Acuity-Matching Resolution Degradation Through Wavelet Coefficient Scaling

Andrew T. Duchowski

Abstract

A wavelet-based multiresolution image representation method is developed matching Human Visual System (HVS) spatial acuity within multiple Regions Of Interest (ROIs). ROIs are maintained at high (original) resolution while peripheral areas are gracefully degraded. Variable resolution images are generated by selectively scaling wavelet (detail) coefficients prior to reconstruction. The technique is equivalent to linear interpolation MIP-mapping which involves smooth subsampling (decomposition) prior to texture mapping (reconstruction). Multiple ROI degradation is achieved through wavelet coefficient scaling following Voronoi partitioning of the image plane.

Keywords

Multiresolution, ROI, Texture Mapping, Wavelets, Image Processing.

EDICS: 1.2, 1.6, 1.8, 1.10

A. Duchowski is with the Department of Computer Science, Clemson University. E-mail: andrewd@cs.clemson.edu .

I. INTRODUCTION

In order to maximize display rates of CPU-intensive applications such as flight simulators, a *gaze-contingent* management scheme may be used to direct resources towards representing a high-fidelity foveal Region Of Interest (ROI), while degrading peripheral detail [1]. Recently, gaze-contingent approaches have been proposed for ROI-based video coding [2], [3], [4], [5]. These schemes concentrate on the representation of the foveal ROI, processing the periphery through smoothing or quantization of transform coefficients. Here, a multiresolution method suitable for gaze-contingent display is introduced, emphasizing graceful peripheral degradation to match the spatial acuity of the Human Visual System (HVS). The technique extends previous work on MIP-mapping to the wavelet domain by appropriately scaling wavelet coefficients prior to reconstruction and allowing multiple ROI representation through Voronoi partitioning.

Selective scaling of wavelet coefficients is not a new approach. A similar method to the one described here was shown by Nguyen et al [4]. In order to enhance relative reconstruction quality, the authors introduced *a priori* weighting factors defining a region-based weighted l^2 metric. The weighting factors were considered as quantitative decimating factors in the relative distortion contributions in each region. Only region-based spatial weighting was considered. The work focused on video encoding, where each frame was synthesized from a fixed subband representation. ROIs were selected according to a motion criterion, where candidates were obtained from a segmentation map which isolated moving objects from the background. To preserve the hierarchy of relevant spatial information in the decimation process, the ROIs were projected onto the subband domain. Uniform threshold quantization was used on wavelet coefficients obtained using Daubechies-4 filters.

The main aspect in which Nguyen et al.'s approach differs from the present technique is the uniform decimation of the ROI coefficients. Preserving coefficients within ROIs and decimating coefficients elsewhere results in abrupt resolution modulation at the reconstructed ROI boundary. Boundaries between levels of resolution in the reconstructed image are clearly visible. This may be suitable for the purposes of compression, but it does not match the spatial sensitivity of the HVS. In contrast, the present goal is to match the abrupt but smooth gradient of the HVS spatial acuity function. Instead of uniform decimation, resolution is linearly interpolated within and between ROI boundaries. The resultant images possess larger Mean Squared Errors (MSE) than those generated by uniform ROI-projection, however, the transition between foveal and peripheral regions is more gradual.

II. Wavelet Interpolation

The Discrete Wavelet Transform (DWT) can be used to reconstruct images at variable resolution by selectively scaling wavelet coefficients. Provided appropriate wavelet filters can be found, reconstruction exactly matches linear MIP-mapping, a well known interpolative texture mapping algorithm used extensively in computer graphics [6]. Identical interpolation results can be generated through the DWT by appropriately scaling wavelet coefficients prior to reconstruction. Wavelet coefficient scaling results in the attenuation of the signal with respect to the average (low-pass) signal. Full decimation of the coefficients (scaling by 0) results in a lossy, subsampled reproduction of the original. Conversely, scaling wavelet coefficients by 1 preserves all detail information producing lossless reconstruction. Selectively scaling the coefficients by a value in the range $[0,1]$, at appropriate levels of the wavelet pyramid, produces a variable resolution image upon reconstruction. This approach is equivalent to MIP-mapping reconstruction with linear interpolation of pixel values [7]. The proof is intuitive since subimages in the MIP-map pyramid correspond to the low-pass subimages recovered at each stage of the wavelet reconstruction. In fact, the low-pass subimages generated at each level of reconstruction are identical to the subsampled images used in MIP-mapping provided both approaches use equivalent decomposition filters and the DWT is guaranteed to be lossless (e.g., orthogonal wavelets are used).

