# On the Perceptual Factors Underlying the Quality of Post-Compression Enhancement of Textures

*Yusizwan M. Yaacob, Yi Zhang, and Damon M. Chandler; Laboratory of Computational and Subjective Image Quality, Department of Electrical and Electronic Engineering, Shizuoka University; Hamamatsu, Shizuoka, Japan*

## Abstract

*The addition of white noise to an image has been shown to increase the perceived sharpness of the image's blurred regions under certain conditions. Additive white noise has also been shown to increase the visual quality a compressed image, a finding which has been attributed, in large part, to the noise's ability to simulate textures that have been lost via the compression. To explore the perceptual underpinnings of this enhancing effect, in this paper, we tested whether the noise can be tuned based on properties of the source texture to provide even greater improvements in quality as compared to white noise. We used a parametric texture-synthesis algorithm to generate statistically and spectrally shaped noise patterns, which were scaled in contrast and then added to corresponding compressed texture regions. Subjects reported both the optimal contrast scaling factors and the associated quality improvement scores relative to the distorted regions. Our results indicate that the addition of the shaped noise can provide markedly greater quality improvements compared to white noise, a finding which cannot be explained by the mere presence of high-frequency content. We discuss how the optimal contrast scalings might be predicted, and we examine the performances of existing quality assessment algorithms on our enhanced images.*

## Introduction

Nearly every digital image/video viewed by today's consumers has undergone some form of lossy compression, and thus contains some form of compression distortions. When bitrate constraints and/or uncorrectable communication errors force these distortions into the visible regime, the only hope for improvement is to attempt to restore or otherwise enhance the decompressed image. To this end, numerous artifact reduction techniques have been proposed to reduce the visibility of blocking artifacts and/or aliasing artifacts [1] and/or attempt to enhance the appearances of degraded edges (e.g., [2]). Even more general image restoration techniques have been proposed, and although these techniques are not specifically designed for compressed images, some have demonstrated the ability to partially reverse the blurring and aliasing caused by compression. Yet, despite all of this previous work, very little effort has focused on improving the visual qualities of the texture regions.

During compression, it is most often the textures that are the first to be degraded (usually via blurring), whereas edges tend to remain relatively intact, even under moderate compression. Thus, a technique focused on improving the appearances of the textures could potentially yield substantially greater quality improvements than what is currently possible. Achieving this goal requires a better understanding of the perceptual factors that underlie human judgments of texture quality in an enhancement context.

A clever but lesser-known technique of improving the quality of textures in a compressed image is to add a subtle amount of white noise to the image. Kayargadde and Martens were the first to demonstrate that adding noise to a blurred image can improve the perceived sharpness of the image [3]. In [4], Kashibuchi *et al.* demonstrated that additive white noise can improve the perceived sharpness in printed photographs. In [5], Kurihara *et al.* investigated the influence of additive noise on the perception of both sine-wave patterns and textures; Kurihara *et al.* demonstrated that white noise added only to the texture regions consistently yields an increase in the perceived sharpness, though further hypothesis regarding the underlying perceptual factors for this improvement were not explored in [5]. In a follow-up study, Wan *et al.* investigated how the addition of noise affects the perceived sharpness of various natural textures [6]. Wan *et al.* proposed that the noise increases the high-frequent content in the image, thus boosting the perceived sharpness. The prevalence of high-frequency content has also been employed by several no-reference image sharpness estimators (e.g., [7]). Wan *et al.* also demonstrated that the presence of noise facilitates the subject's ability to recall the texture from memory, thus boosting the perceived sharpness.

In the context of compression, in [8], Chandler *et al.* demonstrated that additive noise can improve the visual quality of images subjected to wavelet-based compression. Chandler *et al.* proposed that the noise has two additional effects on quality: (1) the noise helps to masks the compression artifacts; and (2) the noise helps to synthesize some of the details lost in the textures.



|  |  |
|---|---|
| Original | Distorted (Compressed) |
| Distorted + White Noise | Distorted + (Filtered) Syn. Texture |

**Figure 1.** *Demonstrative example of the potential of additive, filtered synthesized texture to improve the visual quality of a compressed image to a greater extent than additive white noise.*

Still, beyond these explanations, a deeper investigation of the relations between additive noise and its potential for visual quality enhancement remains relatively unexplored.

Assuming that the addition of noise helps to synthesize lost details, it should be possible to achieve even better quality improvements (and perhaps also higher perceived sharpness) by using "noise" that has been shaped to better synthesize the lost details (see Figure 1). To test this hypothesis, in this study, we measured quality improvement ratings for regions of compressed images enhanced via the addition of spectrally shaped synthesized textures. Specifically, we used a popular parametric texture synthesis algorithm [9] to shape the noise such that its pixel statistics and local and joint wavelet statistics matched those of the original texture; furthermore, we used only the middle and high frequency bands during the texture synthesis because these are the bands that are the most affected during compression/blurring. As a control condition, we also measured quality ratings for these same compressed images containing additive white noise. In each case, the subjects were allowed adjust the contrast of the texture or noise to maximize the perceived quality improvement. To further investigate the spectral influences, we measured quality improvement ratings for these same image regions enhanced via three more control conditions: (1) noise that was white in terms orientation, but matched to the texture in terms radial frequency information; (2) noise that was white in terms of radial frequency, but matched to the texture in terms of orientation; and (3) noise whose full 2D magnitude spectrum matched the texture, but as with (1) and (2) still had a random phase spectrum.

