

Real-time Simulation of Arbitrary Visual Fields

Wilson S. Geisler
University of Texas at Austin
geisler@psy.utexas.edu

Jeffrey S. Perry
University of Texas at Austin
perry@psy.utexas.edu

Abstract

This report describes an algorithm and software for creating and displaying, in real time, arbitrary variable resolution displays, contingent on the direction of gaze. The software produces precise, artifact-free video at high frame rates in either 8-bit gray scale or 24-bit color. The software is demonstrated by simulating the visual fields of normal individuals and low-vision patients.

Keywords

variable resolution image, eye movements, gaze contingent display, visual fields, low vision, eye disease, foveated imaging

1. Introduction

During the last twenty years there has been considerable interest in gaze contingent video displays. Of particular interest have been variable-resolution displays where spatial resolution is contingent upon the direction of gaze.

One area of application for such display technology is image compression. The spatial resolution of the human visual system falls off precipitously from the point of gaze (the fovea) and hence it is possible to reduce considerably the spatial resolution of video images away from the line of sight with little effect on perceived image quality or on performance in visual tasks. Such displays are most effective when used in conjunction with an eye tracker, but they are still useful with less sophisticated pointing devices (such as a mouse) or with appropriate algorithms for the automatic selection of foveation points. Another area of application is experimental research on visual processing. Specifically, real-time variable resolution displays allow the researcher to manipulate the spatial information across the visual field, and hence to test hypotheses about the role of peripheral and foveal vision in different tasks (e.g., visual search or reading). Another area of application is in education. For example, real-time simulation of visual field defects can be used to educate students, physicians and patients' family members about the perceptual and performance consequences of vision loss.

Early real-time variable resolution display systems [Juday and Fisher 1989; Weiman 1990; Waner et al. 1993; Silsbee et al. 1993; Barnett and Bovik 1996; Kortum and Geisler 1996] suffered from blocking artifacts, limited display sizes, or limited control of resolution (e.g., only a high resolution area of interest and a low

resolution background area).

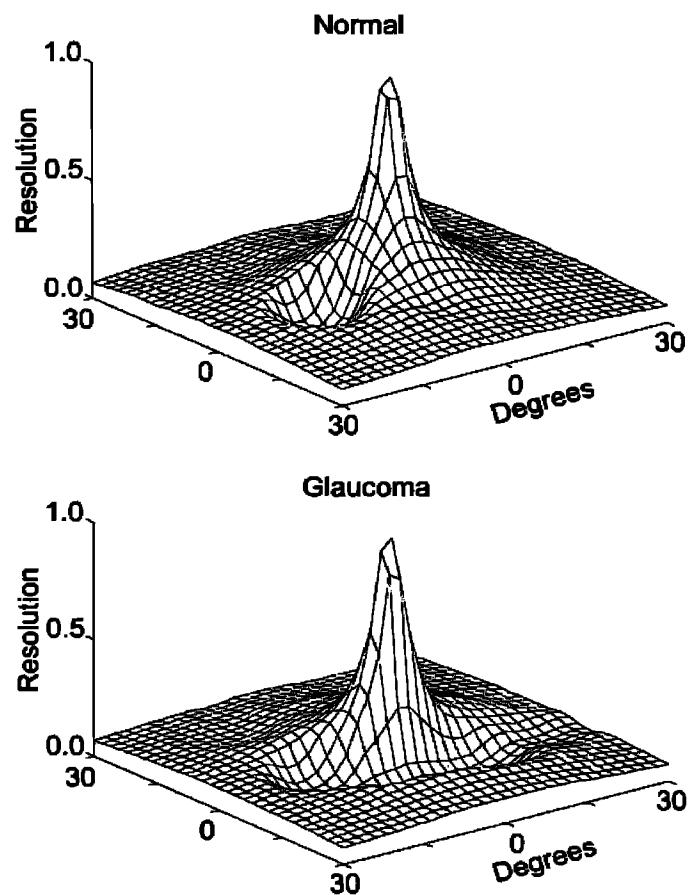


Figure 1. Resolution maps for a normal individual and for a patient with glaucoma. The dip to the left of the foveal peak in the normal patient's map is the "blind spot," where the optic nerve leaves the eye.

