Quality of Small Images

Karin Töpfer, Eastman Kodak Company, Rochester, New York/USA

Abstract

Today, digital cameras and many handheld electronic devices contain small image displays. Predictive system image quality models are needed to facilitate the optimization of such systems. Therefore, the quality of small images was investigated as a function of the angular subtense of the displayed images for several important image quality attributes. At constant levels of physically measurable degradations, the perceived image quality losses increased with decreasing image size. Moreover, at constant size reduction, the largest degradation was observed for attributes involving spatial frequency distortion, followed by sharpness and noise; tonal clipping showed no size variation. All results were quantified in terms of 50% just-noticeable differences of quality loss for use in multivariate system image quality models.

Introduction

In connection with the development of multiresolution image formats for PhotoCD, FlashPix, and other applications, it was observed that small images looked less sharp than predicted based on the overall system modulation transfer function (MTF) and common acutance metrics, which were found to be good predictors of perceived sharpness for larger images [1]. A possible cause for this effect is that important detail in the images, e.g., eyes, hair, letters, and texture, decreases in proportion to the decreasing image size and thus becomes relatively more obscured if blurred. This effect is of particular interest today as small images have become ubiquitous on camera backs, cell phones, and other handheld devices. In order to optimize the overall quality of images displayed on such devices under a variety of system constraints (e.g., display resolution, cost, speed), it is necessary to trade off many different image quality attributes, such as sharpness, noise, aliasing, and compression artifacts, against each other. The goal of this investigation was to quantify the loss in image quality with image size at a given physically measurable level of degradation for a variety of different image quality attributes. Using these results, predictive image quality models can be developed for imaging systems that display small images.

Theoretical Background

As described previously [2a], the concept of a just-noticeable difference (JND) is central to our image quality framework. It allows us to describe the effect of all attributes on image quality in similar terms so that they can be rigorously combined into a prediction of overall quality in the presence of multiple attributes. A 50% JND of quality is defined as the stimulus difference that would lead to a 75%:25% outcome in a paired comparison task.

One successful way to build up a comprehensive image quality model is to study many different perceptually independent image quality attributes in isolation, and to develop objective metrics, which are in turn based physical measurements, e.g., the system MTF for sharpness. An objective metric is a single number that may be determined through objective means and is correlated with a perceived attribute of quality in an image, accounting for its viewing conditions and the properties of the human visual system [2b]. For example, acutance metrics, which cascade the system MTF with the contrast sensitivity function of the human visual system, are good correlates of sharpness [3].

Once the effect of the individual attributes on image quality, ΔQ_i, is known, either through experimentation, or from predictions based on the objective metrics, a Minkowski metric with a variable exponent n_m is applied to predict the overall image quality degradation, ΔQ_m, of a system containing multiple attributes [2c]:

$$\Delta Q_m = -\left(\sum_i \left(-\Delta Q_i\right)^{n_m} \right)^{1/n_m} \tag{1}$$

$$n_m = 1 + 2 \cdot \tanh\left(\frac{-\Delta Q_{max}}{16.9} \right), \tag{2}$$

where ΔQ_{max} is the most severe individual degradation.

This approach allows us to optimize the overall image quality of complex imaging systems by trading off contributions from different attributes. For example, it is possible to compensate for unsharpness by applying unsharp masking, however, this may increase the perceived noise and may introduce edge artifacts caused by oversharpening. By obtaining image size-dependent corrections of the 50% JNDs of quality for the individual attributes, these powerful predictions can be extended to systems displaying small images.

Experimental Design

While it was desirable to include a wide variety of different image quality attributes in the study, the selection was limited by practical considerations, in particular, if the attribute could be simulated in isolation without causing image quality degradations of other attributes, and whether it was possible to generate a continuously varying series of image quality levels for the attribute. The attributes included in the study were: (1) unsharpness (Sh); (2) isotropic noise (N), giving a grainy, speckled, or salt-and-pepper appearance; (3) misregistration (MR), i.e., the imprecise alignment of the color planes of an image, leading to unsharpness or, in extreme cases, to superimposed sharp images of different color; (4) tonal clipping (TC), the irreversible loss of highlight or shadow detail caused by tonal limitations; (5) oversharpening (OS), which appears as harshness, edge overshoot, and/or auras caused by excessive frequency boost; and (6) reconstruction error (RE), i.e., high-frequency artifacts caused by poor interpolation methods such as pixel replication, which manifest themselves as ringing on edges, jaggies (a stair-step appearance in diagonal lines), and, in extreme cases, pixellization.