Based on these results, we propose a semi-automated approach for enhancing compressed images via the addition of synthesized textures. Our approach requires a texture segmentation map and sample textures from the original image (prior to compression), from which statistical parameters can be measured and transmitted as side-information. After decompression, the sampled textures can be synthesized (and spectrally shaped), scaled in contrast, and then added to the compressed image. Toward this application, we provide preliminary results of a feature-based approach for predicting the optimal contrast scaling factors, and a preliminary analysis of the performances of various quality assessment algorithms in predicting the quality improvements.

This paper is organized as follows. Details of the experimental methods are provided in Section 2. Results of the experiment are provided in Section 3. A discussion and preliminary efforts towards using the results to guide enhancement and quality assessment are provided in Section 4. Conclusions are provided in Section 5.

## Methods

Our aim was to investigate the differences in visual quality afforded by the addition of synthesized textures vs. white noise. To this end, we used as the additive test patterns: white noise, synthesized texture, and three hybrid versions which were partially white and partially synthesized. Each of these test patterns was scaled in contrast and added to individual texture regions of HEVC-compressed images. The preferred contrast scaling factors and the associated quality improvement scores were collected from the subjects.

### *Stimuli*
### *Original and Distorted Images*
Fifteen images were selected from the various categories of the McGill Image database [10] for use as original images in the study. The images were selected such that they contained at least one relatively large texture region. The images were then hand-segmented to isolate different textures and to merge perceptually similar textures (i.e., those from the same source material at roughly the same spatial scale), even if the regions were non-contiguous. Thus, each image was segmented into 1-7, possibly non-contiguous textures regions. We refer to these regions as the *source textures*. The images were resized in Matlab by using bicubic resizing to sizes of either $768 \times 576$ or $576 \times 768$ pixels.

Let $X_i$, $i \in [1, 15]$, denote the $i^{\text{th}}$ source image. Let $M_i^j$ denote a binary mask corresponding to the $j^{\text{th}}$ source texture of $X_i$, where $j \in [1, J]$, and $J \in [1, 7]$ denotes the total number of source textures for a given image.

Distorted versions of the images were created by using HEVC single-frame compression (BPG format) to obtain roughly medium-low-quality images as judged by the first author. Specifically, we used quantization scaling factors ranging from 33 to 40, resulting in bit-rate ranging from 0.03 to 0.36 bpp. In all cases, the textures were largely blurred, but still recognizable given the context. Let $\hat{X}_i$ denote the distorted (decompressed) version of $X_i$.

### *Test Patterns*
The following test patterns were used in the experiment:

1. *White noise*: White in terms of the magnitude across radial frequency and orientation; random phase spectrum.
2. *Synthesized texture*: A synthesized version of the original (uncompressed) texture, generated via a parametric texture-synthesis algorithm, but omitting the coarse subbands during the synthesis (see the Stimuli section for details).
3. *Hybrid R*: A partially synthesized and partially white texture that was matched to the synthesized texture in terms of the magnitude across radial frequency, but white in terms of the magnitude across orientation, and white (random) in terms of the phase spectrum.
4. *Hybrid O*: A partially synthesized and partially white texture that was matched to the synthesized texture in terms of the magnitude across orientation, but white in terms of the magnitude across radial frequency, and white (random) in terms of the phase spectrum.
5. *Hybrid 2D*: A partially synthesized and partially white texture that was matched to the synthesized texture in terms of the magnitude spectrum, but had a white (random) phase spectrum.

Although color images were used in the experiment, each of the test patterns was defined in terms of grayscale and was added only to the grayscale component of the images.

