NASA/CR-1998-208207

IN-33-CR
151455

Sarnoff Corporation

FINAL REPORT: SARNOFF JND VISION MODEL FOR FLAT-PANEL DESIGN

Prepared for: National Aeronautics and Space Administration
Ames Research Center
Moffett Field, CA 94035-1000

Contract No.: NAS2-14257

Period of Performance: Jan. 1996 - Jan. 1998

Prepared by: Visual Information Systems Research Group
Information Sciences Laboratory
Sarnoff Corporation
CN 5300
Princeton, NJ 08543-5300

May 20, 1998

TABLE OF CONTENTS

1. Introduction

This document describes the adaptation of the basic Sarnoff JND Vision Model created in response to the requirement of the NASA/ARPA program for a general-purpose model for predicting the perceived image quality attained by flat-panel displays. The software was delivered to NASA Ames and is being integrated with LCD display models at that facility.

2. The Sarnoff JND Vision Model

The Sarnoff JND Vision Model is a method of predicting the perceptual ratings that human subjects will assign to a degraded color-image sequence relative to its nondegraded counterpart. The model takes in two image sequences and produces several difference estimates, including a single metric of perceptual differences between the sequences. These differences are quantified in units of the modeled human just-noticeable difference (JND). A version of the model that applies only to static, achromatic images is described by Lubin (1993, 1995).

The Sarnoff Vision Model can be useful in a general context (see Figure 1). An input video sequence passes through two different channels on the way to a human observer (not shown in the figure). One channel is uncorrupted (the reference channel), and the other distorts the image in some way (the channel under test). The distortion, a side effect of some measure taken for economy (or a necessary effect of the display technology), can occur at an encoder prior to transmission, in the transmission channel itself, or in the decoding process. In Figure 1, the box called "system under test" refers schematically to any of these alternatives. Ordinarily, evaluation of the subjective quality of the test image relative to the reference sequence would involve the human observer and a real display device. This evaluation would be facilitated by replacing the display and observer by the JND model, which compares the test and reference sequences to produce a sequence of JND maps instead of the subjective comparison.

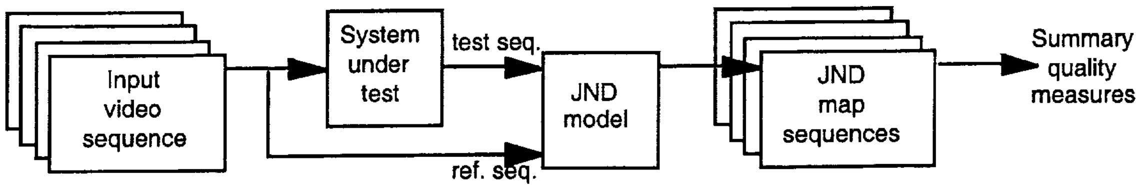


Figure 1. JND Model in System Evaluation

Figure 2 shows an overview of the JND Model architecture. The inputs are two image sequences of arbitrary length. For each image of each input sequence, there are three data sets, labeled X, Y, and Z (using the 1931 CIE System) at the top of Figure 2. To model a display (e.g., an LCD), these data can be sampled at many times the pixel resolution and at many times the digital frame rate. The result is a consecutive stack of images of CIE 1931 X, Y, Z values from a test and a reference image sequence.

The first stage of the model, labeled Front-End Processing in Fig. 2, first downsamples each sequence in time and in space to physiologically reasonable rates. The result is saved in a stack of four tristimulus images (X, Y, Z) representing four time slices at the chosen physiological rate. The luminance arrays Y are passed to luma processing, and all the arrays are transformed so ensure (at each spatial point) approximate perceptual uniformity of the color space to isoluminant color differences. To accomplish this goal, the individual pixels are mapped into CIELUV, an international-standard uniform-color space (see Wyszecki and Stiles, 1982). The chroma components u^*, v^* of this space are passed to the chroma processing steps in the model.[1]

[1] The luminance channel L^* from CIELUV is not used in luma processing, but instead is replaced by a visual nonlinearity for which the vision model has been calibrated over a range of luminance values. L^* is used in chroma processing, however. to create a chroma metric that is approximately uniform and familiar to display engineers.

