Perceptual error measure and its application to sampled and interpolated single-edged images

Citation for published version (APA):

DOI:
10.1364/JOSAA.14.002111

Document status and date:
Published: 01/01/1997

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 30. Sep. 2023
Perceptual-error measure and its application to sampled and interpolated single-edged images

M. R. M. Nijenhuis and F. J. J. Blommaert

Institute for Perception Research, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Received May 1, 1996; revised manuscript received March 10, 1997; accepted May 19, 1997

Error metrics quantify the difference between a reproduced image and the corresponding unprocessed original image. A drawback of the commonly used metrics such as the mean square error is their poor correlation with the perceived quality of the reproduced image. We present a framework for an alternative metric that uses the distance in a perceptual space to predict the perceived impairment of reproduced images. The perceptual space is spanned by perceptual artifacts that are introduced by image-reproduction techniques. For image reproduction using sampling and interpolation it is shown how such a multidimensional space can be determined from the image. The sensory strengths of the artifacts’ periodic structure and blur are two of the orthogonal dimensions of this space. In addition, we demonstrate that, after the perceptual-error metric is calibrated to a particular observer, this metric can successfully predict experimentally determined subjective image quality of sampled and interpolated simple black-and-white images. © 1997 Optical Society of America [S0740-3232(97)04109-4]

1. INTRODUCTION
Displays often reproduce images that are visually impaired compared with the originals, i.e., images produced by a hypothetical ideal display. For example, the majority of electro-optical displays, including liquid-crystal video displays and CRT’s, produce sampled and interpolated images in which typical artifacts such as blur, periodic structure, and moiré may be visible. Measurements of perceptual image quality, i.e., the degree of excellence of an image, should be used to evaluate the images reproduced by displays. Perceptual image quality also makes possible the perceptual optimization of the physical parameters that specify the display.

Error metrics that predict perceptual image quality are popular, because they make time-consuming experiments involving human subjects redundant. In general, the perceptual-image-quality values predicted by error measures such as the mean square error (e.g., Oakley and Cunningham) do not correlate well with the results of subjective tests, because these measures do not incorporate properties of the human visual system (e.g., Pratt and Mannos and Sakrison).

At present, there is a growing interest in error measures that use properties of the visual system. A description of one of these properties is the modulation transfer function of the eye (e.g., Campbell and Robson and Kelly), which specifies the threshold modulation depth of sine gratings as a function of spatial and temporal frequencies of the sine grating. It is a popular description in the frequency domain that can easily be included in error measures by using a Fourier description of the sampling and interpolation problem (e.g., Watson et al.).

Although such error metrics successfully model properties of the front end of the visual system, these models generally cannot handle properties of high-level vision such as personal preference and cognitive aspects. Therefore we introduce a perceptual-error measure that is characterized by an intermediate perceptual space spanned by the relevant perceptual attributes underlying perceptual image quality. Front-end properties are then used to determine the strengths of these attributes from physical parameters that specify the image and the display. The process that combines the underlying attributes into perceptual quality reflects properties of high-level vision. The perceptual-error measure uses only a single parameter to account for individual differences in high-level vision. The perceptual-error measure is based on global knowledge about the structure of the perceptual process that leads to judgments about the quality of images. This new measure models the perceptual processes that are relevant to quality judgments rather than modeling the functions of (groups of) cells in the visual system (e.g., Ref. 8). Although the latter approach may work well for low-level vision, it is difficult to model high-level vision (grouping, cognition) in such a way.

In this paper we will specify the perceptual-error measure and work it out in detail for sampling and interpolation. Before the performance of the model is tested, we will first determine and study its parameters. We restrict ourselves to a simple black-and-white image consisting of a single edge. This single-edged image is relatively easy to describe yet it does not oversimplify the problem since both of the most dominant sampling and interpolation artifacts, periodic structure and blur, can appear in this image.

2. RATIONALE OF SINGLE-EDGED IMAGES
A single-edged image as indicated in Fig. 1(a) is formed by two adjacent uniform regions with different lumini.
In this section we review the perceptual error measure for sampled and interpolated images presented in Refs. 9–11. The diagram in Fig. 2 shows the computational steps used for determining the perceptual image quality $Q$ from the physical parameters $\Phi_i$ that specify both the image and the sampling and interpolation process. The perceptual-error measure is distinguished from conventional error measures by a perceptual space with the sampling and interpolation artifacts as orthogonal dimensions. Hence the perceptual-error measure explicitly incorporates the dimensions underlying the visual impairment of the images.

As mentioned, typical sampling and interpolation artifacts that may occur in the single-edged image are periodic structure and blur. If we vary the Michelson contrast or the average luminance of this image, we also expect some effect of these variations on perceptual image quality. Therefore we also include in the error measure the perceptual attributes induced by the Michelson contrast and the average luminance. These attributes are brightness contrast and brightness, respectively. The influence of these attributes is intuitively expected to be second order if the values for both $L_1$ and $L_2$ are representative for standard monitors and TV sets, which have a relatively small luminance range. We expect that the total perceptual impairment of the single-edged image will decrease if the brightness or the brightness contrast increases. We therefore use the complementary perceptual attributes as artifacts. Since proper terminology for these artifacts is lacking, we adopt the terms complement of brightness and complement of brightness contrast.

Perceptual image quality $Q$ is related linearly to perceptual impairment $I^{12,13}$:

$$I = 1 - Q.$$  

(1)

The perceptual attributes are defined in the interval $[0, 1]$ for computational convenience.

Besides judging total perceptual impairment, people can also distinguish among the various underlying artifacts and are able to judge the perceptual impairments of these underlying artifacts separately. Minkowski metrics can be used to combine the $M$ underlying perceptual impairments into the total impairment $I^{13,14}$.

$$I^a = \sum_{i=1}^{M} I_i^a,$$

(2)

which reduces to

$$I^a = I_p^a + I_b^a + I_{C^*}^a + I_{B^*}^a,$$

(3)

if we consider only the impairments of the artifacts’ periodic structure $I_p$, blur $I_b$, complement of brightness contrast $I_{C^*}$ and complement of brightness $I_{B^*}$. Since the artifacts are the orthogonal dimensions of a perceptual space, the total impairment can alternatively be interpreted as the distance in this perceptual space between the sampled and interpolated image and the original image that is placed in the origin. The original image is an image produced by a hypothetical ideal display. For $a = 2$ the distance is Euclidean.

The perceptual impairment of the underlying artifacts periodic structure and blur was found to be linearly re-

---

**Fig. 1.** (a) Single-edged image with two uniformly illuminated regions separated by a horizontal straight edge. The image is completely specified by these two physical parameters or, alternatively, by the average luminance $L = (L_1 + L_2)/2$ and some luminance-contrast measure such as the Michelson contrast $C = (L_1 - L_2)/(L_1 + L_2)$. The horizontal orientation of the edge has been chosen for convenience.

(b) Magnified part of the single-edged image with a columnar sampling structure with sampling distance $d$ and column width $w$. The luminance of the columns is a factor $d/w$ higher than the uniform regions in (a), so that the average luminances of the lower and upper regions remain equal to $L_1$ and $L_2$, respectively.

(c) Impulse response of an elliptical optical Gaussian interpolation filter with spread parameters $\sigma_v$ (horizontal) and $\sigma_h$ (vertical).

**Fig. 2.** Schematic diagram of the computational steps of the perceptual-error measure that relates perceptual image quality $Q$ or total perceptual impairment $I$ to the physical parameters $\Phi_i$, with use of an intermediate perceptual space with sensory strengths $S_i$, or impairments $I_i$ of the individual sampling and interpolation artifacts along the orthogonal axes. The most dominant artifacts, periodic structure and blur, are indicated by the indices $p$ and $b$, respectively.
lated to the perceptual strengths of these artifacts.\textsuperscript{11} We assume that this linear relation holds for other artifacts as well:

\[ I_i = a_i S_i. \]  

(4)

The interpretation is that the strengths of the perceptual attributes are attributes of low-level vision, whereas the impairments are cognitive attributes. The constants \( a_i \) represent the relative weights of the artifacts in the total perceptual impairment.

According to Eqs. (3) and (4) the total perceptual impairment of a sampled and interpolated image can be written as

\[ I^a = a_p^p S_p^a + a_p^o S_p^o + a_c^c S_c^a + a_b^b S_b^a. \]  

(5)

We assume that \( I = 0 \) for the original image. If both the average luminance and the Michelson contrast have fixed values, then we can use a simplified version of Eq. (5) for calculating the difference between the original and the sampled and interpolated images:

\[ I^a = a_p^p (S_p^a + \lambda_b S_b^o). \]  

(6)

The perceptual parameter \( \lambda_b = a_b^b/a_p^p \) expresses the observer’s weight with respect to the quality of the blur artifact relative to the observer’s weight with respect to the quality of the periodic-structure artifact. Equation (6) shows that the optimization of the sampling and interpolation problem is equivalent to minimizing the cost function

\[ S_p^a + \lambda_b S_b^o. \]  

(7)

Minimization of cost functions is a well-known variational regularization solution method for ill-posed problems.\textsuperscript{15}

The solution is a compromise between conflicting demands, i.e., between the absence of periodic structure and the absence of blur. The regularization parameter \( \lambda_b \) controls the relative importance of the constraints. It tells us something an observer’s weight for periodic structure compared with the observer’s weight for blur. Note that the value of \( \lambda_b \) may depend both on the observer and on image content.