The synthesized textures were generated via Portilla and Simoncelli's parametric texture-synthesis algorithm [9]. Thus, each source texture was analyzed prior to compression to obtain the algorithm's required statistical parameters. For these analyses, we employed a steerable pyramid with three scales and four orientations, and we recorded the full five sets of parameters: (1) Marginal statistics of each subband; (2) auto-correlation of low-pass bands at each scale; (3) cross-correlation of magnitudes be-
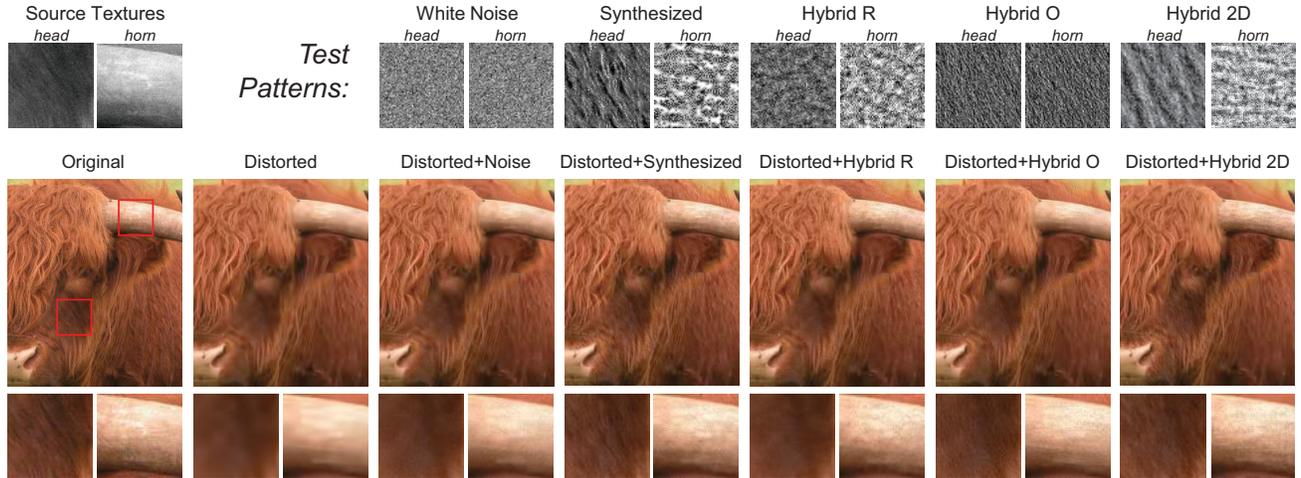
**Figure 2.** An example of the stimuli used in the experiment for one texture of one image (cropped for display purposes). First row: Source textures and test patterns. Second row: Original, distorted, and enhanced images.

tween scales; (4) cross-correlation of magnitudes between orientations; (5) cross-scale relative phase.

Synthesized versions of the source textures, equal in size to each source image, were then created. However, we did not use the low-frequency information (coarsest scale) during the synthesis. During our pilot experiments, we found that more realistic results were obtained by synthesizing only the medium- and high-frequency components. Low-frequency information made the restored textures look unnatural and in most cases resulted in blotchiness. We attribute this finding to the fact that the compression leaves intact the lower-frequency information, and thus adding a synthesized texture with more of this lower-frequency information leads to over-enhancement in this band. Thus, in this study, the synthesized textures were all zero-mean textures and devoid of the lowest-band content.

Figure 2 shows some examples of the noise- or texture- enhanced conditions. The *Hybrid R* shaped-noise patterns exhibit matched amounts of content across radial frequency, but they contain a flat distribution across orientation. The *Hybrid O* shaped-noise patterns exhibit matched amounts of content across orientation, but they contain a flat distribution across radial frequency. The *Hybrid 2D* shaped-noise patterns exhibit matched amounts of content across both radial frequency and orientation. All of the noise conditions (white noise and shaped-noise patterns) retain their random phase spectrum; thus, they lack spots, lines, and edges that manifest as a result of phase coherence.

### *Additive Enhancement Process*

Let $T_i^j$ denote the test pattern (white, synthesized, hybrid) generated for the $j$<sup>th</sup> source texture of the $i$<sup>th</sup> image. The enhancement of this $j$<sup>th</sup> texture in the distorted image $\hat{X}_i$ was performed as follows:

$$Y_i^j = \hat{X}_i + c_i^j T_i^j M_i^j \tag{1}$$

where $Y_i^j$ denotes the version of the compressed image with only the $j$<sup>th</sup> texture enhanced, and where $c_i^j$ denotes the contrast scaling factor of the synthesized texture required to achieve the highest

possible visual quality under the constraints of this enhancement process. In Experiment 1, we measured these quality-maximizing contrasts. Let $\bar{c}_i^j$ denote the quality-maximizing contrast for the $j$<sup>th</sup> source texture of the $i$<sup>th</sup> image averaged across all subjects.

Following the main experiment, we performed a follow-up quality-rating experiment in which we measured subjective quality scores for the distorted images and for the enhanced images with all source textures simultaneously enhanced. Let $Y_i$ denote the version of $\hat{X}_i$ with all source textures enhanced, obtained via

$$Y_i = \sum_{j=1}^J X_i + \bar{c}_i^j T_i^j M_i^j. \tag{2}$$

### *Procedures*

We measured two values for each source texture: (1) the contrast scaling factor required to achieve the highest possible local visual quality; and (2) the local visual quality improvement rating relative to the distorted image.

The subjects were presented with an interface in which the distorted and enhanced images were shown side-by-side. For each texture, the subjects were instructed to adjust a slider that varied the value of $c_i^j$ such that the enhanced texture appeared as high in quality as possible. Subjects then continued this process for the next texture of the same image, with all prior enhancements still intact, and this process was repeated until all source textures were enhanced. Subjects were instructed to move the slider below and above their choices (to avoid hysteresis effects), and to readjust prior slider values after all textures were enhanced (to take into account quality-level-based contextual effects). After all textures were enhanced, the subjected were asked to provide a quality score, on a scale from 0-10, for each restored texture, where 10 denoted perfect quality.