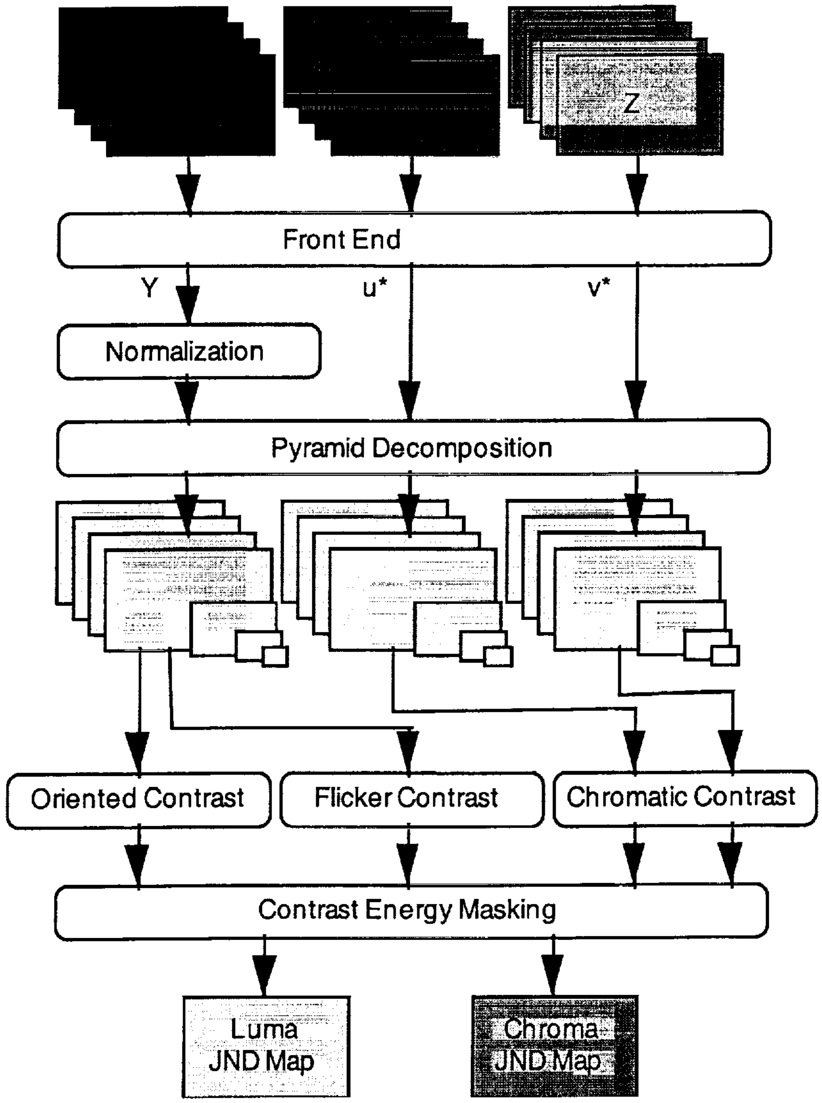


Figure 2. Architecture of the Sarnoff Vision Model. Note that one further step, the single-number summary of the JND map, is not represented in this figure.

Luma processing in the JND model accepts two images (test and reference) of luminances Y, expressed as fractions of the maximum luminance of the display. First, a point nonlinearity (which depends on overall light level) effects luma compression. Next, each sequence is filtered and down-sampled using a Gaussian pyramid operation (Burt and Adelson, 1983) to efficiently generate a range of spatial resolutions for subsequent filtering operations. Then contrast arrays (local differences divided by local sums) are calculated at each pyramid level, and scaled to be 1 when the image contrast is at the human detection threshold. Finally, these scaled contrast arrays are subjected to masking nonlinearities (to desensitize in the presence of image "busyness") and compared between test and reference to produce a JND map. This map is an image whose gray levels are proportional to the number of JNDs between the test and reference image at the corresponding pixel location. The parameters of the contrast computation are fit according to contrast-detection data of van Nes et al. (1967) and Koenderink et al. (1979). The point non-linearity of masking is fit according to contrast discrimination data (Carlson and Cohen, 1978).