On the basis of Eq. (1) the difference in perceptual quality between an original image and a sampled and interpolated version of this image simply equals the total impairment caused by sampling and interpolation. Equation (6) can be written to represent this difference as

\[ I = a_p^p (S_p^a + \lambda_b S_b^o)^{1/\alpha}. \]  

(8)

The interpretation of the error measure as a distance measure in a perceptual space is straightforward. If both periodic-structure and blur artifacts are visible, the resulting loss in perceptual quality is proportional to the \( \alpha \) norm of a vector \((S_p^a, \sqrt{\lambda_b S_b^o})\) in a perceptual plane spanned by these artifacts.

4. SENSORY-STRENGTH FUNCTIONS

Explicit expressions for the sensory strengths of periodic structure and blur for a columnar sampling structure and Gaussian interpolation (see Fig. 1) were derived and experimentally verified in Refs. 9 and 11. Details of the derivation of the expressions can be found in Appendix A. Appendix B contains a method for finding the values of the parameters in the expressions. A columnar structure means that the image is sampled in the horizontal direction only. Gaussian interpolation can, for instance, be implemented by an optical filter placed in front of the display. The expressions for the perceptual strengths of blur and periodic structure are, respectively,

\[ S_b = 1 - \frac{1}{[(\sigma_e/\sigma_g)^2 + 1]^{0.25}}, \]  

(9)

\[ S_p = \frac{c [(1 + (m/m_0)^3)]^{1/3} - 1}{(m/m_0)^2}, \]  

(10)

where \( \sigma_e \) is the spread parameter of the Gaussian interpolation filter in the vertical direction, \( m \) is the modulation depth, \( \beta \approx 0.7 \), and \( c \) is a constant such that \( 0 \leq S_p \leq 1 \). The parameter \( m_0 \) is a periodic-structure-threshold parameter, and \( \sigma_g \) is the spread parameter of a Gaussian interpolation function representing the intrinsic blur of the early visual pathway and including the optics of the eye. These last two parameters depend on the observer and must be determined experimentally. The modulation depth \( m \) is equal to

\[ m = 2m_p \exp[-2(\pi d)^2(\sigma_b^2 + \sigma_e^2)], \]  

(11)

where \( d \) is the sampling distance and \( \sigma_b \) is the spread of the Gaussian interpolation filter in the horizontal direction. We use the modulation depth of the first harmonic of the columnar structure because the first harmonic is a good predictor of the visibility of any periodic structure of not too low frequency.\textsuperscript{5} The attenuation factor \( m_p \) that accounts for the interpolation of the columns of width \( w \) in Fig. 1 equals

\[ m_p = \frac{\sin(\pi w/d)}{(\pi w/d)}. \]  

(12)

This factor is the attenuation of the spectral component at the sampling frequency \((1/d)\) caused by a filter with a rectangular impulse response of width \( w \), i.e. the columns of the structure.

For the sensory strengths of the complement of brightness and the complement of brightness contrast, we use

\[ S_{B^c} = 1 - (\bar{L}/\bar{L}_{\text{max}})^{0.3}, \]  

(13)

\[ S_{C^c} = 1 - C^{0.3}, \]  

(14)

where \( \bar{L}_{\text{max}} \) is the maximum average luminance of a set of single-edged images that is used in an experiment (see Appendix A for further details).

The sensory strengths of periodic structure and blur are a function of the physical display parameters. The expressions for these strengths in Eqs. (9) and (10) do not depend on the physical parameters average luminance and Michelson contrast of the image. Such dependences have, however, been reported. Watt and Morgan,\textsuperscript{16} for instance, found just noticeable differences (jnd) of blur to be a power function of edge contrast with an exponent of \(-0.5\). Watanabe et al.\textsuperscript{17} and van Nes\textsuperscript{18} showed that contrast-detection thresholds for sine gratings are lower for higher average luminances. The influence of the
image-content parameters, average luminance, and Michelson contrast on the impairments of blur and periodic structure will be included in the relative-weight parameters $a_i$ of the perceptual-error measure: $a_p = a_p(C, \hat{L})$ and $a_b = a_b(C, \hat{L})$. Although it would be more appropriate to include the image-content parameters in the sensory-strength functions, the parameters will be included in the relative-weight parameters $a_i$ of the perceptual error measure, because we think that the effects are second-order effects. If experimental data indicate that the relative weight parameters depend heavily on average luminance or Michelson contrast, then this factor should be transferred from the relative-weight parameters in the sensory-strength functions.

5. EXPERIMENTAL VALIDATION

Although it is possible to obtain all data necessary to check the performance of the model as well as to study the values of the parameters and their dependences on image content and observers from one experiment, we opted for a stepwise approach. Instead of one there were two experiments. The first experiment addressed the fundamental problem of the value of the exponent in the combination rule for impairments. In the second experiment we determined the values of the intrinsic blur parameter, the periodic-structure threshold parameter, and the relative weight parameters and studied the dependence of these parameters on observers, average luminance, and Michelson contrast. In addition, we checked how well the model predictions fit the experimental data.

In both experiments, subjects were instructed to rate perceptual image quality of the stimuli. Since it is not trivial for subjects to be able to use the concept of perceptual image quality of a single-edged image, we experimentally verified whether subjects can employ the concept for simple images such as the single-edged image. In a small experiment, three subjects judged the perceptual quality of 24 black-and-white stimuli. The stimuli were sampled and interpolated versions of the portrait of a woman ("Wanda") and a simple ("Mondrian") image. The simple image consisted of seven partially overlapping rectangles of varying size. Only three different luminance levels were used for the seven rectangles. Before the experiment, subjects viewed a test series that included the extreme stimuli to familiarize themselves with the images and to adjust the sensitivity of their scale. Experimental results indicate that the trends for the data of Wanda and the Mondrian image are similar. There are no unexpected effects. We thus conclude that the subjects are able to use a stable perceptual-quality criterion for simple images. Since the Mondrian image already has little meaning from a cognitive point of view, we assume, in addition, that subjects can also use the concept of perceptual quality for other simpler images such as the single-edged image.

Scaled perceptual-quality values can be converted into impairments by use of Eq. (1). In the following, these impairment values are referred to as measured impairments.

6. MINKOWSKI EXPONENT EXPERIMENT

With the experiment described below we determine the value of the Minkowski exponent $\alpha$ in the combination rule for impairments as given in Eq. (3). Since it is sufficient to vary the impairments of two artifacts in order to be able to determine this exponent, we varied only the most dominant artifacts, periodic structure and blur. Therefore the luminances of the two half-planes of the single-edged image are constant for all stimuli. Equation (3) then reduces to

$$I^\alpha = I_p^\alpha + I_b^\alpha.$$  

A. Method

1. Image

Only one type of image, a single-edged image as shown in Fig. 1, was used in this experiment. This single-edged black-and-white image consists of two uniform regions with luminances $L_1$ and $L_2$ separated by a horizontal edge. The average luminance was $\bar{L} = (L_1 + L_2)/2 = 9.6$ cd/m$^2$, and the Michelson contrast was $C = L_1 - L_2)/(L_1 + L_2) = 0.21$.

2. Equipment

Single-edged black-and-white images were generated and displayed with use of a Gould DeAnza IP 8400 image-processing system. An image consisted of 512 × 512 eight-bit pixels. During the experiment only the center 496 × 496 pixels (size 0.28 m × 0.28 m or 4° by 4°) were shown on a Conrac model 2400 high-resolution 50-Hz interlace monochrome monitor. The viewing distance was 4 m and the pixel pitch was 0.49 arc min. The system was calibrated in such a way that there was a power-law relation $L \propto g^\gamma$, with an exponent $\gamma = 2.5$, between the eight-bit pixel value $g$ and the luminance $L$ of a pixel.

