

# The Image Fidelity Assessor

*Christopher C. Taylor, Jan P. Allebach*  
*Electronic Imaging Systems Laboratory*  
*Purdue University*  
*School of Electrical and Computer Engineering*  
*West Lafayette, IN*  
*{taylor,allebach}@ecn.purdue.edu*

*Zygmunt Pizlo*  
*Visual Perception Laboratory*  
*Purdue University*  
*Department of Psychological Sciences*  
*West Lafayette, IN*  
*pizlo@psych.purdue.edu*

*URL: <http://www.ecn.purdue.edu/EISL/>*

## Abstract

In this paper we describe the Image Fidelity Assessor (IFA), a model of the human visual system designed to evaluate perceived fidelity. In particular, we describe the structure of the model, present results from a contrast discrimination experiment used within the model, and demonstrate preliminary results of the model. The IFA is designed to be both physiologically and psychophysically plausible. It consists of a multichannel Gabor pyramid decomposition. The fidelity at each spatial location within each channel is evaluated based on psychometric functions. The psychometric functions were determined experimentally and describe the discrimination ability of the visual system as a function of spatial frequency, orientation, and adaptation level. Limited memory probability summation across the visual channels is used to obtain a measure of perceived image fidelity. Preliminary results of the IFA are consistent with human evaluation of image fidelity.

## 1. Introduction

A number of models incorporating some type of frequency selective channels have been proposed for image fidelity assessment[1, 2, 3, 4, 5]. These models have made significant progress towards attaining meaningful image fidelity assessments. They go far beyond simple mean-squared error, which is well-known to be unsatisfactory. Our work has been heavily influenced by these models and attempts to build on them.

In order to motivate our approach, we begin by consid-

ering some of the characteristics of these models. All of these models depend on psychophysical data obtained by others. Frequently the task performed in the experiment used to obtain model parameters is not directly related to the function of the corresponding portion of the model. For example, psychophysical experiments that determine the contrast sensitivity function (CSF) in Daly's Visible Differences Predictor[2] (VDP) lead to the relation between contrast and percept, i.e. the probability of seeing a sine grating at a particular frequency and contrast. However, the output of the CSF in the VDP is not treated as probability. This means that this portion of the model is not a model of percept. Other models share these characteristics. We have endeavored to take a more integrated approach with the goal of having our model more closely match the psychophysical experiments upon which it is based.

A number of physiological studies on cats and primates have reported that mammalian visual systems contain neurons whose receptive fields closely resemble Gabor patches [6, 7]. However, since Gabor decompositions are not orthogonal, prior image fidelity assessment methods have preferred to decompose the image by sectioning the frequency domain into a number of pie pieces. On the other hand, Gabor decompositions have been used to address a number of image understanding problems. Many texture segmentation and classification techniques have relied on Gabor decompositions[8]. The IFA differs from previous image fidelity models in that it uses a physiologically plausible Gabor pyramid decomposition.

Another fundamental difference between the IFA and its predecessors is founded on psychophysical grounds. With

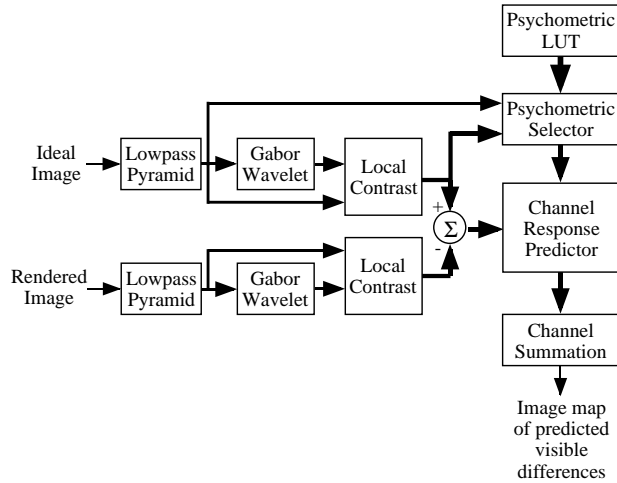


Figure 1: Block diagram of the Image Fidelity Assessor.

the exception of the Sarnoff Model[5], prior image fidelity measures implicitly assume that contrast detection and contrast discrimination are equivalent[2, 3, 4]. In contrast detection experiments the subject is asked whether or not the stimulus is visible. In contrast discrimination experiments the subject is asked to discriminate between two stimuli. In previous models the difference of the two contrast images is computed just prior to the channel response predictor, and information about the contrast in the original images is lost. Then contrast detection thresholds are used to predict their discriminability. These models discard information about the absolute contrast levels in the images; however, this discarded information is relevant to visibility thresholds. A masking calculation is required to compensate when detection thresholds are used. The IFA uses discrimination thresholds in order to obtain a more accurate estimate of image fidelity. We differ from the Sarnoff Model in that we use contrast discrimination data that is specific to our modeling of the visual system.