Similar processing, based on the CIE L*u*v* uniform-color space, occurs for each of the chroma images u* and v*. Outputs of u* and v* processing are combined to produce the chroma JND map. Before creation of the chroma JND map, the chroma outputs are subjected to masking from both chroma and luma channels so as to render perceived differences more or less visible depending on the structure of the luma images. The parameters of the contrast computation are fit according to contrast-detection data of Mullen (1985), and the point-nonlinearity of masking is fit according to contrast discrimination data (Switkes, et al., 1988).

The chroma and luma JND maps are each available as output, together with a small number of summary measures derived from these maps. For each field in the video-sequence comparison, the luma and chroma JND maps are first combined to give a total-JND map. Then, each of the three JND maps (luma, chroma, and total) is reduced to a single-number summary, namely a JND-Aggregate-Measure (JAM) value. Finally, three single performance measures for many fields of a video sequence (one for luma, one for chroma, and one for both luma and chroma) are determined from the corresponding single-field JAMs.

Whereas the single summary JAM values are useful to model an observer's overall comparative rating of the test sequence with respect to the reference sequence, the JND maps give a more detailed view of the location and severity of the artifacts.

3. Comparisons with Rating Data

Four image sequences, each with various degrees of distortion, were used to compare the Sarnoff Vision Model with DSCQS rating data. The model accommodated one pixel per image-resolution cell and one inter-field interval per model epoch. The following viewing conditions were assumed: A color CRT display with a gamma of 2.5, phosphor chromaticities as specified in the ITU BT-709 standard, viewing conditions as specified by the ITU-R Rec 500, and a maximum screen luminance of 100 cd/m^2.

The results are plotted in Figure 3, and reveal a correlation 0.92 between the model and the data. For each of the sequences, the Vision Model processed 30 fields. The high correlation instills confidence that the model will be successful in predicting the results of future tests.