Thus, for each texture of each image, we obtained from each subject a quality-maximizing contrast ($c_i^j$) and an associated quality improvement score ($\Delta Q_i^j$) relative to the distorted image. Note that, during the experiment, the original (uncompressed) images were not shown to the subjects.

### Apparatus and Subjects

The stimuli were displayed on an I-O Data LCD-MF276XD 27-inch LCD monitor set to a resolution of $1920 \times 1080$ pixels. The screen had minimum and maximum luminances of 0.3 and 222 cd/m$^2$, respectively; and a luminance gamma of 2.0. Subjects viewed the stimuli in a dark room (following a brief period of darkness adaptation) with normal pupils. The viewing distance was 60 cm, resulting in stimuli which subtended $22.5 \times 17.0$ (or $17.0 \times 22.5$) degrees of visual angle.

Five people, including the three authors (YY, YZ, and DC), served as subjects in the experiment. YZ and DC had previous experience in viewing compressed images and in rating image quality; the other subjects were new to image-quality tests, and the non-author subjects were naive to the purpose of the experiment. Also note that subjects YZ and DC had no knowledge of the original images. Considering the highly subjective nature of this experiment, we acknowledge that this subject pool is extremely limited; at this point, any findings from this work should only be interpreted as preliminary suggestions. Still, the

## Results

Table 1 lists the average contrast factors and quality improvement scores for each texture enhanced with each of the test patterns. The per-subject quality improvement scores were first converted to standardized scores ($z$-scores), then averaged across subjects, and then rescaled to span the range $[0, 10]$. Observe that, as expected, the average contrast factors differed from texture to texture. One obvious explanation for this finding is due to the different perceived contrasts of the synthesized textures. Generally, synthesized textures with lower perceived contrasts required lesser contrast factors as compared to textures with higher perceived contrasts to achieve the maximum quality improvement. Another partial explanation for this finding is due to the different masking capabilities of the source textures. Some source textures were better than others at reducing the visibility of an additive texture, thus resulting in larger contrast factors in order to achieve the maximum quality improvement.

In regards to the quality improvement scores, as reported in previous studies, observe that the addition of white noise, in all cases, improved the quality of each respective texture to at least some extent. On average, the noise provided a quality improvement of 2.5 (on a scale from 0-10). However, in many cases, the addition of synthesized texture improved the perceived quality to a greater extent than the addition of white noise. On average the synthesized textures provided a quality improvement of 3.5. As we mentioned in Section , it has been argued that white noise improves quality in part via the addition of extra high-frequency content, thereby improving the perceived sharpness which counteracts the blurring caused by compression. However, this argument cannot fully explain our findings because a spectral analysis of the enhanced regions indicated that the synthesized textures always added lesser high-frequency content than the white noise. Clearly, and perhaps not surprisingly, the improvement in quality is due to influences beyond just the amount of high-frequency content.

The highest per-texture quality improvements relative to the distorted textures were observed for very stochastic textures (e.g., rocks and carpet/synthetic fur). These textures also received the highest mean-opinion contrast scaling factors ($c_i^j$). We suspect

that this finding can be partially attributed to the subjects unfamiliarity with the mask (entropy masking [11]) and partially attributed to the subjects' tolerances of perturbations of the mid-to-high-frequency bands.

The lowest per-texture quality improvements relative to the distorted textures were observed for the highly regular textures, and more generally for textures which the synthesis algorithm could not yield what would generally be considered a visually faithful representation of the source texture. Some textures required strict alignment of edges in order to be enhanced; these textures thus resulted in very low contrast scaling factors in order to avoid an artificial appearance. Nonetheless, the contrast scaling factors were always greater than zero, and always resulted in a visually detectable positive change in quality, suggesting that even minor improvements are possible given the limitations of the current enhancement technique.

The results of the hybrid test patterns indicate that the spectra of the synthesized textures play large roles in the quality improvements, but the relative importances of radial frequency information vs. orientation information is indeed texture-dependent.