3. Stimuli

The stimulus set consists of different versions of the single-edged image that was defined above. Besides this original image the set contained single-edged images with a blurred edge varying in filtering of the edge and with a periodic structure varying in modulation depth of the structure. The experiment had a complete factorial design, i.e., all combinations of periodic structure and blur were presented. Besides a nonsampled version there were four sampled versions with varying modulation depth. A vertical periodic structure was introduced by imposing a columnar sampling structure with a sampling distance of 8 pixel pitch units (3.92 arc min) and a column width of 4 pixel pitch units (1.96 arc min). The modulation depth was varied by optical Gaussian filtering in the horizontal direction. Gaussian filtering was simulated on the Gould De Anza system by transforming pixel values into luminance values, filtering in the horizontal direction with a binomial impulse response filter of length $L$, and finally transforming the filtered luminance values into pixel values. The filter length is related to the spread parameter $\sigma$ of the Gaussian impulse response filter by $\sigma = 1/2 \sqrt{L - 1}$. Filter lengths were 14, 23, 33, and 43 pixel pitch units. In addition to the five stimuli specified above, the stimulus set contained four blurred ver-
sions for each of these five stimuli. Blur of the horizontal edge was varied by a similar filter in the vertical direction with filter lengths 10, 18, 33, and 60 pixel pitch units. The lengths were chosen such that for both periodic structure and blur the difference in perceived strength between successive filter lengths was approximately equal.

4. Procedure
Six male subjects between 27 and 43 years of age rated perceptual quality of the displayed images on a 10-point numerical category scale ranging from 1 to 10. Subjects had normal or corrected-to-normal vision and visual acuity, measured on a Landolt chart, between 1.25 and 2. Although two of the subjects (FB and GS) had a slight red–green deficiency, their results did not differ significantly from those of the other subjects. Subjects received an instruction form in which the quality of the single-edged image was defined as depending only on the period of structure in the uniform regions and the blur of the edge. Before the start of the actual experiment, subjects judged a test series of four stimuli containing the extreme stimuli in order to adjust the sensitivity of their scale. Each test stimulus was presented twice. All 25 stimuli were presented 4 times. The sequence of the images was random. Images were presented for 5 s and were followed by an adaptation field with a luminance of 15 cd/m² that lasted until subjects pressed a key but had a minimum duration of 2 s. The viewing conditions satisfied Comité Consultatif International Radio (CCIR) recommendation 500 (Ref. 20) except for the viewing distance, which was 4 m. All category data were transformed into an interval scale on the psychological continuum with use of Thurstone’s law of categorical judgment. We applied a class I model involving replications of trials within a single subject with condition D constraints.21 These constraints limit the number of model parameters by the assumption that the correlation between the momentary position of stimuli and the category boundaries as well as the dispersion of both category boundaries and stimuli are constant. Before the Thurstone correction, data were processed in accordance with Edwards’ method22 so that scale values of the extreme categories could be corrected. It was found that the trends in the Thurstone-corrected data of the individual subjects were similar. It was therefore permissible to average the Thurstone-corrected data over subjects.

B. Results
We use the perceptual impairments that can be derived from the scaled quality values to determine the value of the Minkowski exponent. Figure 3 displays the data averaged over subjects of the Minkowski exponent experiment and shows that a Minkowski exponent \( \alpha = 2 \) yields a fair fit. Unless otherwise stated, the length of an error bar in this and other figures is twice the standard error of the mean averaged over the data points. In the geometrical representation in Fig. 3, each data point is characterized by a Euclidian distance from the origin of the plot and a direction. The distance is equal to the total perceptual impairment. The direction is determined by the constituent impairments of the artifacts periodic structure and blur. In Fig. 3 the constituent impairment values are indicated by the horizontal and vertical lines. The positions of the lines are determined by an iteration process. Initially, the positions are equal to the data values on the axes. The intersections of the lines are the predicted impairment values. The measured impairments that have the same direction as the corresponding predictions are decomposed into a periodic-structure component along the horizontal axis and a blur component along the vertical axis. The new position of a line is the average of the five components along this line. In each iteration, new values are determined for all lines. Iteration stops when the positions of the lines do not vary significantly between variations.

Although a Minkowski exponent \( \alpha = 2 \) is satisfactory, it may not be optimal. We therefore determined the value of the exponent for which the model best fitted the data. To this end a measure of fit must be chosen. A measure of fit is the root mean square of the differences between the measured impairments and impairments calculated from constituent impairments with Eq. (15). This root mean square of differences is plotted in Fig. 4 as a function of the Minkowski exponent. There is a curve for each subject and for the data averaged over subjects. The symbols indicate the optimal values of the exponent for which the root mean square of differences is minimal. The curves in Fig. 4 are similar to those found by Ronacher and Bautz,23 who estimated the Minkowski exponent from multidimensional scaling data for the dissimilarity of pairs of disks that differed in both size and luminance. The curves in Fig. 4 yielding the two highest values of the exponent are very flat near the minima.

Ronacher and Bautz23 explained the flat curves from the fact that as the exponent increases, two Minkowski...
metrics of a constant difference in exponents become more and more similar, and therefore differences in the minima shown in Fig. 4 become less significant. Ronacher and Bautz show that the confidence intervals for the optimal values of the Minkowski exponent increase from left to right as shown in Fig. 4. Owing to the less pronounced minima, the estimates for the two highest values of the Minkowski exponent are less accurate.

The optimal values for the exponent of Fig. 4 are in good agreement with optimal values determined by using another measure of fit, namely, the mean of differences between the measured and the calculated impairments.

This mean of differences is plotted in Fig. 5 as a function of the Minkowski exponent. Again there is a curve for each subject and for the data averaged over subjects. The symbols indicate values for which the mean of differences is zero. The use of the mean of differences as a measure of fit is based on the fact that if a model fits the data well, there will be no systematic errors but only random normally distributed errors with an average value of zero. Consequently, we use the zero crossings for determining the values of the Minkowski exponent. Note that none of the curves in Fig. 5 is flat near the zero crossings. Small variations in the experimental data will therefore only slightly change the estimated value of the exponent. The standard deviation of the estimated values is smaller if the slope of the curve near the zero crossing is larger. Since the slopes of the curves in Fig. 5 are approximately equal near the zero crossing, the estimates are approximately equally accurate. In addition, the slopes in Fig. 5 are sharper than those of the corresponding curves in Fig. 4. Therefore we think that the values estimated from Fig. 5 are more reliable.

The values of the Minkowski exponent determined from Figs. 4 and 5 are listed in Table 1. Apart from the values for subjects HR and MB, which have flat curves near the minima in Fig. 4, the results from the root-mean-square-of-differences criterion are consistent with the mean-of-differences criterion.

From Table 1 and Figs. 3–5, we conclude that the Minkowski exponent should have a value \( \alpha > 1 \). Within experimental error, an exponent \( \alpha = 2 \) is acceptable for all subjects, excluding HR and MB. In the following we will use \( \alpha = 2 \), which is consistent with values found by de Ridder. This value has the additional advantage that it allows for an easy geometrical interpretation of the impairments in a Euclidean space. In Section 8 (the discussion) we will indicate that \( \alpha = 2 \) can also be used for the data of subjects HR and MB.

### Table 1. Minkowski Exponent Values Determined with a Root-Mean-Square-of-Differences Criterion (Fig. 4) and a Mean-of-Differences (Fig. 5) for Data of Individual Subjects and for Data Averaged over Subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Root Mean Square of Differences</th>
<th>Mean of Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>GS</td>
<td>2.4</td>
<td>2.9</td>
</tr>
<tr>
<td>HR</td>
<td>6.3</td>
<td>5.2</td>
</tr>
<tr>
<td>MB</td>
<td>8.6</td>
<td>3.6</td>
</tr>
<tr>
<td>MN</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>RS</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>Average</td>
<td>2.3</td>
<td>2.3</td>
</tr>
</tbody>
</table>

7. CONTRAST AND LUMINANCE EXPERIMENT

In the second experiment the perceptual error measure based on Eqs. (1), (5), (9), (10), (13), and (14) with parameters \( \sigma_0, m_0, a_p, a_k, a_{C^*}, a_{B^*} \), and \( \alpha \) is fitted to the experimental data for six subjects and the “average” sub-
ject, with use of the $m$, $\sigma$, $\bar{L}$, and $C$ values corresponding to the stimuli. For the parameter $\alpha$ we use the value $\alpha = 2$ obtained from the Minkowski exponent experiment. The parameters $\sigma_0$, $m_0$, $a_C$, $a_G$, and $\lambda_b = a_g/a_m$ are unique for each subject, and $\lambda_b$ is also unique for each combination of $\bar{L}$ and $C$. The experimental data were obtained in an experiment in which subjects judged perceptual quality of single-edged images that were varied in sampling distance, modulation depth, filtering of the edge, average luminance $\bar{L}$, and Michelson contrast $C$.