## 2. Model

The IFA accepts two grayscale images as inputs and generates a probability map as output. The probability map is a grayscale image that indicates the probability of a human observer detecting a difference between the two input images. Figure 1 shows a block diagram of the IFA. A multiresolution decomposition is performed on each image to generate a number of channels, each containing the response of a particular receptive field. The receptive fields are modeled by Gabor functions of varying frequency and orientation. Significant effort has been expended by the research community to create a more complete description of the receptive fields of the neurons in the primary visual cortex. The characteristics of the Gabor functions used in *Proceedings of the 1998 IS&T Image Processing, Image Quality, and Image Capture Systems Conference*, Portland, OR, 17-20 May 1998, in press. (2)

our model follow those of Lee [9] and are motivated by the work of many others.

The multiresolution pyramid is built by low pass filtering and decimating the original image. We call each of these images the *base image* for a particular pyramid level. The base image for each pyramid level is convolved with even and odd symmetric Gabor functions at eight orientations. A local contrast calculation produces contrast images that describe the response of an ensemble of neurons tuned to a particular spatial frequency and orientation. We call these images the *channel images* as they represent different channels of the visual system. Sixteen channel images are generated for five pyramid levels corresponding to fundamental frequencies of 16, 8, 4, 2, and 1 cycles per degree.

The Psychometric Look Up Table (LUT) consists of a family of psychometric functions that have been empirically determined by psychophysical experiments described in Sec. 3. The Psychometric Selector selects the appropriate psychometric function from the family of psychometric functions in the LUT. The selection is made independently for each spatial location in each channel image. Selection of the psychometric function is based on the local adaptation level, the fundamental frequency of the channel image, and the local contrast level in the channel image. The local adaptation level is obtained by low pass filtering the base image. The channel image for the ideal image is used as a measure of the local contrast level in the channel image.

The difference between the contrast images for each channel is then applied to the appropriate psychometric functions to produce a separate probability map for each channel. All of the probability maps from the different channels are combined in the channel summation stage to produce the final probability map. In this stage, the channel probability maps from higher levels of the pyramid are upsampled to match the resolution of the lowest level of the pyramid. The probability maps are then combined using a method we call *limited memory probability summation*. Limited memory probability summation is motivated by brain physiology. The method uses probability summation to combine the five largest visual channel responses at each spatial location.

## 3. Experiment

A psychophysical contrast discrimination experiment was designed to obtain the psychometric functions used within the model. The stimuli used in the experiment were Gabor patches with the same characteristics as the Gabor functions used in the channel decomposition of the IFA.

Each stimulus consisted of a reference patch and a test patch displayed on either side of a fixation cross. The aver-

age luminance, frequency, and orientation for the test patch was always the same as that of the reference patch. The reference patch had a fixed contrast while the test patch varied in contrast. The stimuli were displayed on a calibrated monitor with a peak luminance of  $79 \text{ cd/m}^2$ . One of the authors (CCT) with corrected to normal vision served as the subject in the experiment. For all but the two lowest frequencies tested, the subject's head was placed in a headrest 2.0 m from the monitor. For the two lowest frequencies tested, the viewing distance was reduced to 1.0 m.

Within a session there were eight test patches, each with different contrast levels. Each contrast level was slightly above reference contrast. A fixation cross remained at the center of a uniform gray image throughout each session. The uniform gray image had the same luminance as the average luminance of the Gabor patches. In each trial the two Gabor patches were presented at a horizontal eccentricity of 2.5 cycles to either side of the fixation cross. Each trial was initiated by pressing the middle mouse button, whereupon a 100 ms presentation of the trial stimulus followed a 50 ms delay. Each test patch was presented 50 times. The test patches were presented in random order, and the side of presentation for the test and references patches was randomized.