Several years ago we described a multi-resolution pyramid method for creating variable resolution displays in real time using general-purpose computers [Geisler and Perry 1998, 1999]. Software implementing this method on Windows 95/98/2000/NT platforms is available at the website, fi.cvis.psy.utexas.edu. This software produces smooth nearly artifact-free images at high frame rates, but it is designed primarily for image compression, and it is limited to displays that mimic the fall-off in resolution of the human visual system.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ETRA'02 New Orleans Louisiana USA
Copyright ACM 2002 1-58113-447-3/02/03...\$5.00

Here, we describe an extension of the previous method that is designed for research and simulation (rather than compression) and allows completely arbitrary variable resolution displays. The new software produces artifact free gaze contingent video at high frame rates in either 8-bit gray scale or 24-bit color. We demonstrate the new software by simulating the vision of patients with various forms of visual field loss (due to eye disease). Gaze directions are measured and images are processed at some selected frame rate (e.g., 30 frames/sec).

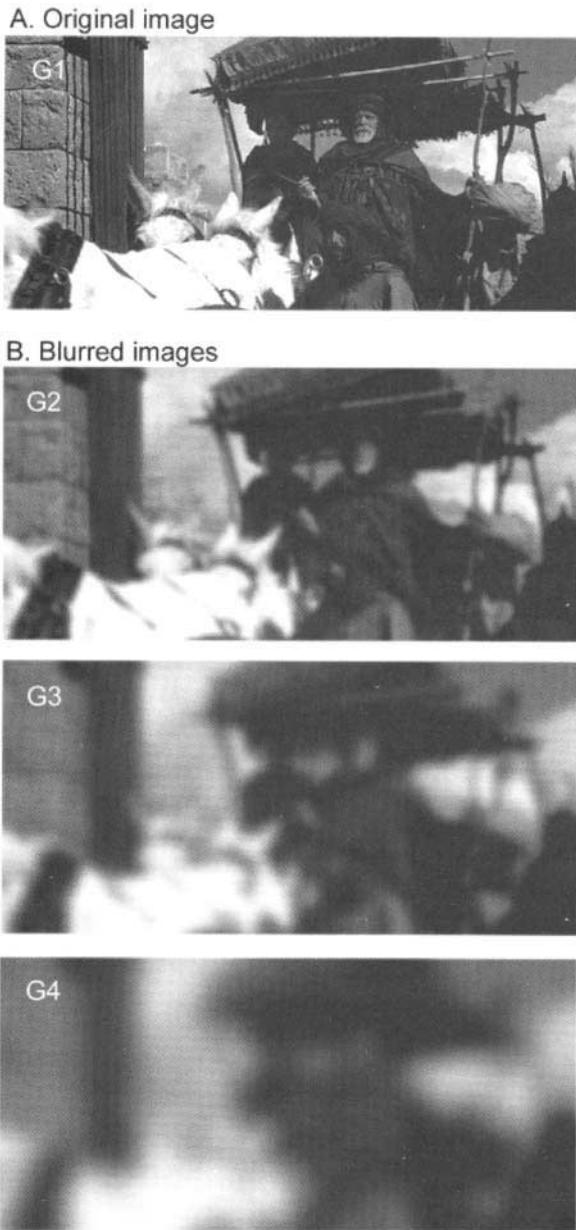


Figure 2. First four levels of the multiresolution pyramid, after up sampling and interpolation to full display size.

2. Methods

The software takes as input the direction of gaze (from an eye tracker or a mouse) and either a single image or a sequence of images. The sequence of images may be retrieved either from memory, a hard drive, or a frame-grabber attached to a camera.

2.1 Variable Resolution Map

The spatial resolution (acuity) desired at each pixel location is specified with a resolution map. This map is specified beforehand, but it is shifted on the fly to be centered on the current gaze direction. The dimensions of the map are typically twice the dimensions of the display, so that the whole display can be rendered for any gaze coordinates within the display region. The resolution at each pixel is specified as a real number between zero and one but is ultimately represented by an 8-bit number.