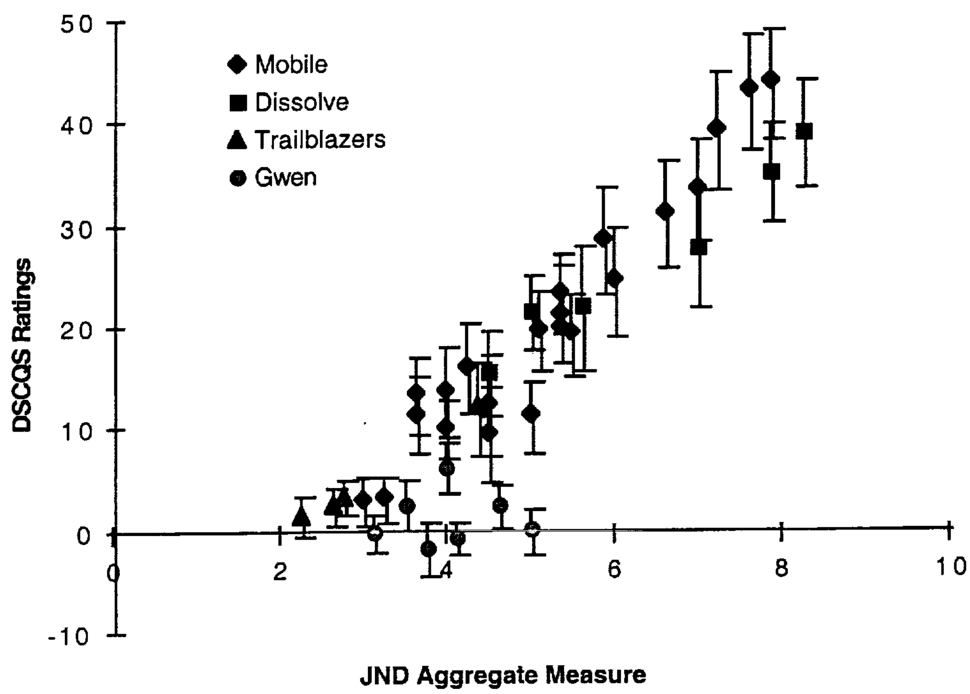


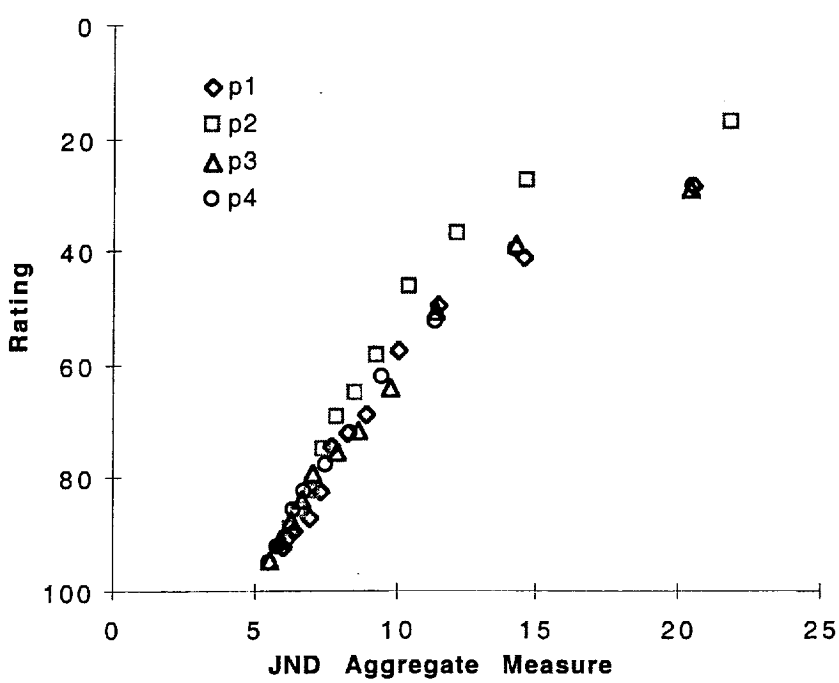
Figure 3. MPEG-2 Rating Predictions, 30 Fields Per Sequence.

In addition to these MPEG rating predictions, we have rerun the latest model on some JPEG rating data first reported in Lubin, 1995. For this task, observers were shown four different scenes (p1, p2, p3, and p4) each compressed at 11 different JPEG levels. Observers were then asked to rate the quality of each resulting still image on a 100-point scale (100 being best).

As shown in Figure 4, the model does a good job predicting the rating data, with excellent clustering across image types and a strong linear correlation over the entire rating range (.94). Even better correlation (0.97) results when one omits the four points above 15 JNDs, for which some saturation at the low end of the rating scale has evidently occurred.

On the other hand, as shown in Figure 5, correlation among ratings and predictions based on the root mean-squared error between the original and compressed images are not nearly as good (.81). Here, the predictions do not track well across image types, even though a monotonic relation between rating and predicted value is observed within each image.

Figure 4. Predictions of Final Model on JPEG rating data



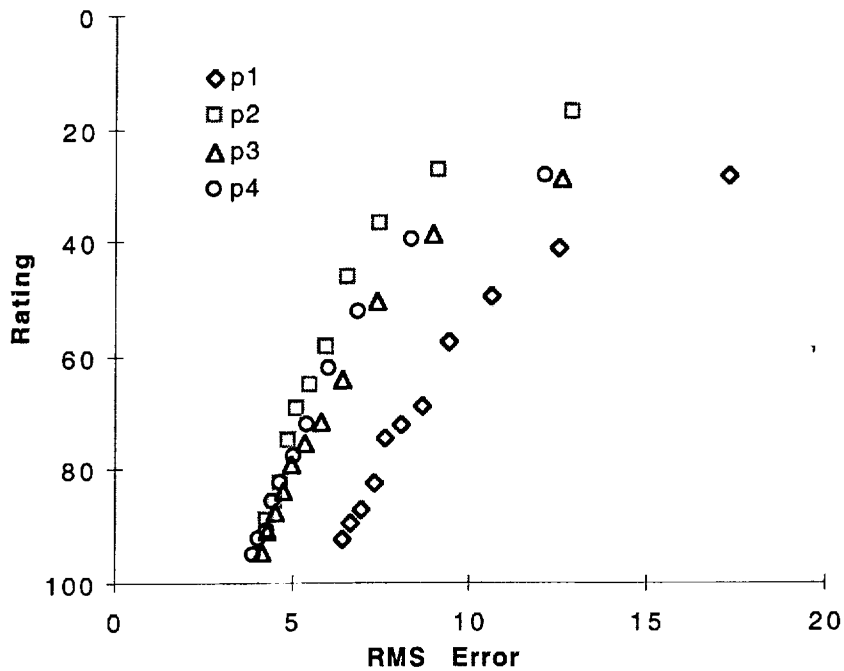


Figure 5. RMS error predictions on JPEG rating data

4. Conclusions

Substantial flexibility has been incorporated into the Sarnoff JND Vision Model so it may be used to model displays at the sub-pixel and sub-frame level. Sub-sampling has been engineered so as to minimize interpolative artifacts and aliasing.