In order to fully characterize the corrections required for small images, a size and magnitude series was generated for each attribute. The image size was characterized by the angular subtense, s, i.e., the angle in degrees subtended by the image diagonal, d, at the eye:

$$s = \frac{2 \cdot 180}{\pi} \cdot \tan^{-1}\left(\frac{d}{2 \cdot v}\right) \approx \frac{180}{\pi} \cdot \frac{d}{v} \qquad (3),$$

where v is the viewing distance. In our previous studies with larger images, 4×6-inch prints were viewed at a viewing distance of 16 inches, resulting in a reference value of the angular subtense, $s_0 = 25.4°$.

For the size series, the viewed angular subtense was reduced by factors of 2, 3, and 4. Three levels representing different severities of the artifact were presented at the medium size reduction by a factor of 3. The sharpness experiment was conducted at viewing distances of 20 and 40 inches to determine if the correction factor was a function of the angular subtense alone. If this were the case, a full-size image viewed at 40 inches should require the same correction as an image subsampled by a factor of two and viewed at half the distance. Table 1 shows the expected 50% JNDs of quality loss of all experimental levels for large images viewed at the reference angular subtense, s_0. Four scenes were selected for the evaluation of each attribute. They were chosen according to the classification into more or less sensitive scenes, which resulted from previous studies with the large images. Each set contained two average scenes, one scene that was less sensitive to the artifact, and one scene that was more sensitive. The data in Table 1 refer to the average scene.

Table 1: Angular subtense and –50% JNDs of quality by attribute for levels 1–6; Sh: sharpness, N: noise, TC: tonal clipping, MR: misregisration, OS: oversharpening, RE: reconstruction error.

	1	2	3	4	5	6
	Viewing distance = 20 inches					
s/degrees	25.4	12.9	8.61	8.61	8.61	6.45
Sh	5.43	5.11	5.09	1.52	8.90	5.00
N	6.00	6.00	6.00	1.05	9.00	6.00
TC	6.00	6.00	6.00	2.04	10.1	6.00
MR	5.75	5.75	5.75	2.40	9.75	5.75
RE	7.13	7.17	7.51	3.61	13.1	7.54
OS	5.16	5.18	5.17	1.10	5.14	5.16
	Viewing distance = 40 inches					
s /degrees	12.9	8.61	6.45	6.45	6.45	4.30
Sh	3.86	3.77	3.74	0.43	7.66	3.71

Image Evaluation

The images were evaluated using the softcopy ruler technique. The test images and the ruler images, which represented a series of reference images of varying severity of the artifact under investigation, were displayed side by side on an IBM T-221 LCD monitor with 200 pixels per inch (ppi) resolution, an NVidia Quatro FX 3000 graphics card, and 3840×2400 addressable pixels. The series of reference images was displayed at the reference angular subtense, $s_0 = 25.4°$, and calibrated based on previous studies such that consecutive images varied by exactly 1 50% JND of quality. In order to achieve this calibration, the MTF of the monitor was measured, the monitor was color-calibrated for a CIE Standard Illuminant D65 white point, and the viewing environment was tightly controlled in terms of viewing distance and luminance levels (the white point luminance of the monitor was set to 178 cd/m^2, and the images were displayed using a 20%

gray surround). A slider on the user interface allowed the user to quickly bring up consecutive ruler images and to compare them with the test images. Normally the judges are asked to match the quality of the test print with the quality of the ruler. In this case, however, they were instructed to match the appearance of the artifact in the test image with that of the ruler image. The goal was to obtain correction factors for different image quality attributes using rulers, which were previously calibrated in terms of 50% JNDs of quality. The calibration of the rulers eliminates the problem usually encountered in appearance matching, where such a technique produces attribute JNDs instead of quality JNDs [2d].

The presentation sequence of the images was randomized by scene and level. Approximately 10 judges drawn from the research community, and tested for visual acuity and color vision, participated in each of the attribute studies.