**Table 1: Average contrast factors and quality improvements**

| Image | Segment | Contrast Factor $\bar{c}_i^j$ | | | | | Quality Improvement $\Delta \bar{Q}_i^j$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Noise | Synth | Hyb R | Hyb O | Hyb 2D | Noise | Synth | Hyb R | Hyb O | Hyb 2D |
| **Apples** | Background | 1.7 | 1.9 | 1.8 | 2.0 | 1.8 | 2.2 | 3.3 | 3.1 | 2.5 | 3.1 |
| | Bags | 0.8 | 0.7 | 0.9 | 1.0 | 0.7 | 1.9 | 2.3 | 2.0 | 2.3 | 2.4 |
| | Foreground | 2.6 | 2.7 | 2.5 | 2.8 | 2.6 | 2.9 | 4.7 | 4.4 | 3.1 | 4.5 |
| | Umbrella | 1.6 | 0.7 | 1.2 | 0.9 | 0.9 | 2.1 | 2.0 | 2.3 | 2.3 | 2.1 |
| **Bark** | Bark | 1.3 | 2.7 | 1.8 | 2.0 | 1.5 | 2.0 | 3.6 | 2.8 | 2.3 | 2.8 |
| **Bird** | Bird | 0.9 | 0.9 | 0.8 | 1.2 | 0.7 | 2.5 | 4.8 | 2.6 | 3.1 | 3.8 |
| | Hand | 1.0 | 0.6 | 0.8 | 0.9 | 0.6 | 2.0 | 2.6 | 2.1 | 2.3 | 1.7 |
| **Bison** | Body | 1.8 | 2.3 | 2.5 | 3.2 | 2.8 | 2.3 | 5.5 | 4.2 | 3.1 | 4.9 |
| | Hair | 2.0 | 2.5 | 2.2 | 2.3 | 2.3 | 2.1 | 2.9 | 2.6 | 2.3 | 2.5 |
| | Head | 2.4 | 3.4 | 2.6 | 3.0 | 2.5 | 2.9 | 5.8 | 3.4 | 4.0 | 3.5 |
| | Horn | 1.7 | 2.0 | 2.3 | 2.0 | 2.2 | 3.4 | 5.4 | 4.7 | 3.4 | 5.3 |
| | Nose | 1.5 | 1.5 | 2.0 | 2.0 | 2.3 | 2.4 | 2.5 | 2.8 | 3.1 | 2.8 |
| **Field** | Cloud | 0.4 | 0.4 | 0.5 | 0.4 | 0.3 | 1.7 | 1.9 | 2.0 | 2.0 | 1.7 |
| | Grass | 2.4 | 3.2 | 2.9 | 2.6 | 2.4 | 3.4 | 5.3 | 4.4 | 3.9 | 4.6 |
| | Mountain | 0.7 | 0.6 | 0.7 | 0.9 | 0.5 | 2.1 | 2.4 | 2.0 | 2.0 | 2.4 |
| | Background | 0.5 | 0.4 | 0.4 | 0.6 | 0.3 | 1.9 | 1.9 | 1.8 | 1.8 | 2.1 |
| | Sky | 0.4 | 0.5 | 0.5 | 0.6 | 0.4 | 1.2 | 1.2 | 1.3 | 1.3 | 1.7 |
| | Trees | 2.4 | 2.5 | 2.2 | 2.1 | 1.7 | 2.5 | 2.7 | 2.3 | 2.5 | 1.7 |
| **Flower** | Flower | 1.3 | 1.0 | 0.9 | 1.4 | 0.9 | 3.1 | 4.5 | 3.6 | 3.6 | 3.5 |
| | Leaves | 1.0 | 0.8 | 0.7 | 0.9 | 0.7 | 1.9 | 1.9 | 2.0 | 1.8 | 2.1 |
| | Pistil | 2.4 | 1.5 | 2.3 | 2.2 | 2.6 | 2.1 | 3.2 | 2.6 | 3.1 | 2.8 |
| | Tip | 2.4 | 3.1 | 2.9 | 2.6 | 2.5 | 2.1 | 2.8 | 3.0 | 2.8 | 3.5 |
| **Gate** | Gate | 2.5 | 3.3 | 3.6 | 3.3 | 3.0 | 2.4 | 2.9 | 2.5 | 2.3 | 2.1 |
| | Left object | 1.9 | 2.0 | 2.1 | 2.1 | 1.9 | 2.3 | 3.6 | 2.6 | 2.3 | 2.1 |
| | Right object | 2.5 | 2.8 | 2.0 | 2.3 | 2.1 | 4.0 | 7.2 | 6.0 | 4.4 | 5.6 |
| **Hydrant** | Back snow | 1.4 | 1.6 | 1.5 | 1.5 | 1.5 | 2.7 | 5.2 | 4.1 | 3.4 | 3.9 |
| | Fore snow | 1.0 | 1.0 | 1.3 | 1.4 | 1.0 | 2.5 | 2.6 | 2.8 | 2.8 | 2.4 |
| | Hydrant | 2.1 | 2.0 | 2.1 | 2.0 | 1.9 | 2.1 | 3.1 | 2.3 | 2.0 | 1.7 |
| | Pole | 1.8 | 1.3 | 1.3 | 1.5 | 1.7 | 1.6 | 2.0 | 2.3 | 2.5 | 2.7 |
| **Leaf** | Leaf | 1.2 | 1.0 | 1.1 | 1.4 | 0.7 | 2.1 | 2.6 | 2.0 | 1.8 | 2.8 |
| | Shadow | 1.9 | 2.1 | 2.5 | 2.7 | 2.6 | 1.5 | 2.0 | 2.0 | 1.8 | 2.8 |
| **Pumpkin** | Back shadow | 2.2 | 2.9 | 2.3 | 2.6 | 2.7 | 1.6 | 2.6 | 1.5 | 2.0 | 1.7 |
| | Burlap | 1.6 | 1.8 | 1.5 | 1.8 | 1.4 | 2.4 | 2.9 | 2.0 | 2.8 | 2.1 |
| | Pumpkin | 0.7 | 0.4 | 0.8 | 0.9 | 0.5 | 2.3 | 2.0 | 2.0 | 2.3 | 2.1 |
| | Shadow | 1.0 | 1.2 | 1.2 | 1.7 | 1.3 | 1.2 | 1.2 | 1.3 | 1.3 | 1.7 |
| | Stem | 3.1 | 4.2 | 3.2 | 3.6 | 3.6 | 2.9 | 3.6 | 2.5 | 2.8 | 2.4 |
| **Rocks** | Rocks | 2.5 | 3.7 | 3.4 | 2.9 | 3.6 | 3.0 | 5.4 | 4.6 | 4.1 | 5.2 |
| **Teddy** | Bench | 1.5 | 2.5 | 1.9 | 2.5 | 2.1 | 2.3 | 3.6 | 2.6 | 3.7 | 3.2 |
| | Pole | 1.9 | 1.5 | 1.7 | 2.1 | 1.5 | 2.3 | 3.5 | 3.1 | 2.8 | 3.1 |
| | Teddy | 3.7 | 3.5 | 3.7 | 4.2 | 3.9 | 3.6 | 8.2 | 6.3 | 4.7 | 6.4 |
| **Valve** | Deck | 1.9 | 3.3 | 2.2 | 3.6 | 3.0 | 4.2 | 6.8 | 4.7 | 5.0 | 4.9 |
| | Gold valve | 1.8 | 1.9 | 2.2 | 2.3 | 1.5 | 2.9 | 3.1 | 3.6 | 2.8 | 2.7 |
| | Green hose | 2.8 | 3.7 | 3.4 | 3.8 | 3.3 | 1.9 | 3.1 | 1.5 | 1.8 | 2.5 |
| | Hose | 1.6 | 1.4 | 2.0 | 2.2 | 1.9 | 2.1 | 1.9 | 1.8 | 1.8 | 1.7 |
| | Rusts | 2.3 | 2.1 | 2.3 | 2.2 | 1.8 | 2.2 | 2.8 | 3.4 | 2.3 | 2.8 |
| | Silver valve | 1.9 | 1.6 | 1.9 | 2.5 | 2.0 | 2.3 | 2.5 | 3.6 | 3.1 | 3.5 |
| | T-joint | 1.5 | 1.3 | 1.6 | 2.0 | 1.6 | 3.2 | 3.7 | 3.9 | 3.4 | 3.9 |
| **Angel** | Angel | 2.2 | 2.1 | 2.2 | 2.4 | 2.2 | 3.6 | 4.7 | 4.7 | 3.1 | 4.7 |
| **Cat** | Cat | 3.0 | 2.4 | 2.6 | 3.2 | 2.8 | 4.6 | 7.0 | 6.5 | 5.0 | 6.0 |
| | Ground | 1.5 | 1.0 | 1.2 | 1.6 | 1.2 | 3.2 | 4.6 | 3.9 | 3.4 | 4.3 |
| | **Average** | 1.8 | 1.9 | 1.9 | 2.1 | 1.8 | 2.5 | 3.5 | 3.0 | 2.8 | 3.1 |

