A SAMPLING APPROACH TO REGION-SELECTIVE IMAGE COMPRESSION

Shahrnaz Azizi†, *Douglas Cochran*†, *and John N. McDonald*‡

† Department of Electrical Engineering
‡ Department of Mathematics
Arizona State University
Tempe, AZ 85287-7206, USA

e-mail: azizi@asu.edu; cochran@asu.edu; mcdonald@math.la.asu.edu

ABSTRACT

This paper proposes a region-selective image compression method that is motivated by the nonuniform sampling structure of the human eye. The approach is based on the nonuniform sampling theorem of Clark for time-warped bandlimited functions. The theoretical underpinnings of the method are explained and examples are presented to illustrate its implementation and performance.

1. INTRODUCTION

In the human visual system, the image formed on the retina is sampled nonuniformly by the photoreceptors (cones), with high sampling density in the center of the visual field (foveal region) and a decreasing density outward. Based on this observation, it is natural to consider reconstruction of images from nonuniform samples taken at points whose distribution matchs the distribution of photoreceptors in the retina. This could have applications in, for example, image/video compression where emerging techniques attempt to exploit biological and psychological perception models. This paper proposes a new method for region-selective image compression that is based on the use of the nonuniform sampling theorem of J.J. Clark [1]. This method provides a means to model the retinal sampling pattern as a two-dimensional warping of a regular sampling lattice. Through use of Clark's theorem, this would enable reconstruction of images from nonuniform samples that match the structure of the retina.

Exploiting the foveated nature of the human eye for data compression and image processing has already received attention by signal processing researchers. In "foveated imaging," static or video images are created and displayed with variable resolution across the image with high resolution given to the regions of interest.

The primary value of foveated imaging is in image compression where high-resolution information is transmitted only in the regions of interest. In [2, 3], a multiresolution foveated coder and a multiresolution pyramid method for creating variable resolution displays in real time were developed. The foveated property of the human eye has also been used in "retinal coding" where an image is represented such that the resolution across it is matched with the structure of the human retina.

A major problem of retinal coding and decoding algorithms is known to be the aliasing artifacts which occur in the peripheral region after decoding [2, 4]. In [5], a smoothly decaying resolution is used for retinal coding and B-spline reconstruction is used to control the smoothness of retinally reconstructed images. An image filtering process based on the foveated nature of the human eye and radial sampling is presented in [6]. The method presented in this paper follows the multiresolution idea of [2, 3], and radial sampling of [6], but nonuniform sampling issues are handled with a more natural interpolation method.

In the following, an explanation of the nonuniform sampling theorem of Clark for time-warped bandlimited functions is given. The proposed region-selective image compression method, as well as the circular spatial warping algorithm are explained. Some examples that demonstrate the properties of the proposed method are also presented.

0-7803-6514-3/00/$10.00©2000 IEEE

2. NONUNIFORM SAMPLING THEOREM OF CLARK

Clark et al. [1] introduced a sampling theorem for functions on the real line that are formed by composition of bandlimited functions with continuous invertible real-valued functions, also called time-warped bandlimited functions. Mathematically, a time-warped bandlimited function $h : \mathbb{R} \to \mathbb{C}$ is defined by $h = f \circ \gamma$ where f is Ω-bandlimited (its Fourier transform $\hat{f}$ is supported in $[-\Omega, \Omega]$) and $\gamma : \mathbb{R} \to \mathbb{R}$. Since γ is invertible, the time-warped bandlimited function h can be reconstructed from samples $\{h(t_n) : n \in \mathbb{Z}\}$ where $t_n = \gamma^{-1}(nT)$ and $T = \pi/\Omega$ is the Nyquist interval for f. This follows immediately from the Shannon sampling theorem: $h(t_n) = f\left(\gamma\left(\gamma^{-1}(nT)\right)\right) = f(nT)$ are Nyquist samples of f so that

$$f(t) = \sum_{n \in \mathbf{Z}} h(t_n)\phi_n(t)$$

with $\phi_n(t) = \text{sinc}\left[T^{-1}(t - nT)\right]$, where

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

Hence

$$h(t) = f(\gamma(t)) = \sum_{n \in \mathbf{Z}} h(t_n)[\phi_n \circ \gamma](t) \qquad (1)$$

If γ is affine (i.e., $\gamma(t) = at+b$ with $a \neq 0$), h is bandlimited and the samples $h(t_n)$ are uniformly spaced. If γ is not affine, h will generally not be bandlimited [7, 8] and the sampling times t_n will generally not be uniformly spaced. The reconstruction is exact if f is continuous. Thus, Clark's observation provides a means for reconstructing certain non-bandlimited functions from nonuniformly spaced samples.

Also, it has been observed in [7] that for any signal space that admits a reconstruction formula of the form

$$f(t) = \sum_{n} f(t_n)\phi_n(t) \qquad (2)$$

there is a sampling theorem with reconstruction formula

$$h(t) = \sum_{n} h(\gamma^{-1}(t_n))\phi_n(\gamma(t)) \qquad (3)$$

for time-warped signals. Thus Clark's basic idea applies to time-warped signal spaces in addition to time-warped bandlimited functions.