Results and Discussion

The data collected from the softcopy ruler workstation were calibrated in terms of 50% JNDs of quality, and the mean over all scenes and observers was taken. The additional quality degradation incurred by the smaller angle subtended by the test images was obtained by subtracting the predicted 50% JNDs of quality for images at the reference subtense, s_0, from the corresponding JNDs obtained in the experiment. Figure 1 shows this difference as a function of the angular subtense for unsharpness and tonal clipping.
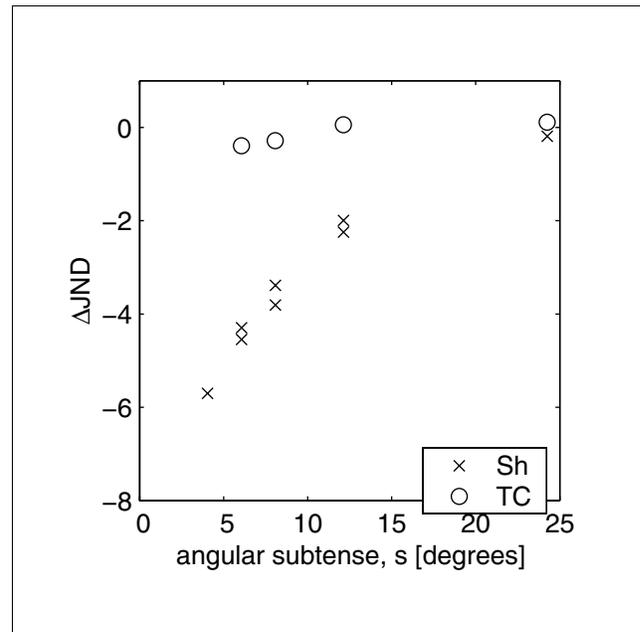
Figure 1. *Additional quality degradation, ΔJND, as a function of the angle, s, subtended by the image diagonal.*

As discussed in the introduction, the additional quality loss for small images is mainly expected to be a perceptual effect that is caused as important detail in the images decreases in size and becomes more obscured by image structure artifacts. Consequently, no additional quality loss at smaller sizes was expected for color and tone attributes, e.g., for tonal clipping. The

data in Fig. 1 confirm this hypothesis. However, significant additional quality loss was perceived for the unsharp small images. Figure 1 shows two sharpness data points for the three intermediate values of angular subtense. These points refer to the two different viewing distances and confirm that the effect is solely driven by the angle subtended by the image diagonal.

In addition, the magnitude of the correction required at a given angular subtense, s, increases in proportion to the magnitude of the 50% JND quality loss, JND, observed for larger images viewed at the reference angular subtense, s_0:

$$\Delta JND = \left| \begin{array}{ll} m \cdot \left| JND + JND_0 \right|^{\delta} \cdot \ln\left(\dfrac{s}{s_0}\right) & \text{for } s \leq 0 \\ \\ 0 & \text{otherwise} \end{array} \right. \quad (4)$$

The attribute-specific parameters m, δ, and JND_0, obtained by nonlinear regression, are summarized in Table 2. Figure 2 illustrates that the measured corrections, ΔJND, for the small images are well predicted according by Eq. 4 in combination with Table 2.

Table 2: Regression parameters for Eq. 4 by attribute.

Attribute	m	δ	JND_0
Sh	2.3772	0.1766	-0.1
N	0.6340	0.3949	0.0
MR	1.3066	0.5380	0.0
RE	2.2669	0.4580	0.0
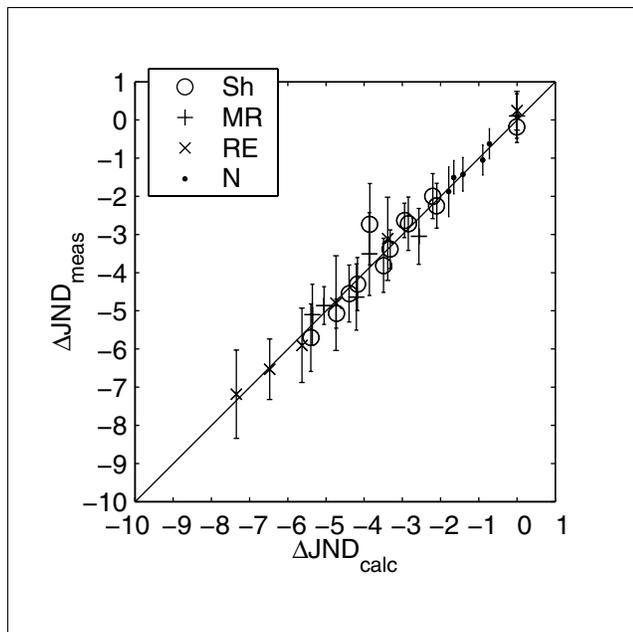
Figure 2. Measured image quality corrections vs predictions according to Eq. 4 and Table 2 for all levels of four different attributes. The error bars correspond to the 95% confidence interval of the measured JNDs.