Not surprisingly, for textures containing a rather stochastic mixture of orientations (e.g., the teddy bear or rocks), equal improvements in quality could be obtained by using noise with a matching radial frequency spectrum, but with a white orientation spectrum, and with a random phase spectrum. For textures containing clearly dominant orientations, (e.g., the bison's hair), equal improvements in quality could be obtained by using noise with a matched orientation spectrum, but with a white radial frequency spectrum, and with a random phase spectrum.

## Discussion and Applications

Previous studies have shown that the addition of a subtle amount of noise can improve the perceived sharpness/quality of blurred/compressed images. In this study, we investigated whether further statistical and spectral shaping of the noise could lead to even greater quality improvements (i.e., by using the mid- and upper-frequency synthesized versions of the source textures).

### Discussion

The results of our first experiment revealed that, for the 50 texture segments tested here, approximately 35 (70%) were better improved via the addition of synthesized texture than noise. Of these successful cases, subjects reported both an increase in the perceived sharpness and the presence of additional structure which had been lost by the compression. Although the noise also helped in these regards, it did so to a much lesser extent due to the fact that too much noise resulted in a grainy appearance. In all cases, the noise had more high-frequency energy than the synthesized texture, suggesting that the presence of additional high-frequency content alone cannot fully account for the ability of noise to increase the perceived sharpness of a blurred region.

For the cases in which noise yielded better improvements, subjects reported that the addition of synthesized texture resulted in a fake appearance. Upon further examination of these unsuccessful cases, we found that either: (1) the synthesized textures themselves appeared to be poor visual representations of the source textures—i.e., they were failure cases for the texture-synthesis algorithm; or (2) the synthesized edges lined up with the distorted image's texture is a way that appeared artificial. Yet, despite this fact, the subjective contrast factors were always above zero, indicating that even the subtle addition of a mismatched texture is better than nothing.

