Predicting Detection Performance on Security X-Ray Images as a Function of Image Quality

Praful Gupta, Zeina Sinno, Jack L. Glover, Nicholas G. Paulter, Jr., *Fellow, IEEE*, and Alan C. Bovik, *Fellow, IEEE*

***Abstract*— Developing methods to predict how image quality affects the task performance is a topic of great interest in many applications. While such studies have been performed in the medical imaging community, little work has been reported in the security X-ray imaging literature. In this paper, we develop models that predict the effect of image quality on the detection of the improvised explosive device components by bomb technicians in images taken using portable X-ray systems. Using a newly developed NIST-LIVE X-Ray Task Performance Database, we created a set of objective algorithms that predict bomb technician detection performance based on the measures of image quality. Our basic measures are traditional image quality indicators (IQIs) and perceptually relevant natural scene statistics (NSS)-based measures that have been extensively used in visible light image quality prediction algorithms. We show that these measures are able to quantify the perceptual severity of degradations and can predict the performance of expert bomb technicians in identifying threats. Combining NSS- and IQI-based measures yields even better task performance prediction than either of these methods independently. We also developed a new suite of statistical task prediction models that we refer to as quality inspectors of X-ray images (QUIX); we believe this is the first NSS-based model for security X-ray images. We also show that QUIX can be used to reliably predict conventional IQI metric values on the distorted X-ray images.**

***Index Terms*— NSS, X-ray images, task performance study, IQI prediction, IEEE/ANSI N42.55, image quality, improvised explosive devices (IEDs).**

I. Introduction

PORTABLE transmission X-ray imaging systems are used by military and civilian bomb technicians to screen suspicious packages and objects for explosives, bombs and other threat items contraband [1]. Their easy deployment and high detection efficiency makes them ideal for screening of hard-to-access places. The quality of the X-ray images captured by these systems serves as an important indicator of the manufacturing quality and overall performance of the imaging system. Several intrinsic and extrinsic factors affect the quality of X-ray images. The geometry of a portable X-ray imaging device, such as the size of detector photosensors and the generator's focal spot, has a strong influence on the quality of captured X-ray image. Photon-limited noise due to the inherent variation of photon influx at each photosensor is a major source of noise in X-ray images. Photon noise increases as a function of the square-root of the number of absorbed photons. The effects of this noise are reduced when the signal is greater than the noise (higher signal-to-noise-ratio or SNR) [2], which may be difficult to achieve in the field. This photon noise significantly affects the object detection and identification accuracy by either a human or an algorithm. There exist other factors that impact X-ray image quality, including but not limited to the voltage-current settings of the imager and the arrangement of the imaging device with respect to the object being imaged [3].

Rapid technological developments have enabled continuous improvements in the speed of acquisition and the quality of images produced by portable X-ray systems. Image quality greatly affects the abilities of trained professionals to make rapid and accurate decisions under challenging field conditions. The performance of these X-ray imagers is generally measured in terms of physical image quality parameters such as resolution, noise and SNR [4], [5]. While physical performance metrics are suitable measures to assess imaging system performance, the task performance of skilled bomb techs on images produced by these systems serve as an ultimate 'gold standard' indicator of system performance [6], [7]. Thus, it is essential to analyze how physical image quality measurements on a system correlate with the task performance of trained bomb technicians. The goals of our work, therefore, are to be able to better understand and model how image quality affects human task performance, to determine how this relationship can be used to create automatic perceptual X-ray image task prediction models that correlate well against human performance, and ultimately to create baseline performance metrics for image quality.

The task-based assessment of image quality has been studied in the medical imaging literature [8]–[10]. The fundamental motivation behind this idea is the dependence of the image quality of a system on the task performance of observers on some specific task. The observer can either be a human, such as a radiologist, or a model observer, such as a Bayesian

Manuscript received June 5, 2018; revised December 29, 2018; accepted January 15, 2019. Date of publication January 31, 2019; date of current version May 14, 2019. This work was supported in part by the National Institute of Technology under Grant 70NANB15H270. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Daniel L. Lau. *(Corresponding author: Praful Gupta.)*

P. Gupta, Z. Sinno, and A. C. Bovik are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78701 USA (e-mail: praful_gupta@utexas.edu; zeina@utexas.edu; bovik@ece.utexas.edu).

J. L. Glover is with the National Institute of Standards and Technology, Gaithersburg, MD 20899 USA, and also with Theiss Research, La Jolla, CA 92037 USA (e-mail: jlglover@nist.gov).

N. G. Paulter, Jr. is with the National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: nicholas.paulter@nist.gov).

Digital Object Identifier 10.1109/TIP.2019.2896488

1057-7149 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

ideal observer [7], [11], that makes best possible use of the knowledge of signal and noise properties. The task can be detection, identification or localization, depending on the purpose against which the imaging system is to be optimized [7]. More complex task-based model observers that take into account the properties of human visual system to assess image quality have been studied, such as the model observers that consider the contrast sensitivity function, the multi-scale and orientation selectivity of cortical neurons, and neural noise models to account for the intrinsic variability of human observers [12]–[14]. While the above models typically assume knowledge of a signal's precise location and may not represent realistic situations, such as clinical trials, models that detect a signal at an unknown location have also been proposed [15]. Model observers that can closely mimic human performance are beneficial to equipment designers for optimizing system design parameters for specific diagnostic tasks.

While a considerable vein of research in this direction has been developed in the medical imaging field, only a little work has focused on the visual task performance in security X-ray imaging. In [16] several factors, including sensitivity and response time, were studied on visual scanning and target detection tasks, where human observers were tasked with searching for a knife inserted at randomly different angles in chromatic X-ray images of cluttered baggages. In another work [17], participants were asked to identify improvised explosive devices (IEDs) in a brief presentation of suspicious baggages, and a model observer developed for a different medical imaging task was adapted to explain the observers' performance. To the best of our knowledge, there is no reported work that deals with task-based image quality assessment (IQA) of security relevant X-ray images. In this work, we contribute to solve this problem by analyzing the detection performance of expert observers on distorted X-ray images, and we build perceptual X-ray image quality models that reliably predict observers' task performance.

There exist internationally standardized methods for objectively measuring the quality of images produced by portable transmission X-ray systems [4], [5]. These objective quality metrics, which we will refer to as image quality indicators (IQIs), operate by making specific quantitative measurements on images of standard test objects obtained under highly specific test conditions. IEEE/ANSI N42.55[1] includes a detailed description of the measurement and performance requirements of these conventional IQIs, which include 'Useful penetration', 'Organic material detection', 'Spatial resolution', 'Dynamic range', 'Noise', 'Flatness of field', 'Image extent', 'Image area', and 'Aspect ratio'. While these IQIs do provide reliable measurements of image quality, their computation also involves the use of precisely defined test objects that are imaged under strictly defined laboratory conditions, which consumes significant amounts of time, cost and effort.

Recently, a no-reference method of objective X-ray image quality prediction was designed, using a generalized linear model to combine various pixel-level sample statistics such as the SNR mean, SNR standard deviation, contrast energy, estimated noise mean and so on [18]. Another no-reference image quality evaluation method was developed that uses the weighted entropy of the grayscale distribution of a region of interest (ROI) to predict objective X-ray image quality [19]. In a related application, five factors affecting human detection performance in X-ray airport security screening systems were analyzed including: fictional threat image (FTI) view difficulty, superposition, clutter, opacity and bag size [20].

Natural Scene Statistics (NSS) models describe the statistical consistencies inherent to images taken of the world, be they of naturalistic or man-made objects or environments, i.e., the image generation process is natural as opposed to computer-generated. NSS have been well studied for various natural imaging modalities including visible light (VL), long-wavelength infrared (LWIR) and X-ray images. The NSS of photographic VL images and videos has been intensively studied and applied to the development of successful perceptual quality models [21]–[23]. A number of these models, such as VIF [24] and NIQE [25] are used in commercial streaming video systems.

Bandpass NSS models of LWIR (thermal) images are also robust and are quite useful in a variety of visual applications [26]. A high performance image classification engine which distinguishes between VL and LWIR images was designed using NSS models [26]. Features extracted from these models on LWIR images have also been demonstrated to be effective for quantifying thermal 'non-uniformity' distortions in LWIR images. Other applications where NSS models deliver excellent performance include the measurement of targeting task performance (TTP) and blind IQA of LWIR, fused VL + LWIR IQA [27], and to analyze TeraHertz (THz) images [28].

In our previous work, we have found that the NSS of X-ray image data is similar to, but different from that of VL images [29]. Here we extend that early work by developing univariate and bivariate X-ray NSS models in both the spatial and wavelet domains, apply them to analyze X-ray image quality and how it affects the task performance of professional bomb technicians. We deploy both traditional lab-measured IQIs and perceptually-relevant NSS models of X-ray images to create algorithms that make reliable predictions of task performance. We also develop a compact, highly efficient set of perceptual quality predictors that we collectively call Quality Inspector of X-Ray Images (QUIX), which are of low computational complexity and suitable for real-time applications.

The remainder of the paper is organized as follows. Section II outlines a task performance study of bomb technicians' ability to detect and identify objects as a function of image quality. Section III presents univariate and bivariate NSS models of X-ray images. Sections IV describes IEEE/ANSI N42.55 standard IQIs and their behavior against image degradations. Section V studies the performance of NSS-based IQA models and IQIs on predicting task performance on distorted X-ray images. Finally, Section VI summarizes this work with suggestions for possible future work.

[1]Nicholas G. Paulter, Jr. and Jack L. Glover served as the Chair and Vice-chair of the ANSI 42.55 working group at the time this standard was approved.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

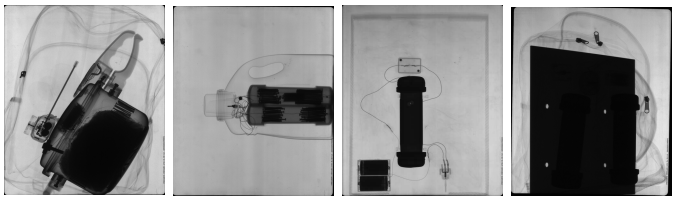


Fig. 1. A few examples of the pristine X-ray images from the NIST-LIVE X-ray image database.

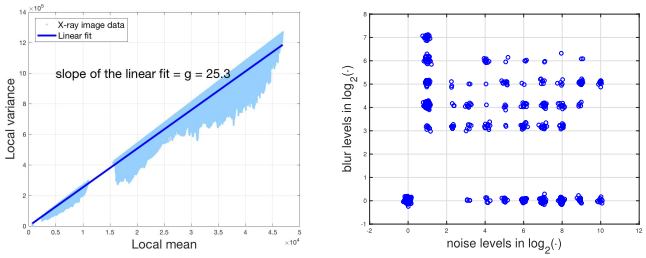


Fig. 2. (a) Scatter plot between the local means and local variances of X-ray images using only 25% of the number of pixels having the least variance-to-mean (VMR) ratio. (b) Scatter plot of different combinations of noise and blur levels used to distort the X-ray images.