A. Method

1. Image

Only one type of image, a single-edged black-and-white image, was used in this experiment. This single-edged black-and-white image consists of two uniform regions with luminances $L_1$ and $L_2$ that are separated by a horizontal edge. We used eight versions of this single-edged image, each having a different combination of values for average luminance and Michelson contrast. The average luminance of the images $\bar{L} = (L_1 + L_2)/2$ was 2.1 cd/m$^2$, $\bar{L} = 9.7$ cd/m$^2$, and $\bar{L} = 43.9$ cd/m$^2$. The Michelson contrast $C = (L_1 - L_2)/(L_1 + L_2)$ was 0.10 and $C = 0.21$. For the images with $\bar{L} = 9.7$ cd/m$^2$ there were two additional contrast levels, $C = 0.43$ and $C = 0.90$.

2. Equipment

The stimuli in this experiment were generated and displayed by use of a Gould DeAnza IP 8400 image-processing system. An image consisted of $512 \times 512$ eight-bit pixels. During the experiment only the center 496 $\times$ 496 pixels (size $0.36$ m $\times$ 0.29 m or $9.7^\circ \times 7.9^\circ$) were shown on a Conrac model 2400 high-resolution 50-Hz interlace monochrome monitor. Although the monitor was of the same type as in the Minkowski exponent experiment, it was a different apparatus. The viewing distance was 2.1 m. The pixel pitch was 1.18 arc min in the horizontal direction and 0.94 arc min in the vertical direction. The system was calibrated in such a way that there was a power-law relation $L \propto g^\gamma$, with an exponent $\gamma = 2.5$, between the eight-bit pixel value $g$ and the luminance $L$ of a pixel.

3. Stimuli

The stimulus set consists of different versions of the single-edged images defined above. For each of these eight “original” (nonsampled and nonfiltered) images with a specific $LC$ combination, the set contained this original image as well as single-edged images with a blurred edge varying in filtering of the edge and with a periodic structure varying in sampling distance and modulation depth of the structure (see Fig. 1). Figure 6 shows some of the stimuli. A vertical periodic structure was introduced by imposing a columnar sampling structure with a sampling distance of 2 and 4 horizontal pixel pitch units (2.36 and 4.72 arc min). In both cases the width of the columns was half the sampling distance. The modulation depth was varied by optical Gaussian filtering in the horizontal direction. Gaussian filtering was simulated on the Gould DeAnza system by transforming pixel values into luminance values, filtering in the horizontal direction with a binomial impulse response filter of length $l$, and then transforming the filtered luminance values into pixel values. The filter length of the binomial filter is related to the spread parameter $\sigma$ of the Gaussian filter by $\sigma = 1/2\sqrt{d - 1}$. Filter lengths were 2, 3, 7, and 15 pixel pitch units. In addition to the stimuli specified above, the stimulus set contained blurred versions of each of these stimuli. Blur of the edge was varied by an identical filter in the vertical direction with filter lengths 5, 11, 25, and 60 pixel pitch units. In order to reduce the number of stimuli, we conducted the experiment without a complete factorial design. For each of the eight luminance contrast combinations there were 27 different combinations of sampling distance, modulation depth, and edge blur as indicated in Table 2.

In practice, the luminance profiles of the sampled images are not exactly rectangular as would be expected for a columnar sampling structure. The rectangular corners of the profile were rounded. Consequently, the modulation depths of the sampling structures are somewhat lower than the theoretical modulation depths. This effect is included in the factor that accounts for the width of the columns relative to the sampling distance. According to Eq. (12) this factor is, ideally, $\sin(\omega/d) = \sin(1/2) \approx 0.637$ for ideal rectangular profiles with a width of half the sampling distance. We measured some of the displayed luminance profiles and calculated this factor. The value that we found and that is used in calculations throughout this paper was 0.50. The factor turned out to be relatively independent of the frequency of the structure and the peak luminance of the profile.

Fig. 6. Four of the 216 stimuli from the contrast-and-luminance-effect experiment. (a) $\bar{L} = 2.1$ cd/m$^2$, $C = 0.21$, zero modulation depth, and no filtering of the edge (V01); (b) $\bar{L} = 43.9$ cd/m$^2$, $C = 0.21$, zero modulation depth, and maximum filtering of the edge (V60); (c) $\bar{L} = 9.7$ cd/m$^2$, $C = 0.10$, maximum modulation depth of the coarsest sampling structure ($d = 4.72$ arc min), and maximum filtering of the edge; (d) $\bar{L} = 9.7$ cd/m$^2$, $C = 0.10$, maximum modulation depth of the finest sampling structure ($d = 2.36$ arc min), and maximum filtering of the edge.
Table 2. 27 Sampling and Interpolation Combinations Used for Each of the Eight Combinations of Average Luminance and Michelson Contrast in the Contrast-and-Luminance Experimenta

| V60 | ○△ | ○△ | □ |
| V25 | ○△ | □ |
| V11 | ○△ | ○△ | □ |
| V05 | ○△ | □ |
| V01 | ○△ | △ | ○△ | □ |
| H01 | H02 | H03 | H07 | H15 | H∞ |

aThe numbers indicate the filter lengths of the binomial impulse response filters in the horizontal (H) and vertical (V) directions in horizontal and vertical pixel units. If the filter length is unity, then the image is not filtered. The infinite filter length in the horizontal direction, □, corresponds to the nonsampled image, which can alternatively be seen as a periodic structure with zero modulation depth. ○ and △ indicate stimuli with a sampling distance of 2 and 4 horizontal pixel pitch units, respectively.

4. Procedure

In four sessions, six male subjects between 27 and 43 years of age rated perceptual quality of the displayed images on a 10-point numerical category scale ranging from 1 to 10. Subjects had normal or corrected-to-normal vision and visual acuity between 1.25 and 2, measured on a Landolt chart. Although two of the subjects (FB and GS) had a slight red–green deficiency, their results did not differ significantly from those of other subjects. Subjects received an instruction form in which the quality of the single-edged image was defined as depending only on the periodic structure in the uniform regions and the blur of the edge. Before the start of the actual experiment, subjects judged a test series of 16 stimuli containing the extreme stimuli in order to adjust the sensitivity of their scale. The 216 stimuli were presented once in each session. The sequence of the stimuli was random. Images were presented for 5 s and were followed by an adaptation field with a luminance of 15 cd/m² that lasted until subjects pressed a key but had a minimum duration of 2 s. The viewing conditions satisfied CCIR recommendation 500. The viewing distance was 2.1 m. All category data were transformed into an interval scale on the psychological continuum with use of Thurstone’s law of categorical judgment. We applied a class I model involving replications over trials within a single subject with condition D constraints. These constraints limit the number of model parameters by the assumption that the correlation between the momentary position of stimuli and the category boundaries as well as the dispersion of both category boundaries and stimuli are constant. Before the Thurstone correction, data were processed in accordance with Edwards’ method so that scale values of the extreme categories could be corrected. It was found that the trends in the Thurstone-corrected data of the individual subjects were similar. It was therefore permissible to average the Thurstone-corrected data over subjects.

B. Parameters of the Sensory-Strength Functions

Below we will determine the values of the intrinsic blur of the early visual pathway parameter \( \sigma_0 \) and the periodic-structure-threshold parameter \( m_0 \) in the sensory strength functions for periodic structure \( S_p \) and blur \( S_b \), as given in Eqs. (9) and (10) by using a method described in Appendix B. In addition, we will illustrate that these functions are adequate for predicting the experimental data.

We will first determine the \( \sigma_0 \) parameter of the blur function by using sets of stimuli that vary only in the degree of filtering and thus only in the degree of blur of the edge. These stimuli are in the same column of Table 2 and have the same symbol. We used the set with the nonsampled stimuli shown in the right-hand column of Table 2. Since there are eight luminance–contrast combinations, there are eight such sets.

For all stimuli within a set, luminance, Michelson contrast, sampling distance, and modulation depth are identical. Hence Eq. (5) can be rewritten as

\[
I^2 = a_0^2 S_b^2 + (a_p^2 S_p^2 + a_c^2 S_c^2 + a_v^2 S_v^2),
\]

where the second term on the right-hand side is constant. Consequently, \( S_b^2 \) and \( I^2 \) are related linearly. Since the stimuli are completely specified, we can make a graph of the calculated strength of blur [Eq. (9)] versus the experimental blur values. The value of the intrinsic blur parameter \( \sigma_0 \) is taken such that the points on this graph are closest to a straight line. More specifically, the value of the parameter \( \sigma_0 \) is taken such that the coefficient of determination of the regression line of \( I_m^2 \) on \( S_b^2 \) is closest to unity. Here \( S_b \) is the calculated strength and \( I_m \) stands for the measured impairment, which, according to Eq. (1), can be calculated from the quality values measured in the experiment. The coefficient of determination measures how well a straight line fits the data. It is equal to the ratio of the explained variation to the total variation (e.g., Chatfield). The total variation is often called the total corrected sum of squares, and the explained variation is equal to the difference between the total variation and the residual sum of squares. See Appendix B for further details of the fitting procedure and the calculation of the variance.