Figure 1 shows two resolution maps, one for a normal individual and one for a glaucoma patient with a visual field loss below the fovea. The normal map is based upon standard formulas for acuity as a function of eccentricity (see Geisler & Perry, 1998). The glaucoma map was obtained by interpolating measurements from a Goldmann perimeter. Another representation of the glaucoma patient's resolution map is shown at the top of Figure 3A (resolution is represented by gray level).

2.2 Multiresolution Pyramid

The first step in processing each frame is to compute a multiresolution pyramid. To do this we convolve (blur) the image using a small kernel, such as the 3×3 weighting function of the Laplacian pyramid [Burt and Adelson 1984], and then down-sample the blurred image by a factor of 2 in each direction to obtain the second level of the pyramid. The second level of the pyramid is then convolved (with the same kernel) and down sampled by a factor of 2 in each direction to obtain the third level of the pyramid. This process is repeated until all the levels of the pyramid are obtained, typically 6-7 levels. The number of levels is determined by the maximum amount of blur required in the simulation.

When the variable resolution image is constructed from the pyramid, the image at each level of the pyramid is interpolated and up-sampled to full scale, but only in certain regions of the display, as determined by the variable resolution map (see below). Figure 2 shows the images at the first four levels of the pyramid after interpolation and up-sampling to full scale (for the entire display).

2.3 Blending and Up-sampling

The next step in processing each frame is to shift the resolution map to the current gaze position obtained from an eye tracker or some other pointing device. Then, the resolution at each pixel location (as defined by the resolution map) is rendered by blending those two images in the pyramid whose spatial resolutions bracket the desired resolution. This blending at each pixel consists of taking the weighted sum of the pixel values in the bracketing images. The blending weights for the glaucoma patient's resolution map are shown in Figure 3B. In these blending maps, pixels that are black indicate a weight of 0.0 on the higher resolution image and a weight of 1.0 on the lower resolution image. Similarly, pixels that are maximum white indicate a weight of 1.0 on the higher resolution image and a weight of 0.0 on the lower resolution image. Obviously, pixels do not need to be blended when the weights are 0.0 or 1.0. All other gray levels indicate weights between 0.0 and 1.0. Two

multiplications and a sum are required to blend each pixel that has a weight between 0.0 and 1.0.

A. Resolution map



B. Resolution map divided into frequency bands

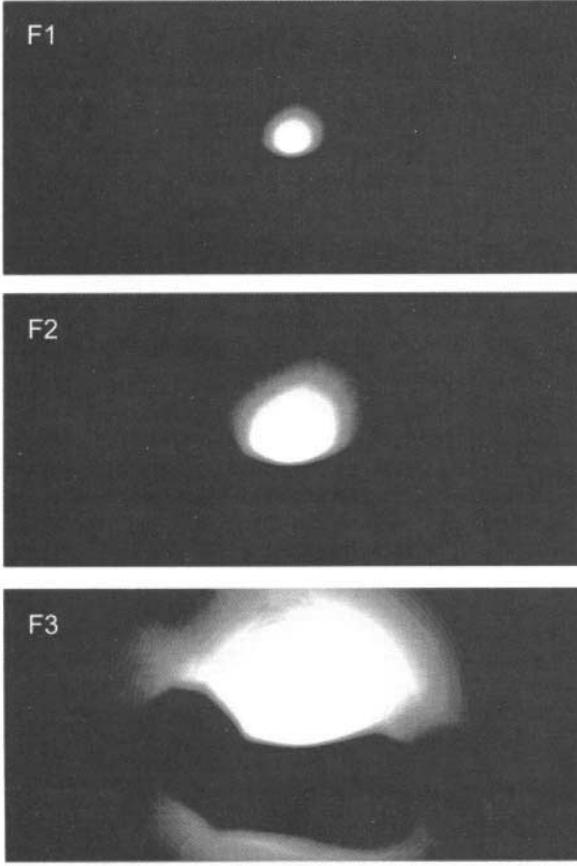


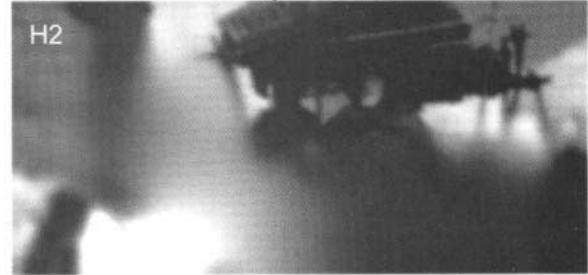
Figure 3. A. Resolution map. B. Blending maps derived from the resolution map, for the first four levels of the pyramid.