II. Task Performance Study

To analyze the effects of X-ray image quality on bomb-technicians' task performance, and to evaluate the predictive performance of the objective X-ray image quality features, we require baseline (or reference) bomb-technician task-performance data. The task considered here is identifying IED components in X-ray images. Since such a database of task performance against picture quality was not available, we conducted our own X-ray task performance study. We measured the performance of bomb-technicians who were asked to detect IED components in distorted X-ray images presented in an interactive viewing environment to which they were accustomed to using. Readers may refer to [30] for a detailed report on the experimental setup and study protocol.

Our primary objective was to find relationships between measures of image quality and human task performance on degraded X-ray images. We collected X-ray images of commercially-available simulated IED threats as well as of benign everyday objects. The simulated threats include a small suitcase IED, a backpack containing pipe bombs, a backpack containing a pipe bomb hidden in a detergent bottle, a box with an IED and an anti-probe IED device, among others. We used clutter objects, including a laptop, cell phones, and a metal sheet to act as shielding. We then captured 35 pristine X-ray images of various combinations of threat objects, clutter, shielding, X-ray source and other imaging parameters. These 16-bit high-resolution grayscale images are shown in Fig. 1.

We next describe the generative model of noise that we used to degrade the X-ray images. This noise is photon limited and spatially -correlated and arises from the absorption of a random number of independent photons, $N(x, y)$, that are incident on a photosensor during the formation of the image. This spatially-correlated noise (SCN) follows a Poisson distribution. Since image noise can be directly related to the number of absorbed photons, it is possible to vary the image noise level by effectively modifying the number of absorbed photons. For the purpose of generating our SCN model, it is sufficient to approximate the X-ray beam as being monochromatic with an effective X-ray energy E_{eff}. The relation between $N(x, y)$ and the image grayscale units, $I(x, y)$, follows a simple linear relation for imaging plate detectors at the dose levels considered in this work [31]:

$$I(x, y) = g \cdot N(x, y), \tag{1}$$

where g is the gain of the imaging device. An important outcome of Eq. (1) is the constant variance to mean ratio (VMR) of the pixel intensities [32]:

$$\frac{\text{Var}[I(x, y)]}{\text{E}[I(x, y)]} = \frac{g^2 \cdot \text{Var}[N(x, y)]}{g \cdot \text{E}[N(x, y)]} = g. \tag{2}$$

Figure 2(a) illustrates the linear relationship observed between the pixel variances and pixel intensities. The gain g, which is the slope of the linear fit to the curve, is computed using that 25% of the pixels that produce the least VMR and which correspond to nominally constant-valued image regions whose variance contribution is largely due to the image noise and not due to the textured regions of the object being imaged. To compute the number of detected photons from each location in the scene, the assumed linear relationship (1) between the photon count and the pixel intensity is used. Once the number of detected photons is known, then to vary the degree of image noise, a multiplicative factor k is used to simulate the effect of a reduced number of detected photons, $N_{eff}(x, y)$, as

$$N_{eff}(x, y) = N(x, y)/k. \tag{3}$$

The noisy photon count field, $N_{noise}(x, y)$, is then calculated using

$$N_{noise}(x, y) = SCN(x, y)\sqrt{N_{eff}(x, y) + \eta}, \tag{4}$$

where $\eta = 1$ is the variance contribution from other sensor-related sources of noise, and $SCN(x, y)$ has power spectra $SCN(f) \propto f^{-2}$ normalized to have unit variance. The square root term is the expected standard deviation for $N_{eff}(x, y)$ using Poisson statistics in the limit of a high number of counts. Finally, a noisy grayscale image is obtained from the noisy photon count field N_{noise} at each pixel location using Eq. (1). Hereafter, we use the multiplicative factor k to denote the image noise level, where higher values of k imply more severe levels of noise degradation. To form the first set of noise-only degraded images, we added eight different levels of SCN, corresponding to $k = \{8, 16, 32, 64, 128, 256, 512, 1024\}$ to each of the 35 pristine X-ray images. The highest levels of degradation caused objects like pipe bombs and batteries to no longer be easily detectable.

The effective spatial resolution can be limited in X-ray images because of many factors, such as detector type and pixel size, source size and geometry, X-ray scattering, and more. These factors may be realized in different ways, from reduced image size to broadening of the point spread function,

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

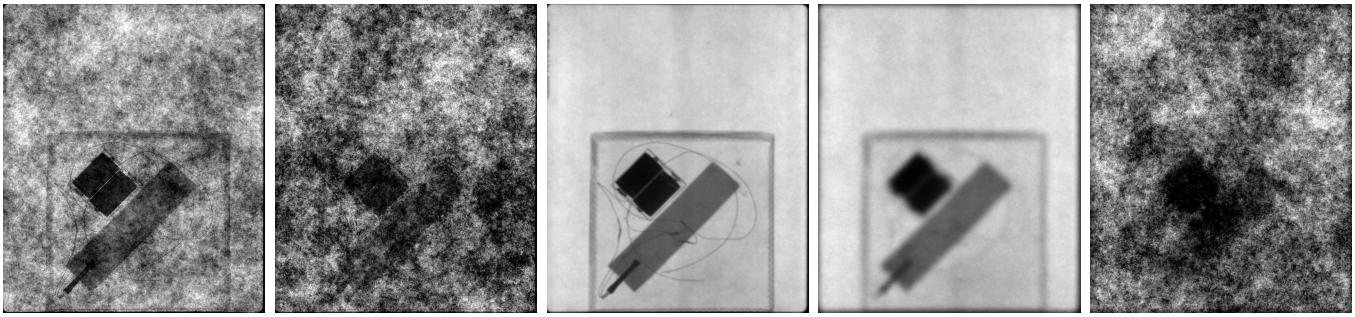


Fig. 3. Example X-ray images distorted by varying degrees of spatially correlated noise (SCN) and blur distortions. The value reported for SCN corresponds to the factor k used in Eq. (3), and the value reported for Blur corresponds to the scale of the Gaussian blur kernel (in pixels). (a) Noise: 128, Blur: 0. (b) Noise: 512, Blur: 0. (c) Noise: 1, Blur: 16. (d) Noise: 1, Blur: 64. (e) Noise: 512, Blur: 64.

all of which limit the high-frequency information available to detect and identify particular objects. We approximate these effects by convolving the images with a Gaussian spatial filter, and use the term *blur* to refer to degradation of the spatial resolution. Different blur levels were implemented by varying the Gaussian distribution to five different widths, given by $\sigma_b = \{8, 16, 32, 64, 128\}$ pixels, followed by additive SCN with $k = 1$. This value of k corresponds to the original image's NEQ factor and simulates photon-limited noise that occurs during the image formation process. The application of low level SCN after Gaussian blur prevents an unrealistic smoothness caused by the blur.

We also created a third group of images by degrading the pristine images with blur followed by higher levels of additive SCN. This third, smaller group was randomly selected from the original degradation levels, since using all combinations of blur and noise would have produced too many images for the human subjects (expert bomb technicians) to view based on the limited time for the test protocol. Figure 2(b) shows a scatter plot of the combination of noise and blur levels across all images shown to the subjects. Noise and blur levels were uniformly sampled on a logarithmic scale to obtain acceptable perceptual separation between distortions. Figure 3 shows a few of the X-ray images degraded by a variety of SCN and blur degradation levels.

The quality of the X-ray images was also varied by employing different X-ray sources operating at different energies. Care was taken to ensure that images were also randomly flipped, rotated (or both) to inhibit the subjects from learning about the image content.

The judgments of 37 subjects were used in the analysis, all of which were either current or former bomb technicians. Each subject viewed an average of 20 X-ray images ranging from a minimum of 5 to a maximum of 39 images. Considering the high proficiency and expertise of the subjects, we presented each image to an average of only 2.27 subjects. The size of the database was limited by geography and the availability of this small and specialized population. Nevertheless, we show the number of collected responses were enough to draw statistically meaningful conclusions. The subjects were each presented with a set of benign and threat-containing distorted and undistorted X-ray images using the X-ray Toolkit (XTK) software [33]. They were asked to locate and identify any potential IED components and annotate the image by drawing a box around it using a mouse. Figure 4 shows examples of human subject responses on a sample of X-ray images, along with baseline annotations of all relevant IED components.

III. NSS Analysis of X-Ray Image Data

A. Univariate Statistical Modeling of X-Ray Images

The Gaussian Scale Mixture (GSM) model provides a robust description of the statistics of bandpass wavelet coefficients of natural VL images [34] and, as it turns out, of X-ray images as well. It has been successfully applied to numerous perception-driven image processing applications [21]–[24], [35]–[39]. Recently, a *generalized* Gaussian scale mixture (GGSM) model was proposed to model the bandpass statistics of distorted VL images [40], and shown to better represent the statistics of both pristine as well as distorted VL images than the GSM model. To demonstrate this, assume that an X-ray image (distorted or not) has been subjected to a bandpass process such as a wavelet filter. The GGSM model of the marginal distributions of bandpass VL (and X-ray) image coefficients are heavy tailed, reflecting the property that natural images are predominantly smooth with sparsely distributed singular structures.

A local neighborhood of adjacent bandpass space, scales and orientations, around a reference subband coefficient can be characterized by a GGSM vector:

$$\mathbf{x} \stackrel{d}{=} \sqrt{z} \cdot \mathbf{u}, \tag{5}$$

where $\stackrel{d}{=}$ denotes equality in probability distribution, z is a scalar random variance field (called a mixing multiplier), and $\mathbf{u}$ is a zero-mean Multivariate Generalized Gaussian (MVGG) random vector with covariance matrix $\Sigma_{\mathbf{u}}$ and shape parameter s. The GGSM vector $\mathbf{x}$ represents a family of infinite Gaussian mixtures with probability density:

$$p(\mathbf{x}) = \int p(\mathbf{x}|z)p(z)dz = \int \frac{\Gamma\left(\frac{N_d}{2}\right) \cdot s \cdot z^{-N_d/2} \cdot \exp\left\{\frac{z^{-s}}{2}(\mathbf{x}^T \Sigma_{\mathbf{u}}^{-1}\mathbf{x})^s\right\}}{\pi^{N_d/2}\Gamma\left(\frac{N_d}{2s}\right) 2^{N_d/2s} |\Sigma_{\mathbf{u}}|^{1/2}} p(z)dz,$$