Since there are eight sets, we obtain eight estimates for \( \sigma_0 \) for each subject. These \( \sigma_0 \) values and the corresponding variances were averaged over the sets. Table 3 shows the \( \sigma_0 \) values and the standard deviations \( s_{\sigma_0} \). Figure 7 shows the corresponding curves of \( S_b^2 \) versus \( I_m^2 \) for the data averaged over subjects with the \( \sigma_0 \) value from Table 3. The fact that the points in Fig. 7 are on a straight line within the experimental error as indicated by the error bars implies that the sensory-strength function for blur is adequate.

The impulse response of a Gaussian filter with the intrinsic blur values of Table 3 as spread parameter is in good agreement with the point-spread function determined by Blommaert et al. Moreover, the intrinsic blur values are consistent with the visual acuity of the subjects (1.25 for subject SP and 2.0 for the other subjects). The variation of the \( \sigma_0 \) values over subjects is not unusual for psychophysical experiments.

For determining the \( \sigma_0 \) and \( m_0 \) parameters of the periodic-structure function [Eq. (10)] we use sets of stimuli that vary only in sampling distance and modula-
tion depth and thus in the visibility of periodic structure. Such stimuli are in the same row of Table 2. We used the set with the stimuli with the nonfiltered edge, which are at the bottom row of Table 2. For each luminance–contrast combination there is a separate set. Since the luminance, Michelson contrast, and filtering of the edge of the stimuli in such a set are constant, Eq. (5) can again be rewritten in a form where the second term on the right-hand side is constant:

$$I^2 = a_p^2 s^2_p + (a_b^2 s^2_b + a_c^2 s^2_c + a_r^2 s^2_r)(17)$$

In this case $s^2_p$ and $I^2$ are related linearly. Since the stimuli are completely specified, we make graphs of the calculated strength of periodic structure [Eq. (10)] versus the experimentally measured periodic-structure values, using a procedure similar to the one discussed above. The $m_0$ and $\sigma_0$ values are taken such that the points in this graph are closest to a straight line. Specific details of the procedure can be found in Appendix B. The parameter values and the corresponding variances were averaged over the sets for different luminance–contrast combinations. The values for each subject and for the data averaged over subjects are listed in Table 4. Curves of $s^2_p$ versus $I^2$ for the data averaged over subjects are drawn in Fig. 8. Since the lines in this figure are straight within experimental error, we conclude that the sensory-strength function for periodic structure describes the data well.

The $m_0$ values in Table 4 are consistent with Wilson’s results. The $\sigma_0$ values yield Gaussian impulse response filters that are consistent with point-spread functions determined by Blommaert et al. Although the intrinsic blur parameters were determined as parameters of two different sensory-strength functions for blur (Table 3) and for periodic structure (Table 4), the values are similar within experimental error, except perhaps for subject MB.

### Table 3. Values for the Intrinsic Blur Parameter $\sigma_0$ and the Standard Deviation $\sigma_{s_0}$ in the Perceptual-Strength-of-Blur Function Determined in the Contrast-and-Luminance Experiment for Data from the Individual Subjects and for Data Averaged over Subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>$\sigma_0$ (arc min)</th>
<th>$\sigma_{s_0}$ (arc min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB</td>
<td>1.09</td>
<td>0.47</td>
</tr>
<tr>
<td>GS</td>
<td>0.50</td>
<td>0.18</td>
</tr>
<tr>
<td>HR</td>
<td>0.70</td>
<td>0.23</td>
</tr>
<tr>
<td>MB</td>
<td>0.36</td>
<td>0.14</td>
</tr>
<tr>
<td>RS</td>
<td>0.60</td>
<td>0.22</td>
</tr>
<tr>
<td>SP</td>
<td>0.50</td>
<td>0.27</td>
</tr>
<tr>
<td>Average</td>
<td>0.59</td>
<td>0.14</td>
</tr>
</tbody>
</table>

### Table 4. Values for the Intrinsic Blur Parameter $\sigma_0$, the Periodic-structure-threshold Parameter $m_0$, and the Standard Deviations $\sigma_{s_0}$ and $\sigma_{m_0}$ in the Perceptual-Strength-of-Periodic-Structure Function Determined in the Contrast-and-Luminance Experiment for Data from the Individual Subjects and for Data Averaged over Subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>$m_0$</th>
<th>$\sigma_{m_0}$</th>
<th>$\sigma_0$ (arc min)</th>
<th>$\sigma_{s_0}$ (arc min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB</td>
<td>0.005</td>
<td>0.008</td>
<td>0.91</td>
<td>0.16</td>
</tr>
<tr>
<td>GS</td>
<td>0.209</td>
<td>0.658</td>
<td>0.42</td>
<td>0.25</td>
</tr>
<tr>
<td>HR</td>
<td>0.024</td>
<td>0.014</td>
<td>1.00</td>
<td>0.10</td>
</tr>
<tr>
<td>MB</td>
<td>0.028</td>
<td>0.018</td>
<td>0.79</td>
<td>0.10</td>
</tr>
<tr>
<td>RS</td>
<td>0.044</td>
<td>0.030</td>
<td>0.56</td>
<td>0.13</td>
</tr>
<tr>
<td>SP</td>
<td>0.010</td>
<td>0.022</td>
<td>0.76</td>
<td>0.56</td>
</tr>
<tr>
<td>Average</td>
<td>0.013</td>
<td>0.007</td>
<td>0.81</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Fig. 7. Squared measured perceptual impairment versus the squared perceptual strength of blur for the data of the contrast-and-luminance experiment averaged over subjects ($\sigma_0 = 0.59$ arc min). The horizontal error bar is twice the standard deviation in $s^2_p$ calculated from $s_{s_0}$ with use of standard error-propagation methods. For each increase in $L$ or $C$ the curves are shifted vertically by 0.2 unit. The bottom curve has not been shifted. The average luminance $L$ is given in candelas per square meter.

Fig. 8. Squared measured perceptual impairment versus the squared perceptual strength of periodic structure for the data of the contrast-and-luminance experiment averaged over subjects ($\sigma_0 = 0.81$ arc min and $m_0 = 0.013$). The horizontal error bar is twice the standard deviation in $s^2_p$ calculated from $s_{s_0}$, $s_{m_0}$, and their covariance with use of standard error-propagation methods. For each increase in $L$ or $C$ the curves are shifted vertically by 0.2 unit. The bottom curve has not been shifted. The average luminance $L$ is given in candelas per square meter.
In the following we use the average of these $\sigma$ values in the sensory-strength functions for blur and periodic structure.

C. Perceptual Strength of Complement of Brightness and Complement of Brightness Contrast

In this subsection we check how well the sensory-strength functions for the artifacts complement of brightness and complement of brightness contrast can be used to predict the total perceptual impairment that is due to variations in average luminance and Michelson contrast. To this end we use the eight stimuli that differ in luminance and contrast but are all nonsampled and nonfiltered in the vertical direction. These stimuli occur in the lower right-hand box of Table 2. Since the sensory strengths of both blur and periodic structure are zero for these stimuli, Eq. (5) for the total perceptual impairment reduces to

$$I^2 = a_{c*}^2 S_{c*}^2 + a_{b*}^2 S_{b*}^2.$$  \hfill (18)

<table>
<thead>
<tr>
<th>Subject</th>
<th>$a_{c*}^2$</th>
<th>$a_{b*}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB</td>
<td>0.16</td>
<td>-0.01</td>
</tr>
<tr>
<td>GS</td>
<td>0.65</td>
<td>-0.08</td>
</tr>
<tr>
<td>HR</td>
<td>0.26</td>
<td>-0.07</td>
</tr>
<tr>
<td>MB</td>
<td>0.31</td>
<td>-0.12</td>
</tr>
<tr>
<td>RS</td>
<td>0.33</td>
<td>0.06</td>
</tr>
<tr>
<td>SP</td>
<td>0.78</td>
<td>0.09</td>
</tr>
<tr>
<td>Average</td>
<td>0.39</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

Table 5. Values of the Squared Relative Weight Parameters $a_{c*}^2$ and $a_{b*}^2$ in the Contrast-and-Luminance Experiment for Each Subject and for the Data Averaged over Subjects

The sensory strengths of complement of brightness contrast and complement of brightness can be calculated directly, since the functions contain no parameters and the stimuli are completely specified. Since the quality and thus the impairments of the stimuli are experimentally measured, we can solve the relative weight parameters $a_{c*}$ and $a_{b*}$ from the set of eight linear equations in $a_{c*}^2$ and $a_{b*}^2$. The results are listed in Table 5. Figure 9 shows a comparison of the measured and the predicted perceptual impairments of the eight stimuli for the data averaged over subjects.