The magnitude of the corrections depends on the type of attribute. No corrections were required for tonal clipping or oversharpening, which is a small correction to the sharpness metric

(so neither is shown in Fig. 2). Small corrections were observed for noise. Larger corrections were required for sharpness-related attributes and misregistration. The largest corrections were found for an attribute that represented a more complex distortion of the spatial frequency spectrum of the image, reconstruction error.

If the corrections are applied in JND space, it is expected that unstudied attributes can be treated like the most similarly studied attribute. The attributes studied can be classified into four categories: color and tone attributes, sharpness-related attributes, noise-related attributes, and attributes related to spectral distortion; it is expected that a common correction is required within a category. Such categorization of attributes is supported by the data for misregistration and unsharpness, which can be fit with a single set of regression parameters. Both attributes have a similar perceptual appearance at small degradations; small shifts between the color records, typical of misregistration, appear as unsharpness.

Figure 3 provides an illustration of the required corrections as a function of the quality loss predicted for the reference subtense, s_0, if the size of the displayed image is reduced threefold. The reference case is equivalent to viewing a 4×6-inch print at a 16-inch viewing distance ($s_0 = 25.4°$). The data from the sharpness experiment suggested some additional image quality loss for small images even if the equivalent larger images had no perceived degradation, described by the parameter JND_0 in Table 2. The actual magnitude of this offset was calculated by setting JND to zero in Eq. 4, and amounted to –1.75 JNDs for the case shown in Fig. 3. There was no visual evidence, however, that any of the other attributes (noise etc.), which were below threshold for large images, went above the perceptual threshold for smaller images ($JND_0 = 0$ for all other cases). This special treatment of sharpness is plausible, because some unsharpness is nearly always evident in pictorial images even at the highest quality levels, so it is generally suprathreshold and amenable to quality shifts.
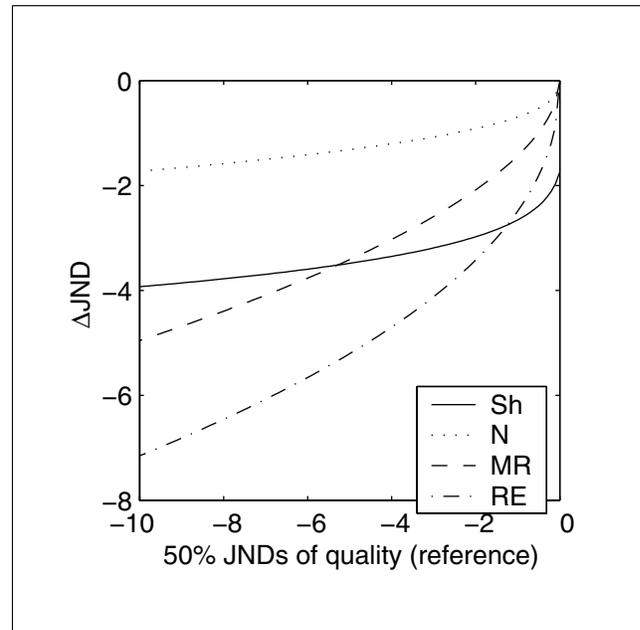
Figure 3. JND corrections for 4 different attributes as a function of the quality loss for large images at the reference subtense, s_0 (3X size reduction).

The following example illustrates how the JND corrections derived in this experiment can be used in conjunction with the multivariate formalism to optimize overall image quality. Let us consider a hypothetical system, which displays images captured with a 5 Megapixel digital camera on a softcopy device. In the first case the image is interpolated by a factor of 0.67 and then displayed and viewed on a high-resolution 200 ppi LCD monitor at a distance of 24 inches. At a 10.82-inch image diagonal, this viewing environment corresponds to the reference angular subtense $s_0 = 25.4°$. In the second case, after further prefiltration and downsampling by a factor of 3, the image is displayed on a LCD camera back with 200 ppi resolution and viewed at a distance of 14 inches. The angle subtended at the eye by this image is 11.03°.