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

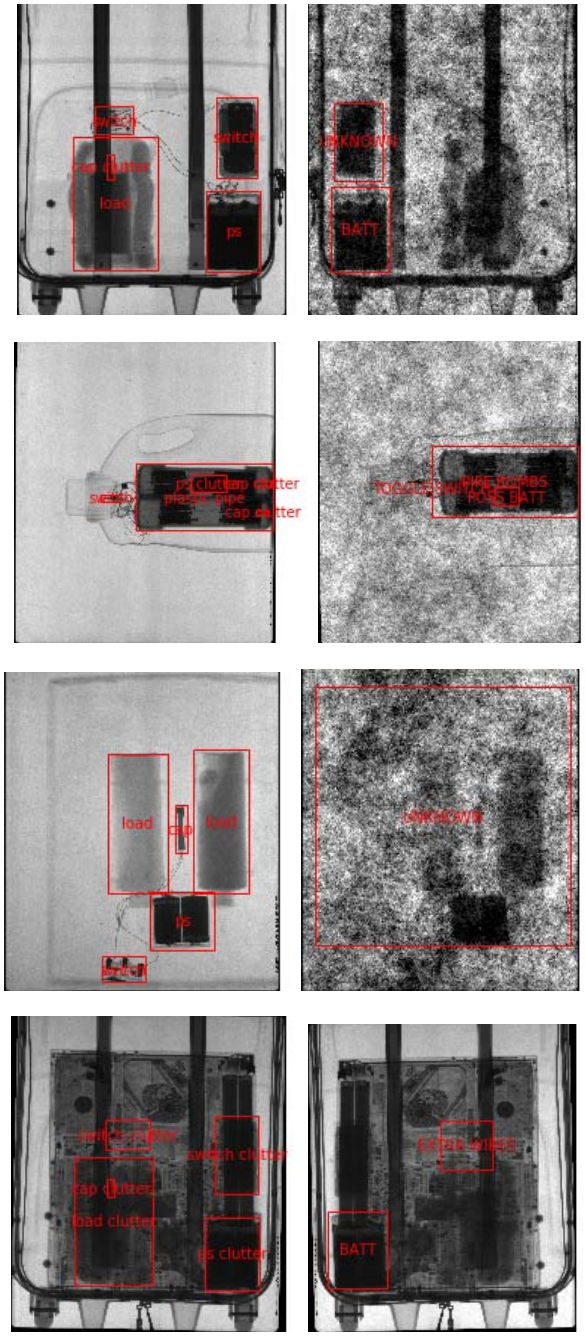


Fig. 4. Left column: Baseline annotations on all relevant IED components on original images. Right column: Human subject responses on distorted images.

where N_d is the dimensionality of $\mathbf{x}$ and $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0.$$

The GGSM model becomes a GSM when $s = 1$ and the MVGG distribution reduces to multivariate Gaussian.

One type of bandpass X-ray image decomposition we use is a steerable pyramid along 2 scales and 2 orientations (vertical and horizontal). To characterize wavelet coefficients as a GGSM vector, we utilize the neighborhood structure in [40] of 27 coefficients: 25 from the same subband (the nearest 5×5 neighbors), 1 from the parent band, and 1 spatially co-located adjacent orientation subband at the same scale.

Like the GSM model, the GGSM statistically motivates the use of a perceptually relevant divisive normalization of the non-linear responses of cortical neurons [41]. Divisive normalization is used in a number of no-reference (NR) [21], [22], reduced reference (RR) [39] and full-reference (FR) [24], [38] IQA algorithms. The density function of a GGSM vector $\mathbf{x}$ becomes generalized Gaussian when conditioned on z. Modeling the conditioning process requires estimation of the variance field z. The maximum likelihood (ML) estimate of z is given by [40]:

$$\hat{z} = \left(\frac{s}{N_d}\right)^{1/s} (\mathbf{x}^T \Sigma_{\mathbf{u}}^{-1} \mathbf{x}). \tag{6}$$

After computing the normalization coefficient $\hat{z}$, the normalized subband coefficients $\mathbf{x}/\sqrt{\hat{z}}$ are computed from each subband. Figure 5 plots the histograms of coefficients afflicted by different degrees of Gaussian blur and Gaussian SCN. Clearly, such structural degradations affect the histograms in characteristic ways. Blur increases correlation, while reducing the coefficient variances, and noise increases the coefficient uncertainty, causing wider histograms.

The second type of bandpass transformation that we use is a simple center-surround, isotropic process. Given an input luminance image I, subtract the local mean field from the image, followed by a divisive normalization step that decorrelates and Gaussianizes the coefficients [42]. This property has been used in many image quality related applications [21], [23], [25]. The normalized coefficients, often referred to as Mean-Subtracted Contrast Normalized (MSCN) coefficients, are defined as

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + c},$$

where $i \in \{1, 2, \ldots M\}$, $j \in \{1, 2, \ldots N\}$ are spatial indices, M and N are image height and width, and the constant $c = 1$ prevents numerical saturation. The weighted sample estimates of the local mean μ and standard deviation σ are:

$$\mu(i,j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} I_{k,l}(i,j)$$

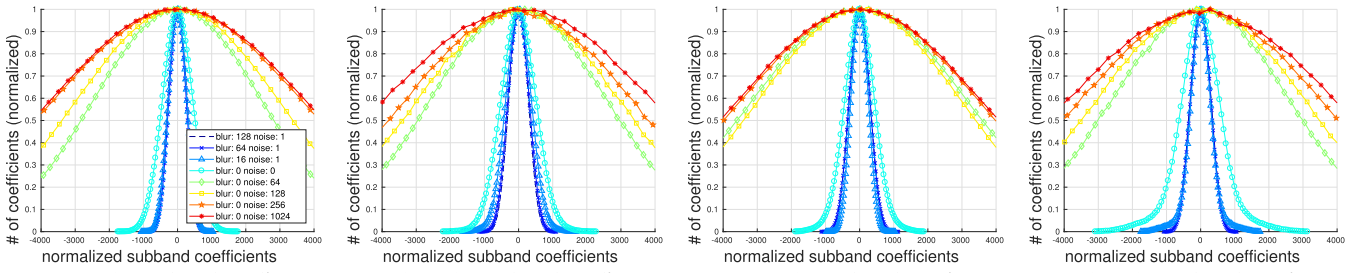
and

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} (I_{k,l}(i,j) - \mu(i,j))^2},$$

where $w = \{w_{k,l} \mid k = -K, \ldots, K \text{ and } l = -L, \ldots, L\}$ is a 2D circularly-symmetric Gaussian weighting function rescaled to unit volume and $K = L = 15$.

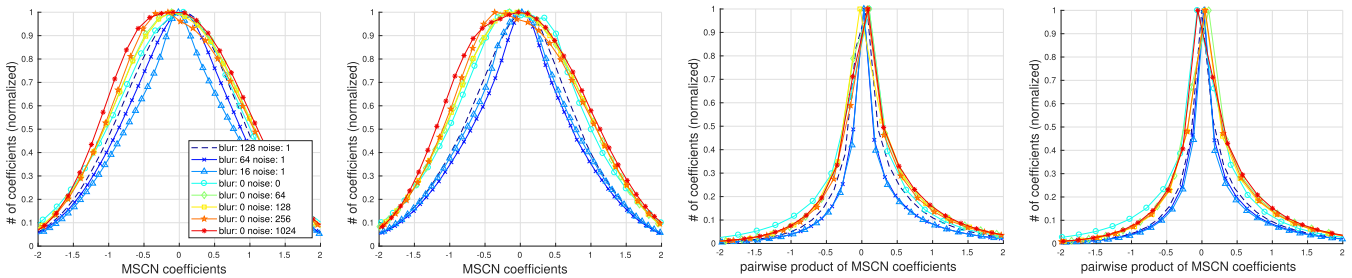
As depicted in Fig. 6, histograms of the MSCN coefficients of X-ray images exhibit Gaussian-like behavior. The effect of degradations as illustrated in Fig. 6(a) and Fig. 6(b), follow a similar trend as the normalized subband coefficients in Fig. 5. Blur results in narrower histograms, while noise produces wider histograms.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.



(a) Scale 1, orientation 45° (b) Scale 2, orientation 45° (c) Scale 1, orientation 90° (d) Scale 2, orientation 90°

Fig. 5. Histograms of GGSM-based normalized bandpass X-ray image coefficients at different scales and orientations. The effect of noise and blur on the normalized subband coefficients is quite consistent across all scales and orientations. (a) Scale 1, orientation 45°. (b) Scale 2, orientation 45°. (c) Scale 1, orientation 90°. (d) Scale 2, orientation 90°.



(a) Scale 1 (b) Scale 2 (c) Scale 1, V pairwise products (d) Scale 1, H pairwise products

Fig. 6. Histograms of the MSCN coefficients and products of spatially adjacent MSCN coefficients along the horizontal and vertical directions. Notice that the effect of noise and blur on MSCN coefficients is consistent across both scales. (a) Scale 1. (b) Scale 2. (c) Scale 1, V pairwise products. (d) Scale 1, H pairwise products.

We also model the products, or simple empirical correlations, of adjacent bandpass X-ray image samples along the horizontal and vertical orientations:

$$H(i,j) = \hat{I}(i,j)\hat{I}(i,j+1)$$
$$V(i,j) = \hat{I}(i,j)\hat{I}(i+1,j)$$

where $i \in \{1,2,3,..M-1\}$ and $j \in \{1,2,3,..N-1\}$ are spatial indices. These also exhibit statistically consistent behavior in the absence of distortions, which is perturbed when distortions are introduced, as shown in Fig. 6(c) and 6(d).

B. Extracting Univariate Statistics of Bandpass X-Ray Images

The GGSM model of bandpass image coefficients prior to normalization [40] implies that the normalized bandpass coefficients should be modeled as following a zero-mean Generalized Gaussian Distribution (GGD). The GGD density function is:

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

where

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

α is the shape parameter, σ^2 is the variance of the distribution. The parameters of the GGD are efficiently estimated using the moment-matching approach [43]. From each image, estimate two features (α,σ^2) from the best GGD fit to the MSCN coefficients (denoted as type f), and four features (of type sp) obtained from the GGD fit to normalized subband coefficients at two orientations.

A zero-mean asymmetric generalized Gaussian distribution (AGGD) models the adjacent products of bandpass MSCN coefficients, defined as:

$$f(x;\gamma,\beta_l,\beta_r) = \begin{cases} \dfrac{\gamma}{(\beta_l+\beta_r)\Gamma(\frac{1}{\gamma})}\exp\left(-\left(\dfrac{-x}{\beta_l}\right)^{\gamma}\right); & x<0 \\ \dfrac{\gamma}{(\beta_l+\beta_r)\Gamma(\frac{1}{\gamma})}\exp\left(-\left(\dfrac{x}{\beta_r}\right)^{\gamma}\right); & x\geq 0 \end{cases}$$

where

$$\beta_l = \sigma_l\sqrt{\frac{\Gamma(1/\gamma)}{\Gamma(3/\gamma)}}, \quad \beta_r = \sigma_r\sqrt{\frac{\Gamma(1/\gamma)}{\Gamma(3/\gamma)}},$$

γ is a shape parameter, and σ_l and σ_r are left and right half scale parameters. These are estimated using the moment matching scheme in [44]. The mean η is also computed:

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

Thus, three features (γ, β_r, η) (paired-product or pp features) are extracted along each orientation, yielding 6 pp features across two orientations. All of these features, which capture both structural degradations and statistical perturbations of an image, are extracted at two scales, hence a total of 24 quality-relevant univariate NSS features are obtained.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

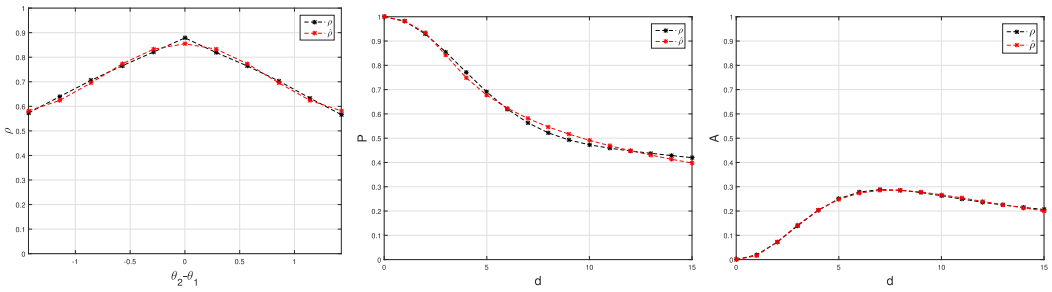


Fig. 7. Sample plots of the bivariate NSS statistics and their fits; the model correlation function ρ, the peak P and the amplitude A. a) $\rho(d=1,\phi=2,\theta_2=0)$, b) $P(d,\phi=2,\theta_2=0)$, and c) $A(d,\phi=2,\theta_2=0)$. Notice the good functional fit between the empirical data and the estimated functional fits.