In MIP-mapping, the value of the interpolant p is determined by an arbitrary mapping function which specifies the desired resolution level l. The two closest pyramid resolution levels are then determined by rounding down and up to find subimage levels $j-1$ and j. The interpolant value is obtained by the relation: $p = l - \lfloor l \rfloor$. Note that the slope of the mapping function should match the resolution hierarchy of the pyramid, i.e., if resolution decreases eccentrically from some reference point, the parameter l should also decrease eccentrically. If it does not, its value may be reversed by subtracting from the number of resolution levels, i.e., $n-l$. To scale wavelet coefficients, p is set to either 0, 1, or the interpolant value $l - \lfloor l \rfloor$ at particular subbands. Note that the expression for the interpolant is analogous to the scale factor used when filtering minified image segments in image warping [8].

III. Resolution Mapping

The ROI-based reconstruction of the image from its wavelet transformation relies on the choice of a mapping function. The mapping function determines the degree of peripheral resolution degradation prior to reconstruction of the image, and is thus crucial to the final appearance of the image. The mapping function, denoted by l, maps resolution from the multiresolution pyramid to *image space*. It is

important to note that resolution information in the pyramid is distributed nonlinearly (by decreasing powers of 2 if the multiresolution pyramid is dyadic in nature). Since reconstruction is carried out in image space (dependent on the pixel location (x,y) in the final image), the resultant percent resolution distribution is obtained by taking the inverse of the constant 2 raised to the mapping function, i.e., $\%\,\text{resolution}/100 = 1/2^l$.[1] In the current implementation, three mapping functions are utilized: linear, nonlinear, and empirical HVS acuity-matching. The linear and nonlinear mapping functions were chosen as approximate respective lower and upper bounds to the HVS matching function, in terms of percent resolution. Each mapping function segments the image into concentric resolution regions, or bands. The dimension of the central resolution region is defined by the image area subtended by the fovea (assumed to be 5° visual angle) at the given viewing distance. In all three implementations, resolution within the central 5° of each ROI is consistent and equal. Although this is not a restriction imposed by the image reconstruction method, the size of each ROI is maintained consistently across mapping functions so that different peripheral degradation methods could be readily compared. The three mapping functions are: (1) linear, $l = d/R$; (2) nonlinear, $l = A(1 - e^{-\lambda(d/R)})$; and (3) HVS acuity-matching, $l = -(\ln(\text{empirical \% resolution at pixel distance}/100))/(\ln(2))$. The parameter d is the pixel distance from the ROI center, and R is the radius of the highest resolution region (foveal region). The derivation of R is based on an empirical HVS acuity function (see below). For the nonlinear mapping function, A is the asymptote approximated at the image boundary (here A = 2.35). To consistently preserve resolution within the radius of the highest resolution region, λ is chosen so that $l = 1$ at pixel distance R. That is, $1 = A(1 - e^{-\lambda})$, giving $\lambda = \ln(\frac{A}{A-1})$. The HVS acuity-matching mapping is derived from empirical MAR (minimum angle of resolution) data [9]. MAR data at the border of the projected foveal ROI (at 5° visual angle) is converted to expected maximum resolution in dots per inch (dpi). Expected resolutions at peripheral eccentricities are derived relative to this maximum. Depending on the viewing distance and the resolution of the display device, relative resolvability values in dots per inch are then converted back to pixel units to give the diameters of resolution bands. Assuming a screen display resolution of 50dpi and a viewing distance of 60cm, the pixel diameters used to specify the (piecewise linear) resolution mapping function at eccentricities are given in Table I. Concentric resolution bands representing the resolution mappings in image space are shown in Figure 1 with 2 ROIs.[2] Lighter areas are reconstructed at higher resolution, black rings are level boundaries.

[1] Percent resolution refers to relative resolution in the reconstructed image assuming 100% resolution in the original.

[2] To exaggerate the spatial distribution effect for presentation in the text, Figure 1 uses $R = 105$.