Table 5 indicates that the squares of the relative weights of complement of brightness ($a_{b*}^2$) are relatively small compared with those of complement of brightness contrast ($a_{c*}^2$) for all subjects and the data averaged over subjects. The negative values for parameter $a_{b*}^2$ are most likely caused by statistical fluctuation, although perhaps for subject MB the $a_{b*}^2$ parameter actually has a negative value, meaning that the impairment increases instead of decreases with increasing average luminance. Since the $a_{b*}^2$ values in Table 5 are relatively small compared with the $a_{c*}^2$ values, we conclude that the influence of average luminance is small. This is in accordance with the data in Fig. 9, where the impairment varies only little within the two groups of stimuli with the same symbol. Figure 9 further shows that the impairment caused by variation of the average luminance and of the Michelson contrast is relatively small yet is too large to be called second order. Note that only a part of the total range of the squared impairment ($0 \leq I^2 \leq 1$) is plotted in Fig. 9. The sensory-strength functions for the complement of brightness in Eq. (13) and the complement of brightness contrast in Eq. (14) are considered satisfactory, since the difference in Fig. 9 between calculated and measured impairments is within the experimental error.

D. Regularization Parameter

Below we determine unique values for the regularization parameter for blur $\lambda_3 = a_{c*}^2/a_{b*}^2$ for each combination of subject, luminance, and Michelson contrast. We start by rewriting Eq. (5), using $\alpha = 2$ from the Minkowski exponent experiment:

$$I^2 - a_{c*}^2 S_{c*}^2 - a_{b*}^2 S_{b*}^2 = a_{c*}^2 S_{c*}^2 + a_{b*}^2 S_{b*}^2.$$  \hfill (19)

Except for $a_{c*}$ and $a_{b*}$, all factors and terms in Eq. (19) are known. The impairments are calculated from the experimentally determined quality data. The relative weights $a_{c*}$ and $a_{b*}$ are taken from Table 5. All four sensory strengths can be calculated with Eqs. (9)–(14) since the stimuli are completely specified and the $\sigma$ and $m_p$ values are known.

For each subject there were 27 stimuli with the same luminance and contrast values (see Table 2). For each luminance–contrast combination we thus have a set of 27 simultaneous equations in $a_{c*}^2$ and $a_{b*}^2$. We solved these sets for the relative weights and calculated the regularization parameter as the quotient of these squared weights. Details on the calculation of the standard deviations of the regularization and weight parameters can be found in Appendix C.

The parameter data are represented in three ways. Figure 10 shows the interaction between the relative...
weight parameters and the regularization parameter as a function of average luminance and Michelson contrast for the data averaged over subjects. Figure 11 plots the regularization parameter for blur as a function of average luminance and Michelson contrast for each luminance–contrast combination for the data averaged over subjects. The average luminance $L$ is given in candelas per square meter.

Fig. 10. Values of the squared relative weights for periodic structure ($a_p^2$) and blur ($a_b^2$) and the regularization parameter for blur ($\lambda_b = a_b^2/a_p^2$) of the contrast-and-luminance experiment for each luminance–contrast combination for the data averaged over subjects. The average luminance $L$ is given in candelas per square meter.

Fig. 11. Regularization parameters for blur in the contrast-and-luminance experiment as a function of the luminance–contrast combination with the observer as a parameter. $\triangle$, FB; $\bigcirc$, GS; $\times$, HR; $+$, MB; $\Delta$, RS; $\bigcirc$, SP; $\square$, average. The average luminance $L$ is given in candelas per square meter. Dashed curve, data averaged over subjects. The error bar is twice the standard deviation in $\lambda_b$ for the data averaged over subjects.

Fig. 12. Regularization parameters for blur in the contrast-and-luminance experiment as a function of observer with the luminance–contrast combination as a parameter. $\square$, $C = 0.10$, $L = 2.1$ cd/m$^2$; $\bigcirc$, $C = 0.21$, $L = 2.1$ cd/m$^2$; $\triangle$, $C = 0.10$, $L = 9.7$ cd/m$^2$; $+$, $C = 0.21$, $L = 9.7$ cd/m$^2$; $\times$, $C = 0.43$, $L = 9.7$ cd/m$^2$; $\bigcirc$, $C = 0.90$, $L = 9.7$ cd/m$^2$; $\bigtriangledown$, $C = 0.10$, $L = 43.9$ cd/m$^2$; $\bigcirc$, $C = 0.21$, $L = 43.9$ cd/m$^2$. The error bar is twice the standard deviation in $\lambda_b$ for the data averaged over subjects.

We conclude that the regularization parameter depends on the observer. If the perceptual-error measure is used to determine optimal display parameters, we must study the effect of changes in the regularization parameter on the optimal display parameters.

Despite the relatively large standard error of the mean, Figs. 10 and 11 suggest that there is a systematic tendency for the regularization parameter for blur to increase if the Michelson contrast increases. The values of the regularization parameter for blur corresponding to the extreme Michelson contrasts are approximately 2 standard errors of the mean apart. To a lesser extent, this tendency may hold also for an increase in average luminance.

According to Fig. 10, the increase of the regularization parameter as a function of Michelson contrast is caused by both an increase of the weight parameter for blur and a decrease of the weight parameter for periodic structure. The increase of the relative weight parameter for blur with Michelson contrast is consistent with the finding, mentioned earlier, that jnd’s of blur are a power function of edge contrast with an exponent of $-0.5$: $\Delta \sigma \propto 1/\sqrt{C}$. This can easily be shown to imply that the relative weight of blur should be proportional to the square root of the Michelson contrast: $a_b \propto 1/\sqrt{C}$. In Fig. 13 we have replotted the square of the relative weight of blur for an average luminance of 9.7 cd/m$^2$ from Fig. 10 as a function of Michelson contrast. There is a systematic tendency for the points to be close to a straight line, indicating that $a_b^2 \propto C$ and hence $a_b \propto 1/\sqrt{C}$. The
E. Performance of the Perceptual-Error Measure

In this subsection we compare the experimentally measured impairment values \( \left( I_m \right) \) with the impairments \( \left( I_c \right) \) calculated with the perceptual-error measure \( \left[ \text{Eqs. (1), (5), (9), (10), (13), and (14)} \right] \). We use the parameter values specified in the previous sections. The Minkowski exponent is \( \alpha = 2 \). The \( \sigma_0 \) values are the average of the corresponding values in Tables 3 and 4. All parameters \( \left( \sigma_0, m_0, a_p, a_B, a_C^*, a_B^* \right) \) are unique for each subject, and \( \lambda_b = a_B^2/a_p^2 \) is also unique for each luminance–contrast combination. Note that for each subject we used all 216 measured data values to estimate the 20 parameters of the perceptual error measure, and we now use these parameters to calculate the impairments of all 216 stimuli.

In Fig. 14 we directly compare measured and calculated impairments for all stimuli of the data averaged over subjects. Figure 15 contains a similar plot but now for the data of a specific subject. From these figures we conclude that the predictions of the perceptual-error measure correlate quite well with scaling data.

The comparison in Figs. 14 and 15 is an overall one that does not give any insight into whether the deviations are random or whether there is some systematic error in the calculated values. Figures 16 and 17 therefore give some comparisons of calculated and measured impairments as a function of the sensory strengths. In these figures the squared values of impairments and squared sensory strengths are along the axes, and only one artifact is varied at a time so that the relations should be linear. Again we conclude that the perceptual-error measure performs well. Since the average value for the intrinsic blur parameter was used for the curves in Figs. 16 and 17, the maximum values of the sensory strengths differ slightly from those in Figs. 7 and 8. The slopes in
Figs. 16 and 17 show little variation, as would be expected since the regularization parameter for blur is only slightly influenced by variations in average luminance and Michelson contrast.

8. DISCUSSION

In this paper we have presented a model that satisfactorily predicts the perceptual image quality of columnar sampled and Gaussian interpolated single-edged black-and-white images. The model uses a Minkowski exponent $a = 2$. According to Figs. 4 and 5 this value was acceptable for all subjects except HR and MB. By reinterpreting the data of the Minkowski exponent experiment and using the results from the contrast-and-luminance experiment, we show that $a = 2$ is also acceptable for subjects HR and MB. In the Minkowski exponent experiment we used a relatively low sampling frequency. Consequently, the columns of the structure were clearly visible. On either side of the columns there is a vertical edge. Hence vertical edges in addition to the horizontal edge were visible. We think that subjects HR and MB, who are considered to be sharpness experts, also used the blur level of the vertical edges for determining the quality of the stimuli. We assumed that the sharpest edge(s) determine(s) the blur in the sampled and interpolated single-edged images, and we corrected the measured impairments in the Minkowski exponent experiment accordingly. This was possible because the filter lengths and the sampling distance of the stimuli as well as $m_0$, and $s_0$, and $l_b$ were known. When we redetermine the Minkowski exponents using the corrected data set, we find values very close to $a = 2$ for both HR and MB.