C. Bivariate Statistical Modeling of X-Ray Images

The power spectra of natural photographic VL images tend to follow a reciprocal power law [45]: $S(f) \propto k/f^{\alpha}$, where $\alpha > 0$ determines the spectral fall-off. Keshner [46] derived models of the stationary autocorrelation functions of one-dimensional $1/f$ processes, arriving at a power law of reciprocal separation. In [47], the peak correlation between bandpass samples of VL images was shown to follow a stabilized reciprocal power law, neatly modeled in a closed form. The parameters of the model provide powerful NSS features useful for distortion classification [47] and NR quality prediction [48],Sinno18SSIAI. We use bivariate features extracted under the closed form NSS model, which requires several processing stages.

1) Steerable Filtering: A steerable filter with frequency tuning orientation θ_1 is defined by:

$$F_{\theta_1}(\boldsymbol{x}) = cos(\theta_1)F_x(\boldsymbol{x}) + sin(\theta_1)F_y(\boldsymbol{x}), \tag{7}$$

where $\boldsymbol{x} = (x, y)$, and F_x and F_y are the gradient components of a 2D unit-energy bivariate isotropic Gaussian function:

$$G(\boldsymbol{x}) = \frac{1}{2\pi\phi^2} e^{\frac{-(x^2+y^2)}{2\phi^2}}, \tag{8}$$

having scale parameter ϕ. We varied the scale parameters ϕ of the bivariate gaussian derivative functions (F_x and F_y) between 1 and 3 to construct a multi-scale bandpass image decomposition broadly resembling the responses of populations of cortical simple cells. We used 13 frequency tuning orientations $\theta_1 \in [0, \pi/13, 2\pi/13, \ldots, 12\pi/13]$, yielding 39 bandpass responses per image.

2) Divisive Normalization: Next, we divisively normalize each subband response [41] as:

$$u_j(\boldsymbol{x}) = \frac{w_j(\boldsymbol{x})}{\sqrt{t + \sum_{\boldsymbol{y}} g(j(\boldsymbol{y}), w_j(\boldsymbol{y}))^2}}, \tag{9}$$

where w_j are responses of filters indexed by j, and $t = 10^{-4}$ is a stabilizing saturation constant. The weighted sum in the denominator is computed over a spatial neighborhood of pixels from the same sub-band, where $g(x_i, y_i)$ is a circularly symmetric, unit volume Gaussian function.

3) Correlation Model: Given a steerable filtered and divisive normalized image, define a window at a fixed position (Window 1) and another sliding window of the same dimensions (Window 2). Denote the Euclidean distance between the center of the two windows by d. Let the windows be separated by horizontal and vertical separations δ_x and δ_y varied over the integer range [1,15], yielding 15 distances $d = \sqrt{\delta_x^2 + \delta_y^2}$. Denote the spatial orientation between the windows by $\theta_2 = \arctan(\frac{\delta_y}{\delta_x})$ (relative to the horizontal). Limit $\theta_2 \in [0, \pi)$ since the quantities being measured are symmetric and π-periodic. Then the relative angle $\theta_2 - \theta_1$ takes 13 values for each fixed spatial angle, θ_2. The correlation function model expresses a periodic behavior in the relative angle $\theta_2 - \theta_1$, as depicted in Fig. 7(a):

$$\rho(d, \phi, \theta_2) = A(d, \phi, \theta_2)cos(2(\theta_2 - \theta_1)) + c(d, \phi, \theta_2) \tag{10}$$

where $A(d, \phi, \theta_2)$ is amplitude, $c(d, \phi, \theta_2)$ is an offset, d is the spatial separation, and ϕ is the scale parameter. As in the case of VL images [47], the shapes of ρ, A, and c vary in a consistent way with d, ϕ and θ_2 on X-ray images. Next, define the peak correlation function:

$$P = \max(\rho) = A + c. \tag{11}$$

wherein we may rewrite (10) as:

$$\rho(d, \phi, \theta_2) = A(d, \phi, \theta_2)cos(2(\theta_2 - \theta_1)) + [P(d, \phi, \theta_2) - A(d, \phi, \theta_2)], \tag{12}$$

where A, P are obtained by fitting (12) to the data.

4) Modeling the Amplitude and the Peak Functions: We also found closed form expressions for A and P [47]. Lee *et al.* [50] observed that the covariances of bandpass pixels follow an approximate reciprocal power law $\frac{1}{|d|^b}$. We modify the model by expressing the peak correlation function in the more stable form $\frac{1}{|d|^{\beta}+1}$, as depicted in Fig. 7(b). Thus, given a fixed spatial orientation θ_2 and a scale ϕ, define

$$\hat{P}(d, \phi, \theta_2) = \frac{1}{(\frac{d}{\alpha_0(\theta_2)*\phi})^{\beta_0} + 1} \tag{13}$$

where $\{\alpha_0, \beta_0\}$ are parameters that control the shape and fall-off of the peak correlation function, and which depend on the spatial orientation θ_2. Given the similarity of the graph of A to the difference of two functions of the *same form* as (13) but of different scales (Fig. 7 (c)), model A as

$$\hat{A}(d, \phi, \theta_2) = \frac{1}{(\frac{d}{\alpha_1(\theta_2)*\phi})^{\beta_1(\theta_2)} + 1} - \frac{1}{(\frac{d}{\alpha_2(\theta_2)*\phi})^{\beta_2(\theta_2)} + 1} \tag{14}$$

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

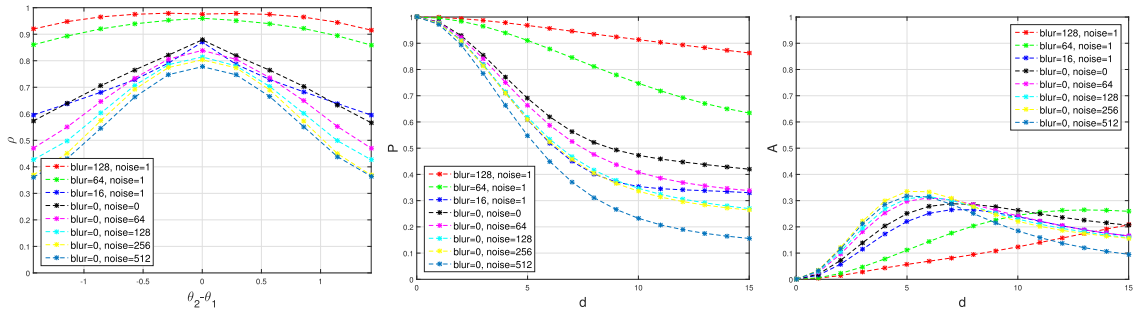


Fig. 8. The impact of distortions on the bivariate NSS parameters. a) $\rho(d=3,\phi=3,\theta_2=0)$, b) $P(d,\phi=3,\theta_2=0)$, and c) $A(d,\phi=3,\theta_2=0)$. Notice how distortions lead to consistent and systematic changes in the shapes of ρ, A and P.

where $\{\alpha_1, \beta_1,\alpha_2, \beta_2\}$ are parameters that are functions of θ_2 that control the shape of A. As may be observed in Fig. 8, distortions lead to changes in the shape of ρ, A and P. Introducing blur leads to an increase in the correlation, as expected. Noise, on the other hand, results in less similarity between neighboring pixels and consequent reduced correlation.

The shape parameters $\{\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha_2, \beta_2\}$ respond in unique ways to each distortion type. To find the values of the parameters $\{\alpha_0, \beta_0\}$ that produce the best fit to (13), and the parameters $\{\alpha_1, \beta_1,\alpha_2, \beta_2\}$, that yield the best fit to (14), we formed least-squares optimization systems for P and A. We applied unconstrained nonlinear regression using the quasi newton method [51]. The four functions $P(d,\sigma,\theta_2)$, $A(d,\sigma,\theta_2)$, $\hat{P}(d,\sigma,\theta_2)$, and $\hat{A}(d,\sigma,\theta_2)$ form vectors of size $m\times 1$, where m is the number of occurrences of θ_2. Denote by D the set of distances for a given spatial orientation θ_2. For the case $\theta_2=0$ or $\pi/2$, $D=\{0,1,2,3,\cdots,15\}$. For the case $\theta_2=\pi/4$ or $3\pi/4$, $D=\{0,\sqrt{2},\sqrt{8},\sqrt{18},\ldots,\sqrt{450}\}$. Our optimization systems are then expressed as:

$$\min_{\alpha_0,b_0}\sum_{d\in D}\sum_{\phi=3}^{3}(P(d,\phi,\theta_2)-\hat{P}(d,\phi,\theta_2))^2 \tag{15}$$

and

$$\min_{\alpha_1,b_1,\alpha_2,\beta_2,b_2}\sum_{d\in D}\sum_{\phi=3}^{3}(A(d,\phi,\theta_2)-\hat{A}(d,\phi,\theta_2))^2 \tag{16}$$

We observed that the third scale, $\phi=3$, captures distortions extremely well, which is not surprising given the heightened sensitivity of the visual system to middle frequencies. While the model also applies to other scales, we found that excluding them when solving the optimization systems (15) and (16) yielded $\{\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha_2, \beta_2\}$ that are more sensitive to distortions. It also produces a smaller, more convenient and discriminative set of features. Among these features, we have found that α_0 and β_0 (denoted as type *bi*), corresponding to $\theta_2=\frac{\pi}{4}$, are the most sensitive to distortions, and lead to the best prediction performance. Hence these two features are included in our model. Overall, a total of 26 features are extracted from the univariate and bivariate NSS models of X-ray images, as summarized in Table I.