IV. Multiple ROI Image Segmentation

To include multiple ROIs within the reconstructed image, the image is partitioned into multiple regions. Image filtering is performed on a per-pixel basis, where the desired resolution at each pixel location is determined by the mapping function, relative to the center of a particular ROI. To select the appropriate ROI, each pixel is subjected to a membership test. This test involves measuring the distance from the pixel location to each ROI center. Using the Euclidian distance metric, the resolution level of the pixel is determined by the mapping function with respect to the closest ROI center.

Formally, the set $S = \{p_1, \ldots, p_n\}$ of n points in the plane, corresponding to ROI centers, defines a partition of the plane into n regions $V_1, \ldots, V_n$ such that any pixel in the region V_i is closer to the point p_i than to any other $p_j \in S$. This definition of the planar partitioning specifies the Voronoi diagram where each V_i is the convex Voronoi polygon of the point p_i in S [10]. An example of the Voronoi diagram is shown in Figure 2(a). A graphic representation of wavelet coefficient scaling of an arbitrary image at two resolution levels is shown in Figure 2(b). White regions represents coefficients scaled by constant 1, black regions represent coefficient decimation (scaling by 0), and intermediate gray regions are scaled by linearly interpolated values in the interval $(0,1)$. Note that the boundaries between linearly interpolated regions, i.e., boundaries between ROIs, are by construction Voronoi edges.

V. Results

Examples of the variable resolution wavelet scaling technique are shown in Figure 3. The *cnn* image was processed with an artificially placed ROI over the anchor's right eye and another over the "timebox" found in the bottom right corner of the image. Haar wavelets were used to accentuate the visibility of resolution bands. Figure 3(b), (d), and (f) show the extent of wavelet coefficient scaling in frequency space. Notice the different distribution spread of the concentric resolution bands under each mapping. The linearly mapped resolution bands are brought together to generate sharp degradation with respect to ROI centers. Nonlinear mapping spreads out the resolution bands to generate gradual degradation. Reconstructed images are shown in Figure 3(a), (c), and (e).

VI. Acknowledgments

This research was supported in part by the National Science Foundation, under Infrastructure Grant CDA-9115123 and CISE Research Instrumentation Grant CDA-9422123, and by the Texas Advanced Technology Program under Grant 999903-124.

REFERENCES

[1] Thomas Longridge, Mel Thomas, Andrew Fernie, Terry Williams, and Paul Wetzel, "Design of an Eye Slaved Area of Interest System for the Simulator Complexity Testbed," in *Interservice/Industry Training Systems Conference*. National Security Industrial Association, 1989, pp. 275–283.

[2] P. Kortum and W. S. Geisler, "Implementation of a foveated image coding system for bandwidth reduction of video images," in *Human Vision and Electronic Imaging*, Bellingham, WA, January 1996, SPIE.

[3] N. Tsumura, C. Endo, H. Haneishi, and Y. Miyake, "Image compression and decompression based on gazing area," in *Human Vision and Electronic Imaging*, Bellingham, WA, January 1996, SPIE.

[4] E. Nguyen, C. Labit, and J-M. Odobez, "A *ROI* Approach for Hybrid Image Sequence Coding," in *International Conference on Image Processing (ICIP)'94*. IEEE, Nov. 1994, pp. 245–249.

[5] Lew B. Stelmach and Wa James Tam, "Processing Image Sequences Based on Eye Movements," in *Conference on Human Vision, Visual Processing, and Digital Display V*, San Jose, CA, February 8-10 1994, SPIE, pp. 90–98.

[6] Lance Williams, "Pyramidal Parametrics," *Computer Graphics*, vol. 17, no. 3, pp. 1–11, July 1983.

[7] Andrew T. Duchowski, "Representing Multiple Regions Of Interest with Wavelets," in *Visual Communications and Image Processing'98 (VCIP)*, Bellingham, WA, January 1998, SPIE.

[8] George Wolberg, *Digital Image Warping*, IEEE Computer Society Press, Washington, DC, second edition, 1990.