As with noise, the hybrid test patterns can be considered low-risk (in terms of inducing a fake appearance) because they lack edges due to their random phase spectra. For the *Hybrid R* condition, subjects reported that the matched radial frequency spectra gave rise to the same visually pleasing clusters as the source texture, but, as designed, lack orientation. For the *Hybrid O* condition, subjects reported that the matched orientation spectra gave rise to similar overall "brush angle" appearances as the source textures, but with a very fine brush that lacked medium- and high-contrast edges. For some textures, either of these attributes alone yielded equivalent or even better quality improvements than the synthesized textures. The *Hybrid 2D* condition yields greater quality improvements than *Hybrid R* and *Hybrid O*, again, owing to the fact that there is less of a chance of improperly aligned edges that might appear artificial.

The fact that synthesized texture can yield equal or better quality improvements (and perceived sharpness) than white noise

suggests that a key reason for the increase in quality/sharpness found in previous experiments is due to the fact that the noise helps to synthesize lost details, and in many cases the synthesized textures did a better job at this task. Furthermore, the results of the second experiment suggest that the synthesis need not be explicit; rather, the additive pattern can facilitate the HVS's ability to perform its own synthesis of the missing details (what we will call "perceptual synthesis"). This conclusion is related to the notion of memory texture proposed by Wan *et al.* [6], though our "perceptual synthesis" may also or alternatively be driven by simpler statistical inferences. Furthermore, it would seem that this perceptual synthesis makes use of or is otherwise influenced by radial frequency and/or orientation information.

### Predicting the Subjective Contrast Factors

In order to use additive synthesized textures (or shaped noise) in an actual enhancement algorithm, it is necessary to predict the mean-opinion contrast factors for the to-be-enhanced textures. Specifically, we seek an automated means of determining the value $\bar{c}_i$ in Equation (2) for any given texture segment. Here, we provide the preliminary results of our current technique for predicting these contrast factors.

We assume that the contrast factors are low when the synthesis is itself is a poor perceptual match to the source texture, and high when the synthesis is a good perceptual match to the source texture. To this end, we use two simple features: (1) The Kullback-Liebler divergence between histogram-of-oriented-gradients for the source vs. synthesized; and (2) a texture-regularity measure applied only to the source texture (under the assumption that the synthesis is best for stochastic, highly irregular textures). These features are briefly defined as follows:

- *Divergence in HoGs*: Let $\mathbf{h}_{org}$ and $\mathbf{h}_{syn}$ denote the HoG vectors for the source texture and synthesized texture, respectively. Our first feature, $f_1$, is defined as the logarithm of the KLD between $\mathbf{h}_{org}$ and $\mathbf{h}_{syn}$.
- *Source Texture Regularity*: We use the algorithm by Anwar *et al.* [12] to obtain a numerical estimate of the regularity of the source texture. This regularity algorithm was designed for larger textures, and thus we applied the algorithm in a block-based fashion ($16 \times 16$ blocks) and combined the per-block regularities into a single overall scalar via a 4-norm. Our second feature, $f_2$, is defined as this scalar.

We performed a linear regression using standardized versions of $f_1$ and $f_2$ as regressors, and using the standardized mean-opinion contrast factors as the dependent (to-be-predicted) variable. Table 2 lists the standardized regression results. The regression demonstrates that both factors are statistically significant, and that the regularity measure has 2-3$\times$ more weight toward the overall prediction. Note, however, that the regression was performed on the entire dataset, and thus this analysis is not meant to be indicative of the performance of this approach on other textures. Instead, this analysis demonstrates that the features show promise in predicting the contrast factors. For reference, the results of a 50%-50% train-test support-vector-machine regression using the same factors yields a correlation coefficient of $R = 0.68$. The correlation drops, but still demonstrates promise under this more realistic setting. The performances are still well below that

**Table 2: Standardized regression results ($\bar{c}_i^j$ on $f_1$ and $f_2$)**

| $R = 0.74$; Std. Error = 0.688; 50 samples | Coefficient | p-value |
|---|---|---|
| Standardized KLD of HoG (standardized $f_1$) | -0.667 | 1.9E-08 |
| Standardized Regularity (standardized $f_2$) | -0.265 | 9.9E-03 |

**Table 3: Performances of various IQA algorithms in predicting the MOS values for the distorted and enhanced images**

| | Distorted images | | Enhanced images | | All 30 images | |
|---|---|---|---|---|---|---|
| Algorithm | SROCC | CC | SROCC | CC | SROCC | CC |
| PSNR [13] | 0.13 | 0.19 | 0.37 | 0.27 | 0.36 | 0.25 |
| SSIM [14] | 0.61 | 0.61 | 0.11 | 0.07 | 0.20 | 0.19 |
| MS-SSIM [15] | 0.62 | 0.73 | 0.16 | 0.14 | 0.24 | 0.25 |
| VIF [16] | 0.63 | 0.66 | 0.41 | 0.32 | 0.12 | 0.21 |
| MAD [17] | 0.36 | 0.37 | 0.06 | 0.05 | 0.41 | 0.37 |
| DIIVINE [18] | 0.25 | 0.25 | 0.13 | 0.03 | 0.44 | 0.45 |
| C-DIIVINE [19] | 0.17 | 0.19 | 0.15 | 0.07 | 0.62 | 0.64 |
| BLIINDS-II [20] | 0.10 | 0.00 | 0.08 | 0.14 | 0.67 | 0.70 |
| BRISQUE [21] | 0.11 | 0.18 | 0.03 | 0.04 | 0.61 | 0.64 |
| DESIQUE [22] | 0.27 | 0.28 | 0.51 | 0.38 | 0.79 | 0.76 |

obtained when using one subject's contrast factors to predict the average of the other subjects' contrast factors ($R > 0.9$).