TABLE I

FEATURE SUMMARY FOR MSCN(f), PAIRWISE PRODUCTS(pp), STEERABLE PYRAMID(sp) AND BIVARIATE FEATURES. UNIVARIATE FEATURES (f, pp, sp) ARE EXTRACTED AT TWO SCALES

Feature ID	Feature Description	Computation Procedure
f_1 - f_2	shape and variance	GGD fit to MSCN coefficients
pp_1 - pp_3	shape, mean and right variance	AGGD fit to H pairwise coefficients
pp_4 - pp_6	shape, mean and right variance	AGGD fit to V pairwise coefficients
sp_1 - sp_2	shape and variance	GGD fit to normalized subband coefficients (orientation = 0°)
sp_3 - sp_4	shape and variance	GGD fit to normalized subband coefficients (orientation = 90°)
α_0 - β_0	shape parameters of the peak function	MSE fit to the peak of the bivariate correlation when $\phi=3$ and $\theta_2=\frac{\pi}{4}$.

IV. X-RAY IMAGE QUALITY INDICATORS FROM THE IEEE/ANSI N42.55 STANDARD

The IEEE/ANSI N42.55 standard is the international standard governing the use of portable X-ray imaging systems used in IED and hazardous device identification [5]. The standard underwent a major revision in 2013 to incorporate a set of objective image quality indicators (referred to here as IQIs). Previous versions of the standard relied on subjective human judgments to score tests. The standard also defines baseline performance requirements on portable X-ray imaging systems. These IQIs must be evaluated on images of a standard test object, captured under controlled test conditions. Here, we review the IQIs defined in IEEE/ANSI N42.55, and analyze the effects of distortions on them.

- **Steel Penetration** characterizes the ability of a portable X-ray system to produce usable images of objects hidden behind shielding. Larger values mean that images with greater contrast can be obtained of objects hidden behind shielding. However, large values require high energy X-rays, which may limit imaging of thinner materials or organics. Noise increases random fluctuations, further reducing the visibility of objects behind steel. The decreasing trend of steel penetration against noise amplitude is depicted in Fig. 9(a).
- **Spatial resolution** describes the ability of an imaging system to resolve the fine details. It is the spatial frequency at which the modulation transfer function falls

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

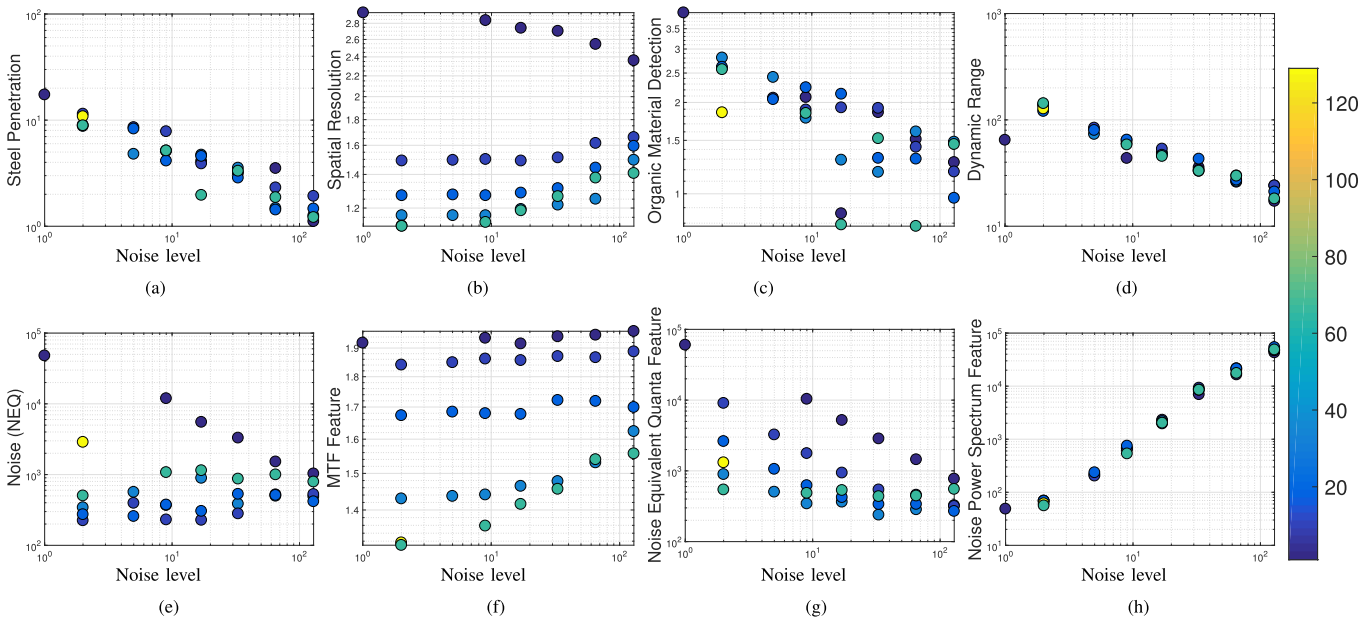


Fig. 9. Log-Log scatter plots of IQI features. Plotted are the values of the IQI feature vs the noise level (k, used in Eq. (3)). Notice that each varies in a characteristic way with increasing severity of noise degradation. The color of each point represents the amount of blur degradation.(a) Steel penetration. (b) Spatial resolution. (c) Organic material detection. (d) Dynamic range. (e) Noise (NEQ). (f) MTF feature. (g) NEQ feature. (h) NPS feature.

to 20% of peak value. While noise does not affect spatial resolution much, it decreases sharply with increasing blur (Fig. 10(b)).

- **Organic material detection** describes the ability of an X-ray imaging system to image thin pieces of low atomic number materials, such as organic compounds that comprise most explosive and energetic materials. A large value typically requires low energy X-rays, which may not permit imaging through thick metals. The organic material detection indicator decreases with increasing noise levels (Fig. 9(c)).
- **Dynamic range** is the useful range of pixel values that an imaging system can capture. It is the ratio of the largest image value in an image to the smallest usable value, which is typically the minimum noise value. Thus, it is inversely proportional to the noise magnitude as shown in the scatter plot in Fig. 9(d).
- **Noise** is measured in terms of noise equivalent quanta (NEQ), which is a form of signal-to-noise ratio (SNR). The noise IQI is defined as the value of NEQ at 1 cy/mm. A robust estimate of the NEQ feature should strongly depend on both noise and blur. The inverse relationship of NEQ against the magnitude of noise is evident from Fig. 9(e).

In addition to the objective IQIs defined by IEEE/ANSI N42.55, other descriptive features were also extracted: one from the spectral distribution of the MTF, another from the noise power spectra (NPS), and another from the NEQ. The 'MTF feature' was derived by integrating the MTF over 0 to 0.25 cy/mm. The NEQ and NPS features were extracted by integrating the NEQ and NPS spectral distributions over the range 0 to 1 cy/mm. The effect of degradations on these features is depicted in Figs. 9(f)-(h) and 10(f)-(h). Overall, a total of 8 IQI features are computed on each image.

V. Predicting Task Performance

A. QUIX: Quality Inspector of X-Ray Images

Our key goal is to identify a compact set of efficient and meaningful image task vs. quality predictors that can be readily adapted to assess X-ray imaging systems. We selected a combination of effective f and pp NSS features (16 features; 8 features per scale) based on their simplicity, minimal computation, and relatively easy interpretability. We refer to a collective set of task-prediction algorithms that use these NSS features as QUIX (QUality Inspectors of X-ray images).

We evaluated the task prediction performance of QUIX models on images of N42.55 test objects. This allows for a fair comparison of QUIX models against N42.55 IQIs, while also providing a way to gauge the intrinsic image-formation quality of an X-ray imaging system. Hereafter we refer to QUIX features computed on images of N42.55 test objects as $\text{QUIX}_{\text{N42.55}}$ features, to distinguish them from QUIX features computed on X-ray images of other, non-N42.55 test objects. We next study and compare the performance of QUIX and $\text{QUIX}_{\text{N42.55}}$ against other quality feature groups, and in combination with N42.55 IQIs.

B. Component-Wise Prediction Performance

We hypothesized that degrading the quality of X-ray images would hinder the abilities of experts to successfully detect and identify objects in them, and that properly designed, distortion-sensitive objective quality prediction engines would correlate well with expert performance on these tasks. Based on this premise, we designed a binary classification framework, whereby a classifier was trained to learn a mapping from a set of quality features to a binary target variable indicating successful identification of an object by an expert.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

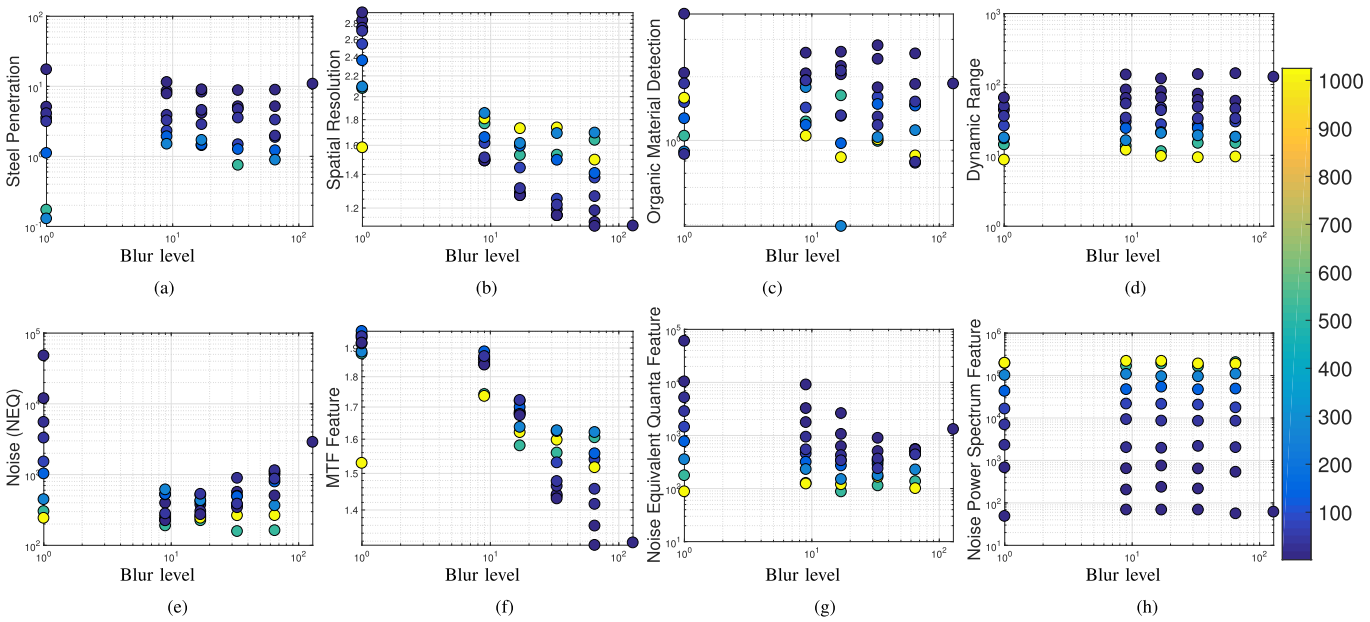


Fig. 10. Log-Log scatter plots of IQI features. Plotted are the values of the IQI feature vs the blur level. Notice that each is characteristically affected (or unaffected) by increasing severity of blur distortion. The color of each point represents the amount of noise distortion.(a) Steel penetration. (b) Spatial resolution. (c) Organic material detection. (d) Dynamic range. (e) Noise (NEQ). (f) MTF feature. (g) NEQ feature. (h) NPS feature.