The steps of blending and up-sampling are illustrated in Figure 4. Again, keep in mind that this figure only shows the steps for four levels of the pyramid; in normal operation there are typically 6-7 levels. The construction of the final image begins at the bottom of the figure. First, pyramid images G3 and G4 (in Figure 2) are blended using the blending map F3 (in Figure 3) to obtain H3. Next, the pyramid image G2 is blended with H3, using the blending map F2 to obtain H2. The final image is obtained by blending the pyramid image G1 with H2 using blending map F1. Figure 5 shows a final full color image (see color plate).

Blend G1 & H2 using F1



Blend G2 & H3 using F2



Blend G3 & G4 using F3



Figure 4. Construction of final image from the multiresolution pyramid and the resolution map.

2.4 Display

After the variable resolution image is created it is sent to the graphics card for display on the monitor. In the case of gray scale images, the 8-bit data are simply copied to graphics memory. In the case of color images, all the processing is done in standard YUV color space [Recommendations of the CCIR 1990] then converted to RGB color space before copying the 24-bit data to graphics memory.

2.5 Calibration

An important feature of our software is that the displays are created with high precision. For the kernel used in computing the pyramid we calculate the transfer function for the entire process of low-pass filtering, down-sampling, blending, and up-sampling. We do this by processing a wide range of sinewave grating images and then comparing the final outputs with the input images. Thus, we know the whole system transfer function for each possible value of the resolution map between 0.0 and 1.0. This allows us to set the blending weights to give exactly the desired resolution at each pixel.

3.0 Results

In general, the quality of the visual field simulation appears to be excellent. We have tested a large number of different still images

and video sequences. There are no visible blocking artifacts, edge artifacts, shimmering or other kinds of image noise (see Figure 6). The real-time performance of the simulator software was measured by processing a grayscale sequence and a color sequence of 60, 640 x 304 video frames. The grayscale video contained 8-bit pixels. The color video contained pixel data separated into luminance (Y) and color planes (U, V), where the color planes were down-sampled by a factor of two in both dimensions, and thus the 24-bit color images were represented with an average of 12 bits per pixel. The multiresolution pyramid was 7 levels deep and the resolution map was the one shown in Figure 3A. The performance measurements, shown in Table 1, include reading the video images, measuring the direction of gaze, processing the video images according to the resolution map,

converting the images to RGB color space and displaying the images. (Note that performance measurements are sensitive to the particular graphics card and its settings, as well as other factors such as the particular CPU targeted by the compiler's optimizer.)

Intel® Pentium® Processor	Grayscale	Color
800 MHz	60	20
1500 MHz	56	28

Table 1. Performance on 640 x 304 video in frames per second.