The latest model extensions--into temporal and chromatic domains--have done well in calibration against psychophysical data and against image-rating data given a CRT-based front-end. Future, more extensive testing of the model remains to be done with LCD displays at various resolutions relative to pixel and frame rates. We are confident that this product will successfully predict subjective ratings for a full range of spatio-temporal and chromatic image sequences.

Acknowledgments: We thank Dr. James Larimer at NASA Ames for guidance and discussions during the course of this effort.

5. References

P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, **COM-31**, 532-540 (1983).

C. R. Carlson and R. W. Cohen, "Visibility of Displayed Information", Report prepared for Office of Naval Research by RCA Laboratories, Princeton, NJ, July 1978. Contract No. N00014-74-C-0184.

C. R. Carlson and R. W. Cohen, "A simple psychophysical model for predicting the visibility of displayed information. *Proc. Soc. Inform. Display* **21**, 229-245 (1980).

C. H. Graham (Ed.), *Vision and Visual Perception..* Wiley, 1966.

J. J. Koenderink and A. J. van Doorn, "Spatiotemporal contrast detection threshold surface is bimodal," *Optics Letters* **4**, 32-34 (1979).

J. Lubin, The use of psychophysical data and models in the analysis of display system performance. In A. B. Watson (ed.), *Digital Images and Human Vision*, MIT Press, 1993, pp. 163-178.

J. Lubin, A visual system discrimination model for imaging system design and evaluation, to be published in E. Peli (ed.),*Visual Models for Target Detection and Recognition*, World Scientific Publishers, 1995.

E. Martinez-Uriegas, "Color Discrimination Thresholds for the Model of Visibility of Color Liquid-Crystal Displays," SRI report to Sarnoff, January 1996.

K. T. Mullen, "The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings," *J. Physiol.* **359**, 381-400, 1985).

J. Nachmias and R. V. Sansbury, "Grating contrast: Discrimination may be better than detection," *Vision Res.* **14**, 1039-1042 (1974).

E. Switkes, A. Bradley, and K. De Valois, "Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings,"*J . Opt. Soc. Am.* **A5**, 1149-1162 (1988).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE 20 May 1998	3. REPORT TYPE AND DATES COVERED Final Report, Jan 1996-Jan 1998

4. TITLE AND SUBTITLE
Final Report: Sarnoff JND Vision Model for Flat-Panel Design

5. FUNDING NUMBERS
Contract No. NAS2-14257

6. AUTHOR(S)
Michael H. Brill and Jeffrey Lubin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Sarnoff Corporation
CN5300
Princeton, NJ 08543-5300

8. PERFORMING ORGANIZATION REPORT NUMBER
Report No. 29

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Ames Research Center
Moffett Field, CA 94035-1000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Unlimited

12. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This document describes adaptation of the basic Sarnoff JND Vision Model created in response to the NASA/ARPA need for a general-purpose model to predict the perceived image quality attained by flat-panel displays. The JND model predicts the perceptual ratings that humans will assign to a degraded color-image sequence relative to its nondegraded counterpart. Substantial flexibility is incorporated into this version of the model so it may be used to model displays at the sub-pixel and sub-frame level. To model a display (e.g., an LCD), the input-image data can be sampled at many times the pixel resolution and at many times the digital frame rate. The first stage of the model downsamples each sequence in time and in space to physiologically reasonable rates, but with minimium interpolative artifacts and aliasing. Luma and chroma parts of the model generate (through multi-resolution pyramid representation) a map of differences between test and reference called the JND map, from which a summary rating predictor is derived. The lastest model extensions have done well in calibration against psychophysical data and against image-rating data given a CRT-based front-end. THe software was delivered to NASA Ames and is being integrated with LCD display models at that facility.

14. SUBJECT TERMS Vision Model, Flat-Panel, Displays	15. NUMBER OF PAGES 11
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT 200 words