TABLE II

MEDIAN LOG-LOSS AND AUC SCORES OF DIFFERENT QUALITY-AWARE FEATURE GROUPS FOR EACH COMPONENT CATEGORY ACROSS 1000 TRAIN-TEST TRIALS. THE BEST SCORE IN EACH CATEGORY IS BOLDFACED. NUMBER OF IMAGES IN EACH COMPONENT CATEGORY IS ALSO REPORTED. NUMBER OF UNIQUE CONTENTS INDICATES THE NUMBER OF UNDISTORTED IMAGES IN WHICH THE GIVEN COMPONENT IS PRESENT

	# of instances, unique contents	QUIX		$QUIX_{N42.55}$		QUIX + sp + bi		IQIs		QUIX +sp+bi+ IQIs		QUIX + IQIs	
		log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC
Power source	464, 27	0.549	0.794	0.552	0.802	0.544	0.796	0.570	0.766	**0.522**	**0.829**	0.550	0.791
Detonator	290, 17	0.425	0.839	0.414	0.822	0.363	0.894	0.429	0.802	**0.344**	**0.907**	0.417	0.839
Load	195, 15	0.462	0.833	0.472	0.815	0.445	0.858	0.493	0.788	**0.442**	**0.864**	0.464	0.828
Switch	399, 26	0.404	0.807	0.390	0.824	0.400	0.819	0.386	0.829	**0.374**	**0.857**	0.385	0.839
Metal pipe	145, 9	0.431	0.857	0.433	0.853	0.401	0.847	0.435	0.825	**0.341**	**0.897**	0.410	0.864
Weighted Average		0.463	0.817	0.460	0.818	0.443	0.835	0.470	0.799	**0.420**	**0.862**	0.455	0.825

We observed that detection and identification accuracy strongly depended on the shapes and sizes of the IED components. For instance, large, high-density components, such as a metal pipe and batteries, were relatively easy to identify, as compared to a detonator or a switch, with the effect becoming more pronounced in highly cluttered or more distorted images. Thus, we studied individual component identification performance, as well as collective performance. We divided[2] the database into five component categories: power source, load, detonator, switch and metal pipe. Table II indicates the number of images present in each category. We evaluated the predictive performance of the various combinations of quality-aware features for each of these separate categories.

In all the experiments, we followed a typical machine learning approach of training a classification engine, on a set of NSS and/or IQI quality features and X-ray image object class labels, to predict the classes of the input. To evaluate the performance of these features, we randomly divided the dataset into non-overlapping sets of 80% training and 20% test samples. This procedure of training and testing on random disjoint splits was repeated 1000 times to avoid bias due to any division of data. In each split, we ensured disjoint content separation, hence images with similar content were not present in both training and test sets at the same time.

Considering the fairly high-dimensionality of the feature sets and the limited number of samples in each category, we made the conservative choice of logistic regression over more sophisticated classifiers. Because of the necessarily limited expert dataset, this reduced the likelihood of model overfitting; we observed that a more sophisticated support vector machine (SVM) could not improve performance, and deeper networks were out of the question. Conversely, logistic regression is a simple and effective probabilistic binary classification approach which outputs easily interpretable class-conditional probabilities.

The evaluation metrics we used to compare the classification performance of the various feature groups are the

[2]In all cases, multiple IED components were present in a single image. Since features were globally extracted from each entire image, components (from the same image) belonging to different categories shared the same features.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

logarithm (log) loss, and the area under the receiver operating characteristic curve (AUC). While log loss quantifies the 'distance' between the distribution of true labels and that of the predicted labels by heavily penalizing confident misclassifications, the AUC is a ranking-based metric which indicates the probability that a classifier will rank positive classes higher than negative classes [52]. Log loss values range from zero to infinity, with perfect classification yielding zero log-loss, while AUC scores lie between 0 and 1, with values close to 1 indicating superior classification performance. An unintelligent classifier with no prior knowledge of class frequencies achieves a log loss of 0.693 and an AUC score of 0.5.

The median log-loss and AUC scores over 1000 train-test trials are reported in Table II for each component category for various combinations of NSS features, IQIs and combinations of NSS and IQI features. The average log loss and AUC scores, weighted by the number of images in each component category, are also reported. Note that IQIs are computed on images of standard test objects, hence using them as training features implies use of the same set of degradations on the test images. Consequently, images corrupted with a same set of distortion parameters share the same IQI features. By comparison, NSS-based quality features have the advantage of being independently computed on each image, and do not require laboratory imaging of any test objects.

It is clear from Table II that the combination of NSS-based QUIX features, along with other NSS-IQI feature combinations performed significantly better than the other feature groups, in terms of both log loss and AUC, across most categories. Since these descriptors capture different aspects of X-ray image quality, they supply complementary information, and task prediction using combinations of them correlates better with expert opinions. However, it is also interesting that QUIX, when used in isolation closely approached the 'combination' models on almost all component-clutter combinations. The original IQIs did as well, and the overall log loss and AUC of QUIX and IQIs are nearly identical. However, QUIX is faster (as reported in Table IV) and much easier to implement and apply. Moreover, QUIX can provide immediate image quality information to the operator, allowing real-time adjustment of equipment.

1) Statistical Significance on Each Component Category: It is important to evaluate whether the differences in Table II are statistically significant. We utilized the log-loss metric to evaluate statistical significance. We used the distribution of the 1000 log loss values to perform significance tests. We applied the one-sided t-test between the log-loss scores computed using the distributions of true and predicted labels. The null hypothesis was no difference in the mean log-loss values of the row and the column at the 95% confidence interval, against the alternative hypothesis that the mean log loss value of the row was greater than or less than the mean log-loss value of the column. Table III reports the results of the statistical significance tests for all component categories. Given that the null hypothesis compares the means of two distributions, Fig. 11 plots the mean log-loss of different feature groups for each component category, along with standard error bars.

TABLE III

STATISTICAL SIGNIFICANCE MATRIX BASED ON LOG-LOSS SCORES. EACH ENTRY IS A CODEWORD CONSISTING OF 6 SYMBOLS FOR EACH COMPONENT CATEGORY. THE POSITION OF THE SYMBOL IN THE CODEWORD REPRESENTS THE FOLLOWING CATEGORIES (FROM LEFT TO RIGHT): POWER SOURCE, DETONATOR, LOAD, SWITCH, METAL PIPE, ALL TOGETHER. '1' SIGNIFIES THAT THE ROW ALGORITHM WAS STATISTICALLY BETTER THAN THE COLUMN ALGORITHM, '0' MEANS STATISTICALLY WORSE AND '−' MEANS STATISTICALLY EQUIVALENT

	QUIX	$\text{QUIX}_{\text{N42.55}}$	QUIX + sp + bi	IQIs	QUIX+sp+bi+IQIs	QUIX + IQIs
QUIX	------	---0--	000--0	1-10-1	000000	---0-0
$\text{QUIX}_{\text{N42.55}}$	---1--	------	000100	1-1--1	000000	---000
QUIX + sp + bi	111--1	111011	------	1110-1	00-000	1110-1
IQIs	0-01-0	0-0--0	0001-0	------	000000	000--0
QUIX+sp+bi+IQIs	111111	111111	11-111	111111	------	111111
QUIX + IQIs	---1-1	---111	0001-0	111--1	000000	------

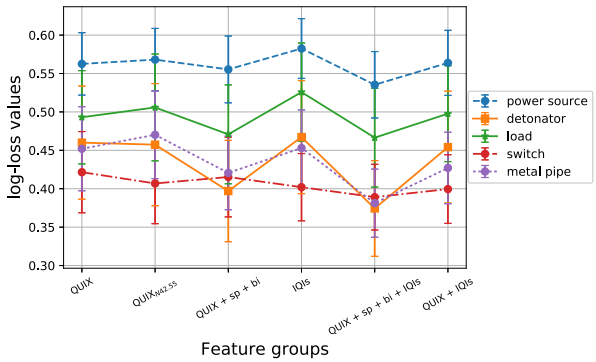


Fig. 11. Mean log-loss values with single standard deviation wide error bars of different feature groups for each component category, across 1000 train-test trials.

TABLE IV

COMPARISON OF MEDIAN TIME TAKEN PER IMAGE TO EXTRACT NSS-BASED FEATURES ON A 4 GHz QUAD-CORE PROCESSOR WITH 32 GBs OF RAM. THE MEDIAN IS COMPUTED OVER ALL IMAGES FROM THE NIST-LIVE X-RAY TASK PERFORMANCE DATABASE

Algorithm	Time (seconds)
QUIX	3.19
sp	56.08
bi (third scale)	3.76

It is obvious from Table III that the combination of QUIX, *sp*, *bi* and IQI features statistically dominates the other feature groups across most component categories.

2) Feature Importance Analysis: It is important to also understand the prediction capabilities of the individual NSS features IQIs. We implemented an iterative *forward feature selection* scheme, whereby different features are progressively incorporated into a growing set of features based on a performance criterion. Begin with an (empty) feature set $\mathbb{S}$. In the first iteration, the best performing feature is selected and included in $\mathbb{S}$. In each subsequent iteration, the feature from the remaining set that best improves model performance is incorporated into $\mathbb{S}$. The process is terminated when no

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

TABLE V

MEDIAN IDENTIFICATION ACCURACIES ACROSS 100 TRAIN-TEST TRIALS ON DIFFERENT COMPONENT-CLUTTER COMBINATIONS. THE BEST TWO FEATURE GROUPS FOR EACH COMPONENT-CLUTTER CATEGORY ARE BOLDFACED