A. Original



B. Ramped visual field

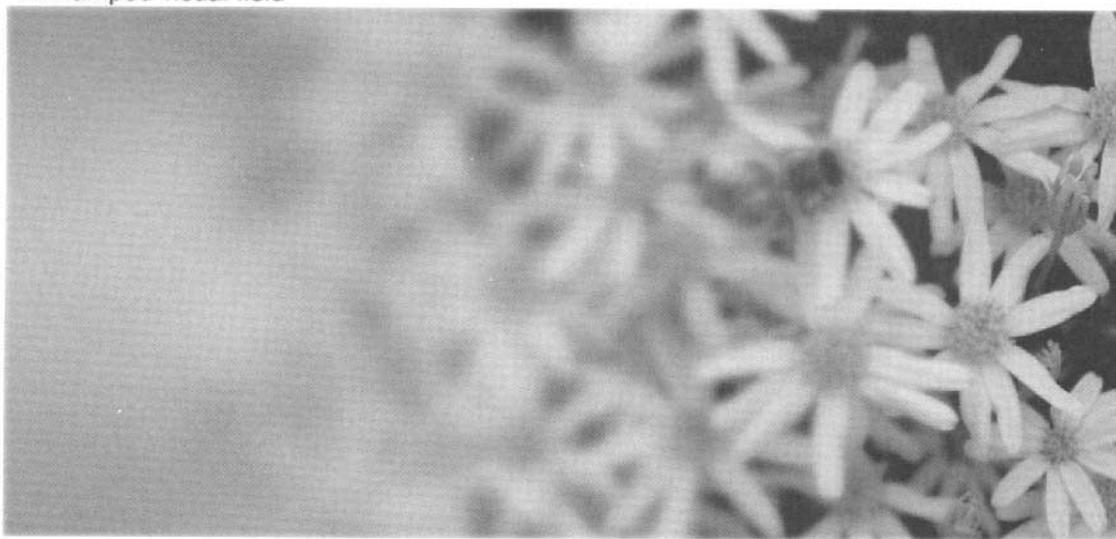


Figure 6. Demonstration of smooth artifact free variable resolution imaging. The image in B was obtained with a resolution map that ramped from .015 to 1.0 on a logarithmic scale.

4. Discussion

The algorithm and software described here show that it is possible to create precision gaze-contingent video displays at relatively high frame rates on inexpensive general-purpose personal computers. To our knowledge there is no other existing algorithm and software that can produce such high-quality variable-resolution video at useful frame rates. The availability of this software raises the possibility of applications in basic and applied research.

The most obvious uses in basic research might be in the study of how the spatial and chromatic information available at different eccentricities affect visual performance in tasks such as visual search or visual navigation (heading perception) [Kortum and Geisler 1996; Van Diepen and Wampers 1998; Duchowski and McCormick 1998; Parkhurst et al. 2000; Loschky and McConkie 2000]. The algorithm and software are sufficiently precise to allow good control of the spatio-chromatic transfer function at each location in the visual field, and hence it is possible to systematically manipulate the spatio-chromatic information across the visual field during complex dynamic tasks. A related application might be in the study of how the spatial and chromatic information available at different eccentricities affect patterns of eye movements during the performance of visual tasks.

The uses in applied research might include studying in normals the effects of the various visual-field defects (produced by eye disease) on task performance and eye movements. It may also be possible to study in normals the learning of fixation strategies in response to the onset of a visual field defect.

Other applied uses may be in education. For example, the simulation software could be used to provide a patient's family members or a physician with a fuller understanding of the consequences of the patient's visual impairment for driving, walking, reading, and so on. Similarly, the software can be used to simulate the spatio-chromatic resolution of the visual system during the course of development. High quality real-time simulations (even without eye tracking) can provide a dramatic and vivid sense of the patient's or infant's visual limitations. This is especially true because the user actively controls gaze direction while viewing the video.

5. Acknowledgments

We thank Eli Peli for supplying us with Goldmann perimeter measurements on the glaucoma patient, and for encouraging us to consider simulations of the visual fields of low vision patients. Supported by NIH grant EY02688.

6. References

Barnett, B.S. and Bovik, A.C. Motion compensated visual pattern image sequence coding for full motion multisession video conferencing on multimedia workstation. *Journal of Electronic Imaging*, 5, 1996, 129-143.

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

Duchowski, Andrew T. and McCormick Bruce H. 1998 "Gaze-Contingent Video Resolution Degradation", in *Human Vision and Electronic Imaging II*, SPIE, Bellingham, WA.

Geisler W.S. and Perry J.S. 1998. A real-time foveated multiresolution system for low bandwidth video communication. *SPIE Proceedings: Human Vision and Electronic Imaging*, 3299, 294-305.

Geisler W.S. and Perry J.S. 1999. Variable resolution displays for visual communication and simulation. *Society for Information Display*, 30, 420-423.