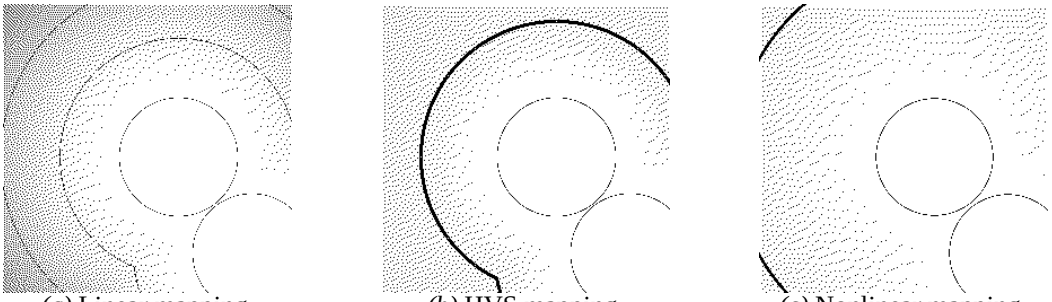
[9] David H. Foster, Salvatore Gravano, and Antonia Tomoszek, "Acuity for Fine-Grain Motion and For Two-Dot Spacing as a Function of Retinal Eccentricity: Differences in Specialization of the Central and Peripheral Retina," *Vision Research*, vol. 29, no. 8, pp. 1017–1031, 1989.

[10] Franco P. Preparata and Michael Ian Shamos, *Computational Geometry: An Introduction*, Springer-Verlag, New York, NY, 1985.

TABLE I

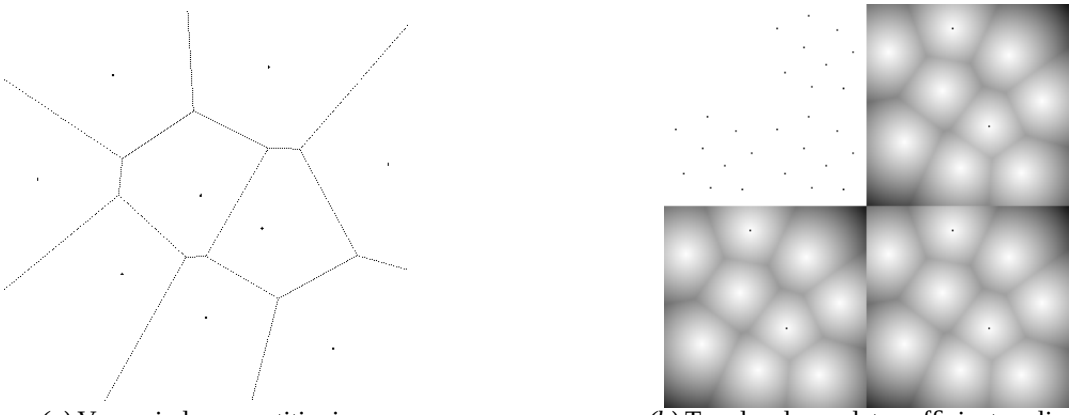
RESOLUTION LEVELS (IN PIXEL DIAMETERS).

	Eccentricity (relative resolution)					
Screen resolution	0-5° (100%)	5° (50%)	10° (26%)	15° (23%)	20° (20%)	25° (17%)
50dpi	–	105	205	310	415	525



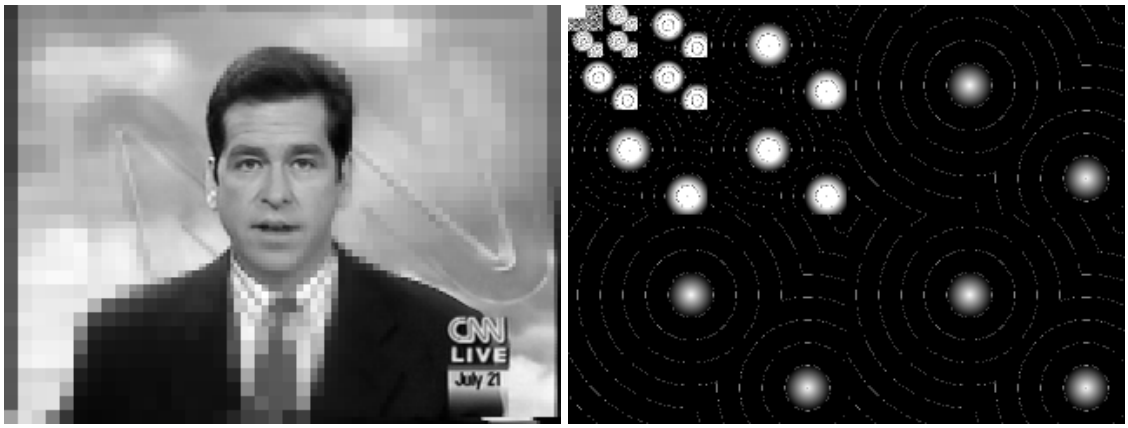
(a) Linear mapping. *(b)* HVS mapping. *(c)* Nonlinear mapping.