		QUIX		$QUIX_{N42.55}$		QUIX + sp + bi		IQIs		QUIX +sp+bi+ IQIs		QUIX + IQIs	
Component	Clutter type	log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC	log-loss	AUC
Power Source	Clutter	0.489	0.831	0.499	0.848	0.503	0.834	**0.485**	0.853	0.486	**0.865**	**0.451**	**0.877**
Power Source	Shielded	0.509	0.900	0.549	0.888	**0.430**	**0.930**	0.555	0.845	**0.444**	**0.925**	0.538	0.892
Power Source	Shield with Clutter	0.205	0.833	0.226	0.833	**0.171**	0.773	0.312	0.591	**0.176**	0.750	0.205	**0.900**
Power Source	No Clutter	0.579	0.865	0.534	0.838	0.540	0.818	0.499	0.886	**0.475**	**0.887**	**0.481**	**0.893**
Detonator	Clutter	0.226	0.944	**0.182**	**0.957**	**0.191**	**0.951**	0.200	0.933	0.209	0.938	0.213	0.939
Detonator	Shield with Clutter	0.526	0.875	0.495	0.875	**0.346**	0.875	0.624	0.875	**0.348**	0.875	0.512	0.875
Detonator	No Clutter	0.395	0.944	0.430	0.889	0.403	0.889	0.455	0.944	**0.346**	0.944	**0.346**	0.944
Load	Clutter	0.137	0.944	**0.125**	0.944	**0.124**	0.944	0.163	0.944	0.127	0.944	0.134	0.944
Load	No Clutter	0.622	**0.883**	0.546	0.852	**0.531**	0.852	**0.505**	0.852	0.549	0.861	0.592	0.861
Switch	Clutter	0.275	0.938	**0.218**	0.938	0.237	0.938	0.318	0.938	**0.222**	0.938	0.249	0.938
Switch	Shielded	**0.311**	0.936	0.337	0.919	0.334	0.964	0.376	0.932	0.342	0.944	0.334	0.927
Switch	No Clutter	0.502	0.877	0.464	0.899	0.467	0.875	0.416	**0.931**	**0.390**	**0.923**	**0.411**	0.907
Metal pipe	Clutter	0.397	0.125	0.386	0.375	0.353	0.625	**0.291**	**1.000**	**0.292**	0.875	0.304	**1.000**
Metal pipe	Shielded	0.511	1.000	**0.457**	1.000	0.551	1.000	**0.394**	1.000	0.533	1.000	0.478	1.000
Metal pipe	Shield with Clutter	0.472	0.905	0.469	0.939	0.456	0.939	0.432	0.905	**0.366**	**1.000**	**0.363**	**0.952**
Weighted Average		0.414	0.875	0.396	0.880	0.377	0.878	0.405	0.879	**0.353**	**0.902**	**0.374**	**0.912**

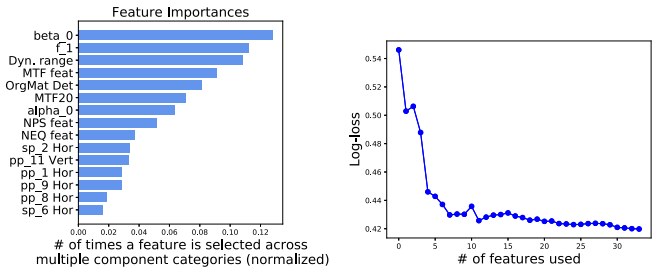


Fig. 12. (a) Horizontal bars showing the number of occurrences (normalized) of best 15 features, cumulatively selected using a forward feature selection scheme aggregated across all component categories. (b) Plot showing the log-loss performance with increasing number of features chosen in a decreasing order of importance.

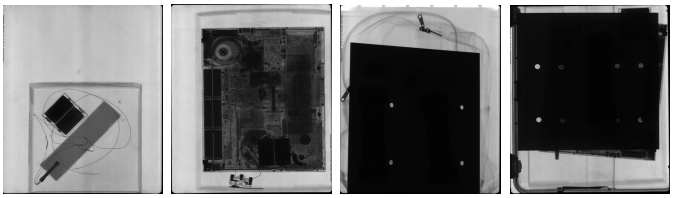


Fig. 13. Example images representative of four distinct clutter categories (left to right): no clutter, clutter (laptop), shield, clutter with shield.

additional feature improves model performance. To conduct feature-importance analysis, we randomly divided the dataset into disjoint 80%-20% train-test sets, then performed 5-fold cross-validation on the training set to obtain 'n' best features from each fold using forward feature selection. Next, we aggregated the features across folds to evaluate their performance on the test set. This process was repeated over 1000 iterations to prevent inconsistencies due to any data division bias. In our experiments, we found that selecting $n = 5$ features per fold was enough to ensure robustness, and that further increasing this number did not affect the results. The chosen features were deemed to provide better predictive power. We used the log-loss as an evaluation metric to compare feature performances. Figure 12(a) shows the relative number of times each feature was selected across all component categories.

We next evaluated the feature performance on the test-set of each train-test trial by progressively adding features in decreasing order of importance. Figure 12(b) plots the median log-loss across 1000 train-test trials, weighted by the number of images from each component category. Evidently, including more features eventually provides diminishing returns, indicating the submodular property of the feature subset.

C. Performance on Component + Clutter Combinations

We also observed that the clutter dimensions and density affected task performance. To better understand the influences of clutter types on the identification task, we further divided each component category into four sub-categories, which are clutter (laptop), shield (metallic shield), clutter with shield, and no clutter. Figure 13 shows example images containing each clutter type. Although this clutter category division reduces the number of samples in each component-clutter subcategory, it is also important to remove variations in prediction performance that occur when different clutter types are not distinguished in the analysis. Hence, we next study identification performance on component-clutter combinations.

We follow a similar binary classification framework as used earlier. Since some sub-categories contain only a few samples, we did not consider them (those with < 30 samples), given the likelihood of overfitting on few samples in a high dimensional feature space. Table V tabulates the prediction performance of the compared objective quality features, for each combination

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

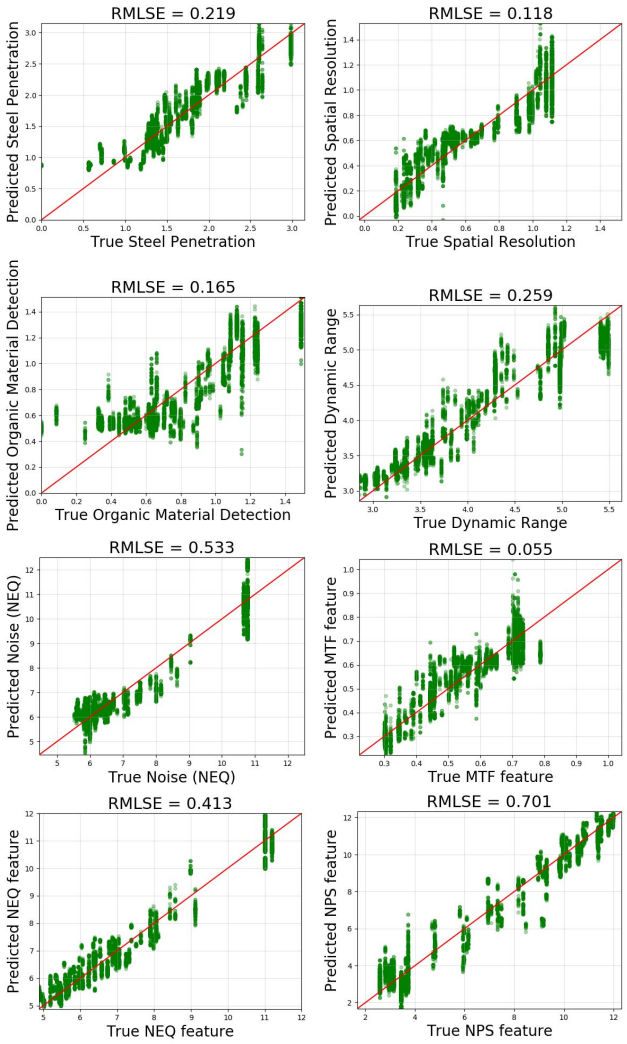


Fig. 14. Log-Log scatter plots between true IQI values and QUIX-predicted values for the given IQIs with median root-mean-log-square-error (RMLSE) across 100 train-test trials. Red line with slope = 1 corresponds to perfect prediction of IQIs.

of component and clutter. It is clear that the combination of IQIs and QUIX features performed better than the other compared feature groups.

D. Predicting IQIs Using QUIX

Given that IQIs can only be applied using a lab set-up, while a QUIX quality predictor can be readily applied to any image of interest, a natural question arises: is it possible to predict IQIs using QUIX? It is also important to understand the relationship between predictions produced by QUIX and IQIs computed on images of test objects. Thus, we designed a regression framework wherein a series of features are mapped to IQI values. Since NSS-based QUIX features are naturally invariant against changes of scenes and objects, we increased the complexity of the prediction task by mapping the QUIX features extracted on X-ray images not containing test objects to IQIs computed on images of test objects.

To evaluate performance, we trained a Support Vector Regressor (SVR) with radial basis function (RBF) kernel to learn a mapping from QUIX features to IQI values. The data was randomly divided into 100 train-test splits of content-disjoint 80% training and 20% testing sets. The target IQIs were log-transformed before training to reduce skewness in their distribution. Figure 14 illustrates an important result – although the QUIX features were extracted from distorted X-ray images not containing test objects, they were quite effective in predicting IQIs computed on a different set of (test) images. This strengthens the notion that QUIX features are invariant across scenes and objects, and can be reliably used to predict IQIs with good accuracy (prediction errors reported in Fig. 14).

VI. Conclusion and Future Work

We studied NSS-based statistical models of natural and distorted X-ray images. NSS models effectively capture statistical consistencies of X-ray images, and provide perceptually relevant tools for estimating the effects of image degradations. We demonstrated the outstanding performance of QUIX human task prediction models, especially when used in conjunction with traditional IQIs when predicting the visual task performances of trained experts. While the computation of IQIs involves significant time, cost and effort, QUIX features rely only on simple statistical feature extraction with low computational overhead. We also demonstrated the predictive abilities of QUIX features by utilizing them to estimate IQIs. NSS features are invariant across scales, scenes and objects, and can be reliably used as powerful and generic quality descriptors of X-ray images.

A potential future direction of research would be to investigate the effects of geometric image degradations on visual task performance in multi-view X-ray images [53]. Studying the statistics of other X-ray modalities, such as computed tomography (CT) and dual energy imaging systems, also offers significant possibilities.

References

[1] W. Yi *et al.*, "Portable pulsed electronic digital X-ray imager," *NDT&E Int.*, vol. 32, no. 4, pp. 215–218, 1999.
[2] S. W. Hasinoff, "Photon, Poisson noise," in *Computer Vision: A Reference Guide*. Boston, MA, USA: Springer, 2014, pp. 608–610.
[3] P. Chatzispyroglou, "Evaluating and modelling the performance of a portable X-ray imaging system," Ph.D. dissertation, Dept. Phys. Astron., The Univ. Leicester, Leicester, U.K., 2016.
[4] *American National Standard for Measuring the Imaging Performance of X-Ray and Gamma-Ray Systems for Security Screening of Humans*, Standard ANSI-N42.47-2010, Jul. 2010.
[5] *American National Standard for the Performance of Portable Transmission X-Ray Systems for Use in Improvised Explosive Device and Hazardous Device Detection*, Standard ANSI-N42.55-2013, Dec. 2013.
[6] R. S. Saunders, Jr., J. A. Baker, D. M. Delong, J. P. Johnson, and E. Samei, "Does image quality matter? Impact of resolution and noise on mammographic task performance," *Med. Phys.*, vol. 34, no. 10, pp. 3971–3981, 2007.
[7] A. Burgess, "Image quality, the ideal observer, and human performance of radiologic decision tasks," *Acad. Radiol.*, vol. 2, no. 6, pp. 522–526, 1995.
[8] L.-N. Loo, K. Doi, and C. E. Metz, "A comparison of physical image quality indices and observer performance in the radiographic detection of nylon beads," *Phys. Med., Biol.*, vol. 29, no. 7, p. 837, 1984.
[9] H. H. Barrett, "Objective assessment of image quality: Effects of quantum noise and object variability," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 7, no. 7, pp. 1266–1278, 1990.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

[10] H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 12, no. 5, pp. 834–852, 1995.