Some additional improvement of the overall image quality can be made in both cases by adding an unsharp masking operation, and the goal is to optimize the boost factor and to assess the magnitude of the improvement. The relevant image quality attributes are unsharpness, chromatic aliasing, i.e., colored fringing, a result of undersampling by the camera color filter array, and oversharpening, which may be introduced by the unsharp masking operation.
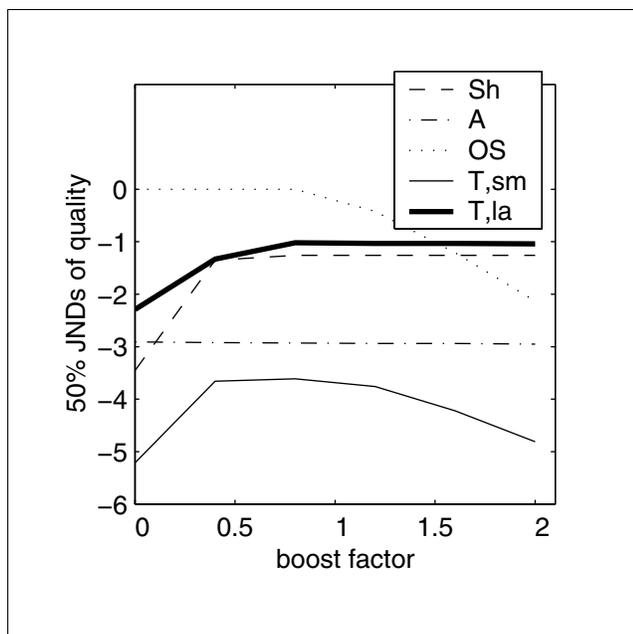
Figure 4. *Quality degradations caused by unsharpness, oversharpening, aliasing (A), and overall quality (T,sm) as a function of the unsharp masking boost factor for small images (subtense s = 11.03^0). Thick solid line: predictions of overall image quality at the reference angular subtense s_0 = 25.4^0 (T,la).*

Figure 4 shows the predicted degradations in terms of 50% JNDs of quality for these attributes, as well as the multivariate sum corresponding to overall quality, calculated according to Eqs. 1 and 2. The quality degradations for the larger image, displayed at the reference angular subtense, consist of some initial unsharpness and 1 JND of aliasing, which is essentially unaffected by the unsharp masking operation, because the aliased artifacts are low

frequency. The thick, solid line in Fig. 4 denotes the overall quality predictions for this case. The sharpness loss can be fully recovered with a boost factor of 0.8. At this point the overall image quality has reached its maximum and further increases in sharpness associated with higher boost factors do not lead to improvements. At the modest boost levels shown in Fig. 4, oversharpening is not an issue for the larger image, although it will eventually decrease the overall image quality.

For the small images, the image quality without additional sharpening is lower. This is mainly a result of the perceptual effect investigated in this study. Both sharpness and aliasing require additional corrections in JND space according to Eq. 4 for smaller size images. The regression parameters listed for reconstruction error in Table 2 were used for aliasing. The unsharp masking initially helps to recover some of the sharpness loss up to a boost factor of 0.5. As discussed, the JNDs for sharpness approach an asymptotic value below zero because of the additional quality offset discussed in connection with Table 2 and Fig. 3. At boost factors above 1.0, the overall image quality quickly deteriorates because of oversharpnening. Although no image size-dependent corrections are required in this case, oversharpening becomes more critical because of the additional downsampling and associated frequency remapping for the small image. This analysis shows that the selection of the correct unsharp masking boost factor is critical for the small image, while wider margins exist for the larger image.

Conclusion

At a constant, physically measurable level of image structure attributes, the perceived image quality decreases with image size, characterized by the angle subtended by the image diagonal at the eye. This effect is strongest for attributes involving spectral distortion, followed by sharpness-related attributes and noise attributes, while no effect is observed for color and tone-related attributes. The quantitative characterization of this effect performed in this study allows us to make predictions of overall quality for complex imaging systems displaying small images.

Acknowledgements

References

[1]   R. Parada, H. Muammar, J. Milch, private communication.
[2]   B.W. Keelan, "Handbook of Image Quality," Marcel Dekker, Inc., New York, 2002, (a) pg. 41, (b) pg. 171, (c) pg. 161, (d) pg. 98.
[3]   E.M. Crane, Proc. SPIE 310, 125 (1981).

Author Biography

Karin Töpfer received her Masters degree in Physics from Dresden University of Technology and a Ph.D. in Photophysics from Dresden University of Technology. Since 1993, she has worked at Eastman Kodak Company, first in the UK and later in Rochester, NY. In recent years, her work has primarily focused on predictive image quality modeling, including color quality, and on psychophysics. She has recently extended this expertise to medical imaging systems. She is the co-author of three chapters in the Handbook of Image Quality: Characterization and Prediction (New York: Marcel Dekker, 2002, ISBN 0-8247-0770-2).