The subject was asked to indicate which patch (left or right) had higher contrast. Auditory feedback was provided after each incorrect response. Cumulative Gaussian distributions were fit to the results using probit analysis [10]. The standard deviation of each distribution was used to estimate the discrimination thresholds. Discrimination thresholds were measured for Gabor patches with six frequencies, eight orientations, five average luminance levels, and nine reference contrast levels. A portion of the experimental results are shown in Fig. 2. The figure contains discrimination functions which consist of discrimination thresholds plotted as a function of reference contrast.

It is apparent from the discrimination functions in Fig. 2 that discrimination and detection thresholds are not equivalent. For if they were, all of the discrimination functions would be horizontal lines. The discrimination functions in Fig. 2 reveal that discrimination thresholds and reference contrasts are inversely related for subthreshold reference contrasts and directly related for suprathreshold reference contrasts. This observation, often referred to as *facilitation* or the *dipper effect*, is consistent with the results of previous experiments on sinewave stimuli. The minimum discrimination threshold occurs when the reference contrast is near the detection threshold (this contrast detection threshold is measured in a contrast discrimination experiment with the reference contrast equal to zero). For example, consider the contrast discrimination function for the 2.0 c/deg Gabor patch in Fig. 2. The minimum for this

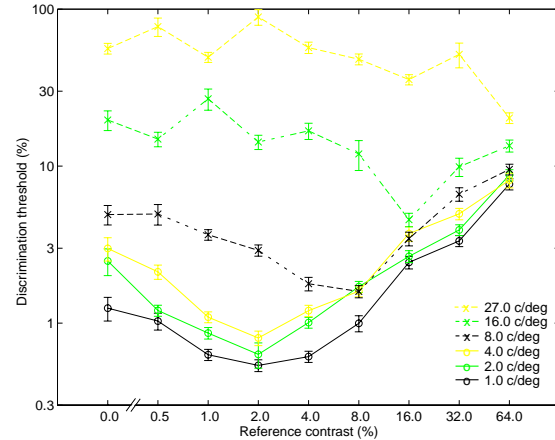


Figure 2: Discrimination thresholds for vertically oriented Gabor patches with spatial frequencies of 1.0, 2.0, 4.0, 8.0, 16.0, and 27.0 c/deg with an average luminance level of  $40 \text{ cd/m}^2$ .

curve occurs at a reference contrast of 2% which is close to the discrimination threshold when the reference contrast is zero.

It is interesting to note that all of the curves in Fig. 2 converge as the reference contrast is increased. In fact, for all but a few Gabor patches studied, the discrimination thresholds are near 10% when the reference patch is at 64% contrast. This implies that contrast discrimination is approximately insensitive to frequency variations in suprathreshold contrast regions. This effect has been referred to as *contrast constancy* and has been widely observed for suprathreshold contrasts [11, 12]. In addition, the increasing portion of the discrimination functions is linear on a log-log plot. The slopes of the linear portion of the discrimination functions obtained by the experiment range from 0.6 to 0.9. This range of slopes is within the range of slopes reported by others [13, 14, 15, 16]. Although discrimination functions for luminance and orientation variations are not shown here, the visual system exhibited little sensitivity to orientation for all reference contrasts and little sensitivity to average luminance for suprathreshold reference contrasts. For subthreshold reference contrasts, discrimination thresholds increased as average luminance decreased.

## 4. Results

In this section we present two examples of image fidelity predictions generated with the IFA. Light areas of the IFA's output probability map correspond to image regions with high likelihood of visible distortions while dark areas correspond to image regions with low likelihood of visible

distortions.

## Acknowledgment

This work was supported by the Hewlett-Packard Company.