The model in this paper was tested only for a particular image, namely, the single-edged image. In other experiments—the complex black-and-white experiment and the complex color experiment—we used several complex color as well as black-and-white images. These experiments have an experimental setup similar to that of the contrast-and-luminance experiment. However, in complex images it is not possible to control the artifacts blur and periodic structure independently. Figure 18 gives an example of a columnar sampled and Gaussian interpolated image used in the complex black-and-white experiment. The sampled and interpolated images in the complex-black-and-white experiment are derived from nonsampled and nonfiltered so-called original images. These originals consist of $512 \times 512$ pixels and depict scenes such as the woman’s portrait “Wanda”, a terrace with a yellow parasol, pseudotext images, Mondrian-like images, and the simple single-edged image. A coarser columnar sampling structure with a sampling distance of 4, 6, and 8 pixels was imposed on these original images. The column width was half the sampling distance. The height of an element in the columns was one pixel. Pixels outside the columns were made black. All pixels in one element have the same luminance, which is equal to the average luminance of the pixels in the original image corresponding to that element. After sampling, the images were filtered with a filter with a binomial impulse response, which is the discrete equivalent of a Gaussian filter. Interpolation is the combined effect of the spatial extent of the columns and filtering, which simulates an optical filter placed in front of the display.

Figure 19 shows both the experimental data and the predictions of the model for the complex black-and-white woman’s portrait Wanda. In this figure we have plotted 1 minus the impairment, i.e., quality versus the spread $\sigma$ of the optical Gaussian interpolation filter. The plots in Fig. 19 are similar to the plots for the other complex images, text images, single-edged images, and Mondrian images in both black and white and color. Figure 19 clearly illustrates the effect of the columnar sampling structure and the optical Gaussian filter on perceptual image quality. For the nonfiltered stimuli ($\sigma = 0$), image quality drops if the sampling distance increases. The effect of filtering on the sampled images is twofold. First, the perceptual quality of the sampled images increases as...
Fig. 19. Comparison of measured (□) and calculated (×) perceptual quality. The perceptual quality averaged over subjects is plotted versus the spread parameter $\sigma$ of the optical Gaussian interpolation filter for the Wanda image of the complex black-and-white experiment. From top to bottom the four pairs of curves are for the nonsampled images and the sampled images with $d = 1.81$ arc min, $d = 2.72$ arc min, and $d = 3.63$ arc min.

the spread parameter increases until it reaches some maximum. Beyond this point, quality decreases as the filtering becomes stronger. The first part of the curves corresponds to a decrease in the perceptual strength of the periodic structure while the sensory strength of blur remains relatively small. In the second part of the curves the perceptual strength of blur increases while the perceptual strength of periodic structure hardly decreases. The value of the spread parameter for which the quality is maximal increases monotonically with the sampling distance. For nonsampled images the quality is highest if the image is not filtered.

The predicted impairment and hence image-quality values were calculated with Eq. (6), which has terms only for the artifacts’ periodic structure and blur. In Ref. 11 we argue how the perceptual-error measure for single-edged images presented in this paper can be generalized into an error measure for complex images that considers only the artifacts’ periodic structure and blur. The basic idea is that a complex image can be seen as a collection of straight edges with various contrasts, average luminances, and orientations. Equation (5) applies to all these local areas. The total impairment of the entire complex image is considered to be a weighted sum of the local impairments for the single-edged images. The weighting coefficients may depend on the position of the local area in the image, the conspicuity of the local area, or the observer’s attention. Note that a cognitive factor such as the recognition of a contour in a complex image may, for instance, yield higher weighting coefficients for the local areas containing the edges that make up this contour. An experiment involving independent blurring of horizontal and vertical edges in an image indicates that a set of weighting coefficients that are all equal or almost equal is more likely than an alternative extreme set in which all coefficients but one are zero. Consequently, the terms for the impairments of complement of brightness and complement of brightness contrast average out into a sampling- and interpolation-independent impairment that expresses the lack of average luminance and average contrast in the image. This impairment has a fixed value for each particular scene. The squares of the relative weight parameters for periodic structure and blur are the weighted sums of the corresponding squares of the relative weights of local areas.

The error measure for complex images does not contain terms for other artifacts such as the staircase effect. In an experiment,11 we showed that the staircase depends closely on orientation. When an edge in an image coincides almost but not completely with the vertical direction of the columns of a columnar sampling structure, the staircase impairment is of the same order of magnitude as the periodic-structure impairment. The staircase impairment decreases as the orientation of the edge is more perpendicular to the columns. The staircase impairment will therefore be large only for a relatively small number of local areas for which the direction of the edge is almost the same as the direction of the periodic structure. Since all local areas have equal weights, the average contribution of the staircase impairment will be small compared with that of the periodic structure and blur impairments, which are present in all local areas.

The error measure for complex images contains one regularization parameter for blur that equals the ratio of weighted sums of the relative weights of periodic structure and blur for the local areas. Since this $\lambda_b$ parameter varies for different $LC$ combinations, as indicated by Fig. 10, it seems sensible to test how well the measured impairments of the contrast-and-luminance experiment can be predicted by using only one $\lambda_b$ value. Figure 20 shows that the model still adequately predicts the experimental data if only one value $\lambda = 1.5$ is used.

The perceptual-error measure for complex images is a specific example from a class of general image-quality models that predict perceptual image quality from physical parameters. An important feature of these models is a perceptual space intermediate between the physical parameter space and perceptual image qual-

![Fig. 20. Squared measured perceptual impairments versus squared calculated perceptual impairments of the contrast-and-luminance experiment for the data averaged over subjects. Contrary to Fig. 14, only one single $\lambda_b$ value is used for all $LC$ combinations. The dashed line indicates points with $I_c = I_m$.](image-url)
ity. Such a concept is also present in Engeldrum’s model. The sensory strengths of the perceptual attributes underlying perceptual image quality are along the axis of this space. In our case we chose periodic structure, blur, brightness, and brightness contrast as the relevant dimensions, because these attributes were dominantly visible when we looked at sampled and interpolated single-edged images. This choice is supported by the predictions of the model.

The sensory-strength functions model the front end of the visual system. Bottom-up processing of images in this part of the visual system is relatively independent of image content and is basically the same for different observers. We therefore expect small variations in the perceptual parameters \((\sigma_0, m_0)\) of the strength functions. The Minkowski metric for combining the strengths of the attributes underlying perceptual quality reflects properties of high-level vision. Top-down processing in this part enables observers to express their personal preferences, which may depend on image content. Hence we expect larger variations in the perceptual parameter \((\lambda_b)\) in the combination rule.

We would like to stress that the variations in the regularization parameter are not a shortcoming of the model but represent real variations that exist among observers. Apart from variations in the low-level parameters \(\sigma_0\) and \(m_0\) and in two parameters \(a_{B_s}\) and \(a_{C_s}\) that account for quality variations of the nonsampled and nonfiltered single-edged images owing to contrast and brightness variations, the current model uses only a single regularization parameter to account for individual differences. The variations put a certain constraint on the optimization of image-reproduction techniques, namely, that optimal solutions should also possess a certain robustness against variations in \(\lambda_b\).

On the basis of the results with sampling and interpolation, we think that the structure of the general image-quality model is an adequate reflection of the perceptual process that leads to image quality. Hence we expect it also to work for other image-processing problems. In order to be of practical value, the relations between the physical parameters and the sensorial strengths must be determined either empirically or analytically as in the case of sampling and interpolation. In addition, the physical parameter values must be known. In the case of sampling and interpolation, the sampling distance and the spread of the Gaussian filter are known. If physical parameter values are not available, these parameters must be extracted from the images. Kayargadde gives an example how the spread of a Gaussian filter can be determined from an arbitrary image.

**APPENDIX A: INFERENCIE OF SENSORY-STRENGTH FUNCTIONS**

In what follows we will show how the strategy introduced by Fechner and recently advocated by Watt can be used to derive quantitative expressions for the sensory strengths of the artifacts as a function of the physical parameters. We start with the derivation of Eq. (9) for the sensory-strength function for blur.