Juday R.D. and Fisher T.E. 1989. Geometric transformations for video compression and human teleoperator display. *SPIE Proceedings: Optical Pattern Recognition*, 1053, 116-123.

Kortum P.T. and Geisler W.S. 1996. Implementation of a foveated image-coding system for bandwidth reduction of video images. *SPIE Proceedings: Human Vision and Electronic Imaging*, 2657, 350-360

Kortum, P.T. and Geisler, W.S. 1996. Search performance in natural scenes: the role of peripheral vision. *Investigative Ophthalmology & Visual Science Supplement*. (ARVO) 37/3, S297.

Loschky L.C. and McConkie, G.W. 2000. User performance with gaze contingent multiresolution displays. *Proceedings of the ACM Symposium on Eye Tracking Research & Applications*, 97-109.

Parkhurst D., Culurciello E. and Niebur E. 2000. Evaluating variable resolution displays with visual search: Task performance and eye movements. *Proceedings of the ACM Symposium on Eye Tracking Research & Applications*, 105-109.

Recommendations of the CCIR 1990. Encoding parameters of digital television for studios. *International Telecommunication Union*: Geneva, XI-1, pp 95-104.

Silsbee P.L., Bovik A.C. and Chen D. 1993. Visual pattern image sequence coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 3, 291-301.

Van Diepen P.M.J. and Wampers M. 1998. Scene exploration with Fourier-filtered peripheral information. *Perception*, 27, 1141-1151.

Warner H.D., Serfoss G.L., and Hubbard D.C. 1993. Effects of area-of-interest display characteristics on visual search performance and head movements in simulated low-level flight. Armstrong Laboratory, Human Resources Directorate, Aircrew Training Division, Williams AFB, AZ, AL-TR-1993-0023.

Weiman, C.F.R. 1990. Video compression via log polar mapping. *SPIE Proceedings: Real Time Image Processing II*, 1295, 266-277.

Visual Field Simulation: Wilson S. Geisler, Jeffrey S. Perry

A. Original



B. Visual field of glaucoma patient

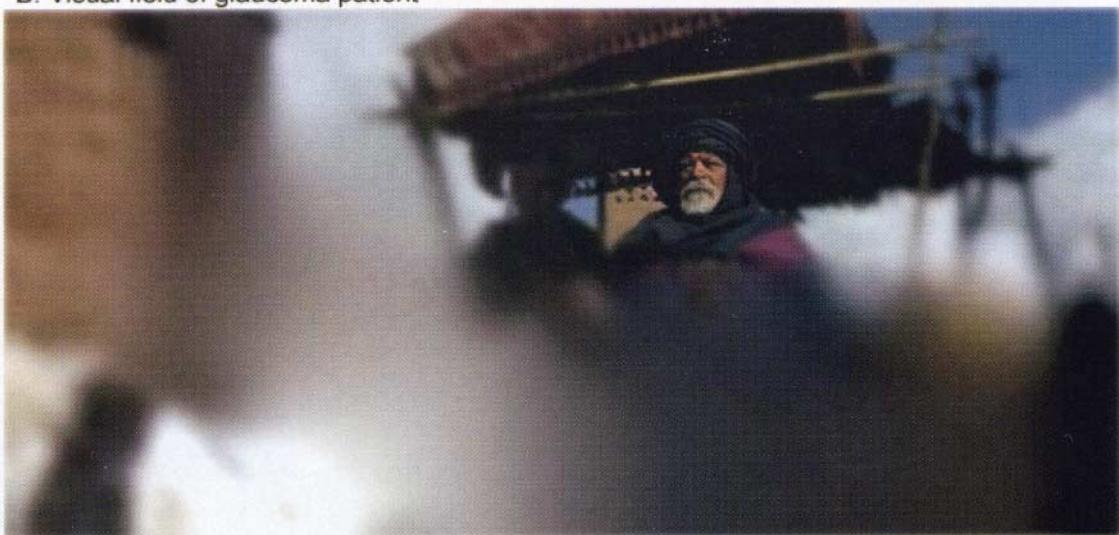


Figure 5. Original and final constructed image simulating the visual field of a glaucoma patient.