### *Predicting the Quality Improvements and Full-Image Quality Scores*

We have not yet explored techniques for predicting the per-segment quality improvements. However, we did obtain MOS values for the distorted and fully-enhanced images (with all segments restored via the proposed technique) obtained in a full-reference setting, where the original images served as the reference images. Thus, there were 15 distorted images and 15 enhanced images, and their associated mean opinion scores (MOS) ranging from 0 to 10.

Here, we briefly report our preliminary survey of using six full-reference and five no-reference image quality assessment algorithms in predicting these MOS values. Due to space limitations, we summarize the results only via Table 3. Not surprisingly, the full-reference techniques can perform reasonably well when given only the distorted images, but the enhanced images appear to confuse most algorithms. Only the no-reference techniques can cope with the two sets as a whole (last columns), suggesting that only the no-reference approaches can begin to predict the enhanced images as being of higher quality than their respective distorted versions.

## Conclusions

This paper presented a study designed to provide insights into the effects that additive white noise and other related test patterns have on the visual quality of compressed (blurred) images. We found that the use of the statistically and spectrally shaped noise can lead to markedly greater quality improvements as compared to white noise, despite the fact that the textures enhanced with the shaped noise patterns, when scaled to their optimal contrasts, contained equal or lesser high-frequency content. This finding suggests that the improvements provided by noise and its related shaped variants are due in large part to their ability to synthesize details lost by the compression/blurring.

## References

[1] M.-Y. Shen and C.-C. Kuo, "Review of postprocessing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2–14, March 1998.

[2] S. A. Golestaneh and D. M. Chandler, "Algorithm for JPEG artifact reduction via local edge regeneration," *J. Electron. Imaging*, vol. 23, no. 1, pp. 013018, February 2014.

[3] V. Kayargadde and J. Martens, "Perceptual characterization of images degraded by blur and noise: experiments," *J. Opt. Soc. Am. A*, vol. 13, no. 6, pp. 1166–1177, 1996.

[4] Y. Kshibuchi, N. Aoki, M. Inui, and H. Kobayashi, "Improvement of description in digital print by adding noise," *Journal of The Society of Photographic Science and Technology of Japan*, vol. 66, no. 5, pp. 471–480, 2003.

[5] T. Kurihara, N. Aoki, and H. Kobayashi, "Analysis of sharpness increase by image noise," in *Proc. SPIE*, 2009, vol. 7240, pp. 724014–724014–9.

[6] X. Wan, H. Kobayashi, and N. Aoki, "Improvement in perception of image sharpness through the addition of noise and its relationship with memory texture," in *Proc. SPIE*, 2015, vol. 9394, pp. 93941B–93941B–11.

[7] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial sharpness measure of local perceived sharpness in natural images," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 934–945, 2012.

[8] D. M. Chandler, K. H. S. Lim, and S. S. Hemami, "Effects of spatial correlations and global precedence on the visual fidelity of distorted images," in *Proc. SPIE Human Vision and Electronic Imaging XI*, B. E. Rogowitz, T. N. Pappas, and S. Daly, Eds., San Jose, CA, 2006.

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

[10] A. Olmos and F. A. A. Kingdom, "Mcgill calibrated colour image database," http://tabby.vision.mcgill.ca.

[11] A. B. Watson, M. Taylor, and R. Borthwick, "Image quality and entropy masking," *Human Vision, Visual Processing, and Digital Display VIII, Proc. SPIE*, vol. 3016, pp. 2–12, 1997.

[12] K. Anwar, A. Harjoko, and S. Suharto, "A new method for measuring texture regularity based on the intensity of the pixels in grayscale images," *International Journal of Computer Applications*, vol. 137, no. 7, pp. 1–5, March 2016, Published by Foundation of Computer Science (FCS), NY, USA.

[13] A. T1.TR.74-2001, *Objective Video Quality Measurement Using a Peak-Signal-to-Noise-Ratio (PSNR) Full Reference Technique*, 2001.

[14] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, 2004.

[15] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, Nov. 2003, vol. 2, pp. 1398–1402.

[16] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.

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

[18] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment:

from natural scene statistics to perceptual quality.," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350–64, 2011.

[19] Y. Zhang, A. K. Moorthy, D. M. Chandler, and A. C. Bovik, "C-DIIVINE: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes," *Signal Processing: Image Communication*, vol. 29, no. 7, pp. 725–747, 2014.

[20] M. A. Saad and A. C. Bovik, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, pp. 3339–3352, 2012.

[21] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. PP, no. 99, pp. 1, 2012, in press.

[22] Y. Zhang and D. M. Chandler, "No-reference image quality assessment based on log-derivative statistics of natural scenes," *Journal of Electronic Imaging*, vol. 22, no. 4, pp. 043025–043025, 2013.