visual perception. His several books include the companion volumes *The Essential Guides to Image and Video Processing* (Academic Press, 2009).

He received the Primetime Emmy Award for Outstanding Achievement in Engineering Development from the Academy of Television Arts and Sciences (The Television Academy) in 2015, for his contributions to the development of video quality prediction models which have become standard tools in broadcast and post-production houses throughout the television industry. He has also 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 was also 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 (2008) and the Joe J. King Professional Engineering Achievement Award (2015) 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 co-founded and was the longest-serving Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING (1996–2002), created and served as the General Chairman of the First IEEE International Conference on Image Processing, Austin, TX, in 1994, along with numerous other professional society activities, including the Board of Governors, the IEEE Signal Processing Society from 1996 to 1998; the Editorial Board of the IEEE Proceedings, 1998–2004; and a Series Editor of *Image, Video, and Multimedia Processing* (Morgan and Claypool Publishing Company, since 2003).

He is a registered Professional Engineer with the State of Texas and is a frequent Consultant to legal, industrial, and academic institutions.