[11] H. H. Barrett, K. J. Myers, C. Hoeschen, M. A. Kupinski, and M. P. Little, "Task-based measures of image quality and their relation to radiation dose and patient risk," *Phys. Med., Biol.*, vol. 60, no. 2, p. R1, 2015.

[12] S. Park, A. Badano, B. D. Gallas, and K. J. Myers, "Incorporating human contrast sensitivity in model observers for detection tasks," *IEEE Trans. Med. Imag.*, vol. 28, no. 3, pp. 339–347, Mar. 2009.

[13] K. J. Myers and H. H. Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 12, pp. 2447–2457, 1987.

[14] X. He and S. Park, "Model observers in medical imaging research," *Theranostics*, vol. 3, no. 10, p. 774, 2013.

[15] C. K. Abbey and M. P. Eckstein, "Observer efficiency in free-localization tasks with correlated noise," *Frontiers Psychol.*, vol. 5, p. 345, May 2014.

[16] J. S. McCarley, A. F. Kramer, C. D. Wickens, E. D. Vidoni, and W. R. Boot, "Visual skills in airport-security screening," *Psychol. Sci.*, vol. 15, no. 5, pp. 302–306, 2004.

[17] A. G. Gale, M. D. Mugglestone, K. J. Purdy, and A. McClumpha, "Is airport baggage inspection just another medical image?" *Proc. SPIE*, vol. 3981, pp. 3981–3989, Apr. 2000.

[18] J. Irvine, M. Young, S. German, and R. Eaton, "Perceived X-ray image quality for baggage screening," in *Proc. Appl. Imagery Pattern Recognit. Workshop (AIPR)*, Oct. 2015, pp. 1–9.

[19] F. Chen, J. Pan, and Y. Han, "An effective image quality evaluation method of X-ray imaging system," *J. Comput. Inf. Syst.*, vol. 7, no. 4, pp. 1278–1285, 2011.

[20] A. Schwaninger, A. Bolfing, T. Halbherr, S. Helman, A. Belyavin, and L. Hay, "The impact of image based factors and training on threat detection performance in X-ray screening," in *Proc. Int. Conf. Air Trans. (ICRAT)*, 2008, pp. 317–324.

[21] 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.

[22] 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.

[23] D. Ghadiyaram and A. C. Bovik, "Feature maps driven no-reference image quality prediction of authentically distorted images," in *Proc. SPIR/IS&T Electron. Imag.*, 2015, p. 93940J.

[24] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[25] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Sign. Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

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

[27] D. E. Moreno-Villamarín, H. D. Benítez-Restrepo, and A. C. Bovik, "Predicting the quality of fused long wave infrared and visible light images," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3479–3491, Jul. 2017.

[28] M. Hu *et al.* (2017). "Terahertz security image quality assessment by no-reference model observers." [Online]. Available: https://arxiv.org/abs/1707.03574

[29] P. Gupta, J. L. Glover, N. G. Paulter, and A. C. Bovik, "Studying the statistics of natural X-ray pictures," *J. Test. Eval.*, vol. 46, no. 4, pp. 1478–1488, 2018.

[30] J. L. Glover, P. Gupta, A. C. Bovik, and N. G. Paulter, "Measuring and modeling the detectability of IED components in X-ray images as a function of image quality," to be published.

[31] Y. Amemiya, T. Matsushita, A. Nakagawa, Y. Satow, J. Miyahara, and J.-I. Chikawa, "Design and performance of an imaging plate system for X-ray diffraction study," *Nucl. Instrum. Methods Phys. Res. A, Accel. Spectrom. Detect. Assoc. Equip.*, vol. 266, nos. 1–3, pp. 645–653, 1988.

[32] L. J. van Vliet, D. Sudar, and I. T. Young, "Digital fluorescence imaging using cooled charge-coupled device array cameras," in *Cell Biology: A Laboratory Handbook*, K. Simons, Ed., 2nd ed. New York, NY, USA: Academic, 1998, pp. 109–120.

[33] *X-Ray Toolkit*. Accessed: Jun. 1, 2018. [Online]. Available: http://www.xraytoolkit.com/

[34] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 12. Cambridge, MA, USA: MIT Press, May 1999, pp. 855–861.

[35] 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.

[36] J. A. Guerrero-Colón, E. P. Simoncelli, and J. Portilla, "Image denoising using mixtures of gaussian scale mixtures," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 565–568.

[37] P. Gupta, C. G. Bampis, and A. C. Bovik, "Natural scene statistics for noise estimation," in *Proc. IEEE Southwest Symp. Image Anal. Interp. (SSIAI)*, Las Vegas, NV, USA, Apr. 2018, pp. 85–88.

[38] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

[39] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517–526, Feb. 2012.

[40] P. Gupta, A. K. Moorthy, R. Soundararajan, and A. C. Bovik, "Generalized Gaussian scale mixtures: A model for wavelet coefficients of natural images," *Signal Process., Image Commun.*, vol. 66, pp. 87–94, Aug. 2018.

[41] M. J. Wainwright, O. Schwartz, and E. P. Simoncelli, "Natural image statistics and divisive normalization: Modeling nonlinearities and adaptation in cortical neurons," in *Probabilistic Models of the Brain: Perception and Neural Function*, R. Rao, B. Olshausen, and M. Lewicki, Eds. Cambridge, MA, USA: MIT Press, 2002.

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

[43] 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.

[44] N.-E. Lasmar, Y. Stitou, and Y. Berthoumieu, "Multiscale skewed heavy tailed model for texture analysis," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 2281–2284.

[45] D. J. Tolhurst, Y. Tadmor, and T. Chao, "Amplitude spectra of natural images," *Ophthalmic Physiol. Opt.*, vol. 12, no. 2, pp. 229–232, 1992.

[46] M. S. Keshner, "1/f noise," *Proc. IEEE*, vol. 70, no. 3, pp. 212–218, Mar. 1982.

[47] Z. Sinno, C. Caramanis, and A. C. Bovik, "Towards a closed form second-order natural scene statistics model," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3194–3209, Jul. 2018.

[48] Z. Sinno, C. Caramanis, and A. C. Bovik, "Second order natural scene statistics model of blind image quality assessment," in *Proc. IEEE Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1238–1242.

[49] Z. Sinno and A. C. Bovik, "On the natural statistics of chromatic images," in *Proc. Southw. Symp. Image Anal. Interpret.*, Apr. 2018, pp. 81–84.

[50] A. Lee, D. Mumford, and J. Huang, "Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model," *Int. J. Comput. Vis.*, vol. 41, nos. 1–2, pp. 35–59, 2001.

[51] P. E. Gill and W. Murray, "Quasi-Newton methods for unconstrained optimization," *IMA J. Appl. Math.*, vol. 9, no. 1, pp. 91–108, Feb. 1972.

[52] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

[53] T. Franzel, U. Schmidt, and S. Roth, "Object detection in multi-view Xray images," in *Pattern Recognition*. Berlin, Germany: Springer, 2012, pp. 144–154.

Praful Gupta received the B.Tech. degree in electrical engineering from IIT Roorkee, Roorkee, India, in 2015 and the M.S. degree in electrical and computer engineering from The University of Texas at Austin, Austin, in 2017, where he is currently pursuing the Ph.D. degree. His research interests include image and video processing, machine learning, and computer vision.

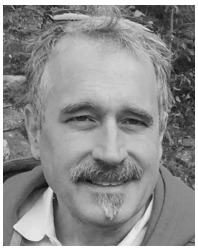
Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.

Zeina Sinno received the B.E. degree (Hons.) in electrical and computer engineering, with a minor in mathematics, from the American University of Beirut in 2013 and the M.S. degree in electrical and computer engineering from The University of Texas at Austin in 2015, where she is currently pursuing the Ph.D. degree with the Laboratory for Image and Video Engineering. Her research focuses on image and video processing and machine learning.

Jack L. Glover received the Ph.D. degree in physics from The University of Melbourne in 2014.

He is currently a Contractor with Theiss Research and is also with the Security Technologies Group, National Institute of Standards and Technology. His main areas of interest are security imaging, X-ray physics, the metrology of image quality, radiation protection, and signal detection and recognition by humans and algorithms.

Nicholas G. Paulter, Jr. was with the Los Alamos National Laboratory, Los Alamos, NM, USA, from 1980 to 1989, where he was involved in the study of fast electrical and optical phenomena. From 1990 to 2005, he was the Leader of the National Institute of Standards and Technology (NIST) High-Speed Pulse Metrology Project, Gaithersburg, MD, USA, and also with NIST-Boulder, where he developed several high-speed electrical pulse generation and sampling systems, electro-optic-based measurement systems, and short optical pulse laser systems for use in pulse metrology. From 2005 to 2013, he was a Program Manager at NIST's Office of Law Enforcement Standards, where he initiated and oversaw several research programs in security imaging, including submillimeter-wave, X-ray, and radio frequency. During his tenure as the Project Leader, the metrology services provided by his team became the best in the world. He has been the Group Leader of the Security Technologies Group, NIST, since 2013. In that capacity, he oversees metrology programs in high-strength fiber characterization, concealed-weapon detection and imaging, through-wall surveillance, characterization of materials and system used in impact mitigation, traffic control devices, and biometrics for identification. He is a Commerce Science and Technology Fellow and an IEEE Fellow. He was a recipient of the NIST Bronze Medal in 2003 for his work in developing minimum performance requirements for metal detectors and the Department of Commerce Silver Medal for his work in security X-ray imaging metrology.



Alan C. Bovik (F'95) is the Cockrell Family Regents Endowed Chair Professor with The University of Texas at Austin. He has authored or co-authored *The Essential Guides to Image and Video Processing*. His research interests are digital video, image processing, and visual perception. For his work in these areas, he was a recipient of the 2019 IEEE Fourier Award, the 2017 Edwin H. Land Medal from the Optical Society of America, the 2015 Primetime Emmy Award for Outstanding Achievement in Engineering Development from the Television Academy, and the Norbert Wiener Society Award and the Karl Friedrich Gauss Education Award from the IEEE Signal Processing Society. He received about ten best journal paper awards, including the 2016 IEEE Signal Processing Society Sustained Impact Award. He is a fellow of the IEEE. He co-founded and was the longest-serving Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING, and also created or chaired the IEEE International Conference on Image Processing which was first held in Austin, TX, USA, in 1994.

Authorized licensed use limited to: NIST Virtual Library (NVL). Downloaded on February 28,2020 at 17:22:22 UTC from IEEE Xplore. Restrictions apply.