Fig. 1. Resolution bands in image space (assuming 100dpi screen resolution).

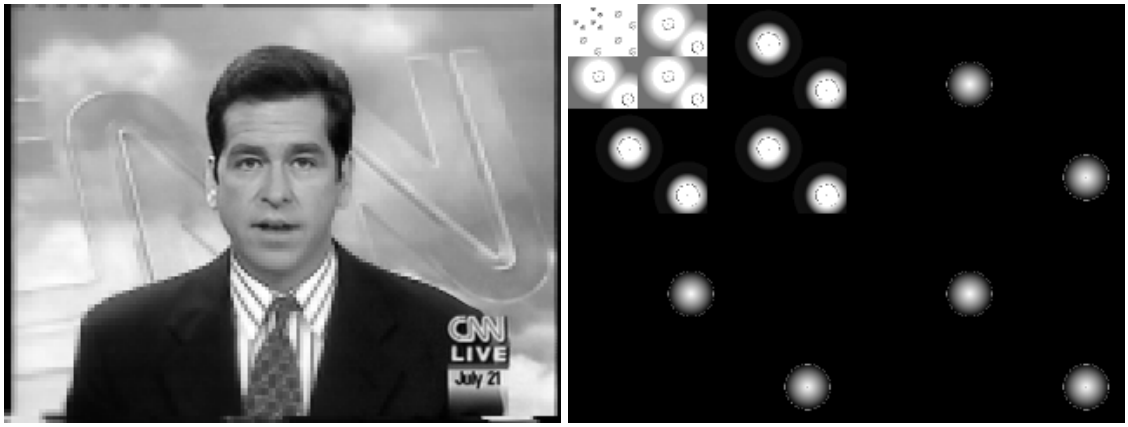


(a) Voronoi planar partitioning. *(b)* Two-level wavelet coefficient scaling.

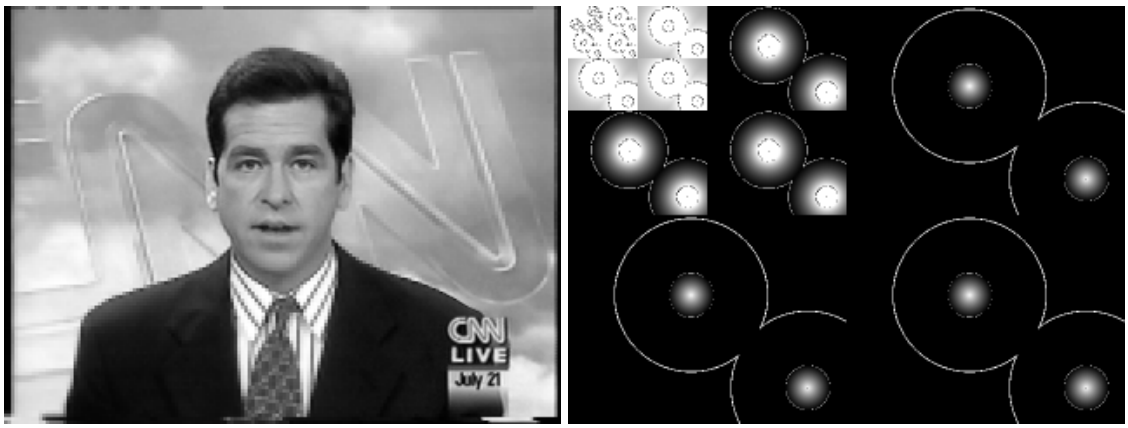
Fig. 2. Example of Voronoi partitioning.



(a) Haar linear mapping. *(b)* Linear mapping.



(c) Haar HVS mapping. *(d)* HVS mapping.



(e) Haar nonlinear mapping. *(f)* Nonlinear mapping.

Fig. 3. Image reconstruction and wavelet coefficient resolution mapping (assuming 50dpi screen resolution).