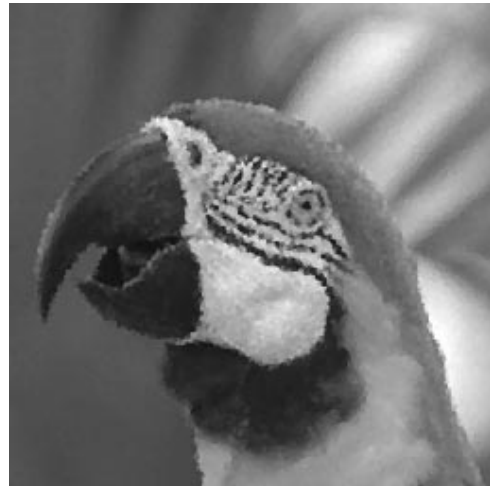
## References

- [1] A. B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vision Graphics and Image Process.*, vol. 39, no. 3, pp. 311 – 327, September 1987.
- [2] S. Daly, "The visible differences predictor: An algorithm for the assessment of image fidelity," in *Digital Images and Human Vision* (A. B. Watson, ed.), pp. 179 – 205, Cambridge, MA: MIT Press, 1993.
- [3] D. J. Heeger and P. C. Teo, "A model of perceptual image fidelity," *Proc. of IEEE Int'l Conf. on Image Proc.*, Oct. 23 – 26 1995, Washington, D.C., USA, pp. 343 – 345.
- [4] S. J. P. Westen, R. L. Legendijk, and J. Biemond, "Perceptual image quality based on a multiple channel HVS model," *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, 1995, pp. 2351 – 2354.
- [5] J. Lubin, "A visual discrimination model for imaging systems design and evaluation," in *Vision Models for Target Detection and Recognition* (E. Peli, ed.), pp. 245 – 283, Singapore: World Scientific, 1995.
- [6] D. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science*, vol. 212, no. 4501, pp. 1409 – 1411, June 18 1981.
- [7] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in the cat striate cortex," *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1233 – 1258, 1987.
- [8] D. F. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Trans. on Image Processing*, vol. 4, no. 7, pp. 947 – 964, July 1995.
- [9] T. S. Lee, "Image representation using 2D Gabor-wavelets," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959 – 971, October 1996.
- [10] D. J. Finney, *Probit analysis*. Cambridge Univ. Press, 3 ed., 1971.

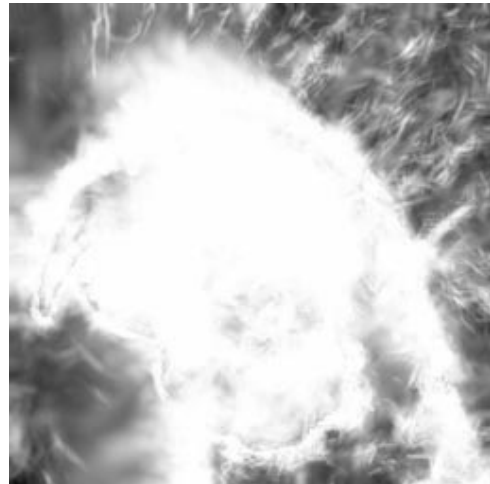
*Proceedings of the 1998 IS&T Image Processing, Image Quality, and Image Capture Systems Conference*, Portland, OR, 17-20 May 1998, in press. (4)



(a)



(b)



(c)

Figure 3: (a) Original image, (b) Image distorted by Adobe Photoshop's "crystallize" algorithm, (c) IFA image fidelity prediction.



(a)



(b)



(c)

Figure 4: (a) Original image, (b) Image distorted by median filtering, (c) IFA image fidelity prediction.

- [11] W. H. Swanson, H. R. Wilson, and S. C. Giese, "Contrast matching data predicted from contrast increment thresholds," *Vision Research*, vol. 24, no. 1, pp. 63 – 75, 1984.
- [12] M. W. Cannon, Jr., "Perceived contrast in the fovea and periphery," *J. Opt. Soc. Am. A*, vol. 2, no. 10, pp. 1760 – 1768, 1985.
- [13] J. Nachmias and R. V. Sansbury, "Letter to the editors, Grating contrast: Discrimination may be better than detection," *Vision Research*, vol. 14, no. 10, pp. 1039 – 1042, 1974.
- [14] G. E. Legge, "A power law for contrast discrimination," *Vision Research*, vol. 21, no. 4, pp. 457 – 467, 1981.
- [15] A. Bradley and I. Ohzawa, "A comparison of contrast detection and discrimination," *Vision Research*, vol. 26, no. 6, pp. 991 – 997, 1986.
- [16] J. Yang and W. Makous, "Modeling pedestal experiments with amplitude instead of contrast," *Vision Research*, vol. 35, no. 14, pp. 1979 – 1989, 1995.