We assume that the perceptual strength of blur is a differentiable function of the physical spread parameter \(\sigma\) of the Gaussian filter:

\[
S_b = S_b(\sigma).
\]  

Variations in the perceptual variable \(S_b\) are then related to variations in the physical variable \(\sigma\) by

\[
\Delta S_b = \frac{dS_b}{d\sigma} \Delta \sigma.
\]

According to Fechner, a key property of perceptual attributes is that jnd of the perceptual variable are independent of the strength of the attribute: \(\Delta S_b = k\). Consequently, the sensory-strength function for blur can be constructed by measuring the jnd’s \(\Delta \sigma\) as a function of \(\sigma\) and deriving \(S_b\) from the equation

\[
\frac{dS_b}{d\sigma} = \frac{k}{\Delta \sigma(\sigma)}. \tag{22}
\]

Experimental results from Morgan show that for larger \(\sigma\) values (2.5 \(\leq \sigma \leq 10\) arc min), \(\Delta \sigma\) is proportional to \(\sigma^{1.5}\). Therefore \(S_b\) can be described by

\[
S_b = \frac{1}{\sqrt[3]{\sigma}} + b. \tag{23}
\]

For smaller values of \(\sigma\) (0 \(\leq \sigma \leq 2.5\) arc min) the data of Watt and Morgan indicate that \(\Delta \sigma\) has a minimum. To describe this dipper-shaped part of the \(\Delta \sigma\) curve, we use a slightly modified version of Eq. (23) in which \(\sigma\) is replaced by \((\sigma^2 + \sigma_0^2)^{1/2}\):

\[
S_b = 1 - \frac{1}{[(\sigma/\sigma_0)^2 + 1]^{0.25}}. \tag{24}
\]

For convenience, the constants are chosen such that 0 \(\leq S_b \leq 1\).

The sensory-strength function for periodic structure \(S_p\) in Eq. (10) is derived similarly from the jnd’s \(\Delta m\) as a function of the modulation depth \(m\). Experimental results of Legge and of Carlson and Cohen show that for larger \(m\) values, \(\Delta m\) is proportional to \(m^{\beta}\). A typical value for the exponent is 0.7. However, the exponent increases slightly with the frequency of the sine grating. Consequently, for larger values of \(m\), \(S_p\) is of the form

\[
S_p = am^{1-\beta} + b. \tag{25}
\]

For small values of \(m\) at threshold, the jnd’s \(\Delta m\) are no longer proportional to \(m^{\beta}\). Wilson derived a sensory-strength function valid for both threshold and suprathreshold values. We use a simplified version:

\[
S_p = c \left[1 + (m/m_0)^3\right]^{1/3} - 1 \left(\frac{m}{m_0}\right)^{\beta}. \tag{26}
\]

The constant \(c\) is chosen such that 0 \(< S_p \leq 1\). Note that for higher values of \(m\) the perceptual strength is proportional to \(m^{1-\beta}\).

In Eq. (13) the second term on the right-hand side in the perceptual-strength function for the complement of brightness is based on the 1976 CIE definition of lightness used in the \(L^*a^*b^*\) color space. The minus sign is included since we are interested in the complement of
brightness. When the luminance is divided by the maximum average luminance of the stimuli in an experiment, the function is normalized such that 0 ≤ S_B ≤ 1. The sensory-strength function for the complement of brightness contrast in Eq. (14) is derived from measurements of ∆C as a function of C, i.e., the contrast-discrimination function. The procedure is identical to the derivation of the sensory-strength function for periodic structure described above. Note that the Michelson contrast of a sine grating is the modulation depth. Legge and Kersten measured contrast discrimination of light and dark bars with a rectangular luminance profile and concluded that the data closely resemble the corresponding result for sine gratings. For the larger C values in which we are interested, ∆C is approximately proportional to C^{0.7}. Hence we use −C^{0.3} for the perceptual strength of complement of brightness contrast. The strength function was normalized such that 0 ≤ S_C ≤ 1.

**APPENDIX B: VALUE AND STANDARD DEVIATION OF THE α₀ AND m₀ PARAMETERS**

The procedures for the estimation of the σ₀ and m₀ parameters in the sensory-strength functions for blur and periodic structure as well as the calculation of the corresponding variances are described in this appendix. They are described only for the periodic-structure case. The procedures for blur are identical but simpler, since we do not have to consider the m₀ parameter.

We want to find the parameter values σ₀ and m₀ such that the squares of the calculated sensory strengths of periodic structure S_p² are identical to the squares of the measured impairments I_m² for a linear transformation. This linear transformation is characterized by the parameters a₀ and a₁. The method of least squares is applied to determine the four parameters:

\[
\min_{\sigma_0, m_0, a_0, a_1} \text{RSQ}(\sigma_0, m_0, a_0, a_1),
\]

where the residual sum of squares (RSQ) is equal to

\[
\text{RSQ}(\sigma_0, m_0, a_0, a_1) = \sum_{sc=1}^{SC} \left[ a_1 S_p^2 (sc; \sigma_0, m_0) + a_0 - I_m^2 (sc) \right]^2.
\]

The minimization is split in two:

\[
\min_{\sigma_0, m_0, a_0, a_1} \text{RSQ}(\sigma_0, m_0, a_0, a_1)
\]

\[
= \min_{\sigma_0, m_0, a_0, a_1} \left[ \min_{\sigma_0, m_0, a_0, a_1} \text{RSQ}(\sigma_0, m_0, a_0, a_1) \right].
\]

The minimization over a₀ and a₁ corresponds to finding the linear regression of I_m² on S_p²(σ₀, m₀). Subsequently, σ₀ and m₀ minimize the linear regression residual sum of squares.

The approximate variances and covariances of the least-squares estimates σ₀, m₀, a₀, and a₁ are specified in the 4 × 4 variance matrix V, which equals \( \langle \text{J}^T (sc; \hat{\sigma_0}, \hat{m_0}, \hat{a_0}, \hat{a_1}) \text{J} (sc; \hat{\sigma_0}, \hat{m_0}, \hat{a_0}, \hat{a_1}) \rangle^{-1} \times \text{RSQ}(\hat{\sigma_0}, \hat{m_0}, \hat{a_0}, \hat{a_1})/SC-NP \). (30)

where \( \hat{\sigma_0}, \hat{m_0}, \hat{a_0}, \) and \( \hat{a_1} \) are the least-squares estimates. The superscripts T and −1 denote matrix transposition and matrix inversion, respectively. The elements of the Jacobian matrix J are the partial derivatives of the residual sum of squares to the parameters evaluated for the stimulus conditions and with use of the least-squares estimates of the parameters. The partial derivatives were computational approximations. The RSQ is divided by the number of stimulus conditions (SC) minus the number of estimated parameters (NP), which is four in this case. Note that the number of parameters is only three for the perceptual-strength function for blur.

The expression for V in Eq. (30) is an approximation that is valid only when the residuals \( a_1 S_p^2 (sc; \sigma_0, m_0) + a_0 - I_m^2 (sc) \) are linear in the parameters or when the residuals are small. Since the residuals are already linear in two parameters and, according to calculations, are small, we adopt this approximation.

**APPENDIX C: STANDARD DEVIATION OF THE REGULARIZATION PARAMETER FOR BLUR**

To determine the standard deviation of the regularization parameter for blur (λ_b) we need the variances of the squares of the relative weight parameters for periodic structure (a_p) and blur (a_b). Since the regularization parameter for blur is the squared quotient of these weight parameters:

\[
\lambda_b = a_p^2/a_b^2.
\]

The squares of the relative weight parameters are solved from a set of linear equations:

\[
I = S \begin{bmatrix} a_p^2 \\ a_b^2 \end{bmatrix}.
\]

The elements of the vector I are the squares of the measured impairments for each stimulus condition. The elements of the matrix S are the squares of the sensory strengths of periodic structure and blur for each stimulus.

The variances and covariances of the squared relative weight parameters are approximated by the 2 × 2 variance matrix V:

\[
V = [S^T S]^{-1} \sigma^2,
\]

where \( \sigma^2 \) is the variance of the measured impairments and the superscripts T and −1 denote matrix transposition and matrix inversion, respectively. The variance of the measured impairments is identical for all stimuli that are due to the Thurstone correction.

The sensory strengths that constitute the elements of the matrix S depend on the parameters σ₀ and m₀. Since there is some variation on these parameters (see Appendix B), the matrix S is a noisy matrix. Hence Eq. (33) for the variance matrix is not correct. Calculation shows that the variation of the matrix elements is relatively small compared with the matrix elements and is of
the same order as the variance of the measured impairments. The variance matrix of Eq. (33) is therefore considered a close enough approximation.

The variance of the regularization parameter for blur can be calculated by using standard error-propagation methods (e.g., Chatfield24):

\[
s_{ab}^2 \approx \left( \frac{a_b}{a_p} \right)^2 s_{ap}^2 + \left( \frac{a_p}{a_b} \right)^2 s_{pb}^2 + 2 \frac{s_{ab}}{a_p a_b},
\]

where \( s_{ap}^2 \) and \( s_{pb}^2 \) are the variances of \( a_p^2 \) and \( a_b^2 \) and their covariance, respectively. The variances are on the diagonal of matrix \( V \). The covariance is equal to the off-diagonal elements.

The authors can be reached at tel: 31-0-43-3557027; fax: 31-0-43-3557444; e-mail: m.nijenhuis@libertel.nl.

REFERENCES

16. CIE, Colorimetry, CIE Publication No. 15.2 (CIE, Vienna, 1986).