Clark's original paper [1] considered the two-dimensional setting and later work of Zeevi and Shlomot [9] developed the two-dimensional perspective further. It is well known (see, for example [10]) that if $f \in L^2(\mathbb{R}^2)$ has compactly supported Fourier transform, it admits a reconstruction formula from certain "regular" sample sets corresponding to the points of a uniform lattice in $\mathbb{R}^2$. In particular, if $\hat{f}(\underline{\omega}) \equiv 0$ outside a bounded region $R \subset \mathbb{R}^2$, and $f(\underline{x}) \in L^2(\mathbb{R}^2)$, then there exist nonsingular matrices $\mathbf{U}$ such that the region of support of $\hat{f}(\underline{\omega} - \mathbf{U}\underline{m})$ does not overlap with the region of support $\hat{f}(\underline{\omega} - \mathbf{U}\underline{n})$ $\forall \underline{n} \neq \underline{m}$ $(\underline{m}, \underline{n} \in \mathbb{Z}^2)$. Under such circumstances, $f(\underline{x})$ can be reconstructed from its samples on the uniform grid $\underline{x}_{\underline{n}} = \mathbf{V}\underline{n}$, where the 2×2 nonsingular sampling matrix $\mathbf{V}$ satisfies $\mathbf{U}^T\mathbf{V} = 2\pi\mathbf{I}$ ($\mathbf{I}$ denotes the 2×2 identity matrix) for any $\mathbf{U}$ with the properties just described. The interpolation formula is

$$f(\underline{x}) = |\det\ \mathbf{V}| \cdot \sum_{\underline{n}} f(\underline{x}_{\underline{n}}) \cdot \Phi(\underline{x} - \mathbf{V}\underline{n}) \qquad (4)$$

where Φ is the inverse Fourier transform of the function $\hat{\Phi}$ defined by

$$\hat{\Phi}(\underline{\omega}) = \begin{cases} 1 & \underline{\omega} \in R \\ 0 & \text{elsewhere} \end{cases}$$

Let $\gamma : \mathbb{R}^2 \to \mathbb{R}^2$ be continuous and invertible, the *spatially-warped function* h is of the form $h = f \circ \gamma$. Defining $\underline{y}_{\underline{n}} = \gamma^{-1}(\mathbf{V}\underline{n})$ yields $h(\underline{y}_{\underline{n}}) = f\left(\gamma(\underline{y}_{\underline{n}})\right) = f(\mathbf{V}\underline{n})$ so that formula (4) gives

$$h(\underline{x}) = f(\gamma(\underline{x})) = |\det\ \mathbf{V}| \cdot \sum_{\underline{n}} f(\mathbf{V}\underline{n}) \cdot \Phi(\gamma(\underline{x}) - \mathbf{V}\underline{n})$$

and hence

$$h(\underline{x}) = |\det\ \mathbf{V}| \cdot \sum_{\underline{n}} h(\underline{y}_{\underline{n}}) \cdot \Phi(\gamma(\underline{x}) - \mathbf{V}\underline{n}) \qquad (5)$$

As with the one-dimensional case, the two-dimensional Clark's formula (5) generalizes the Shannon sampling theorem.

3. REGION-SELECTIVE IMAGE COMPRESSION

Given an image, first a region of it is selected to be the region of interest or "central region." In practice, this should be the region in which the the viewer's gaze is most likely to be centered and hence benefitting most from high quality of reconstruction. The algorithm provides a reconstruction in which the resolution of the reconstructed image decreases outward from the central region. To achieve this, the original image is downsampled nonuniformly with the highest density of sample points in the central region. Specifically, a radial downsampling of the original sampling lattice is done where the original image is divided into several circular regions each of which is downsampled to a different sample density. In the central region, all of the samples of the original image are retained. The second region is a ring around the central region. In this ring, only a fraction of the original samples are retained. In the third region, even a smaller fraction of the original samples are retained. This process can be done for more rings and circular regions until the desired nonuniform sampling pattern is obtained.

The idea is to warp the desired nonuniform sampling pattern to a uniform lattice centered at the central point of the central region (denoted by p_c). By doing so, the uniform lattice will contain the spatially-warped sample points of the original image. Since there are a smaller number of sample points in the warped lattice compared to the original one, the warped image will be a compressed form of the original image. This compression is achieved at the cost of having low quality of the reconstructed image outside of the central region.

Radial downsampling of the original sampling lattice is illustrated in Figure 1. The image is partitioned into a central region surrounded by multiple annular regions. The central region, inside the circle centered at p_c with radius r_1, retains all its original samples. The second region is a ring around the central region, specified by the radii r_1 and r_2. In this ring, only a fraction R_1 of the original samples are retained. In the third region, which is outside of the circle r_2, an even smaller fraction R_2 of the original samples are retained. This process can be done for more rings and circular regions until the desired nonuniform sampling pattern is obtained.

The spatial warping function should map the desired nonuniform sampling pattern to a uniform sampling lattice centered at p_c. It allocates a point in a uniform sampling lattice to each sample point that is taken nonuniformly from the original image and is defined by

$$\gamma(\rho,\theta) = \begin{cases} (\rho,\theta) & \rho \leq r_1 \\ (r_{i-1} + (\rho - r_{i-1}) \times R_{i-1}, \theta) & \\ \qquad r_{i-1} < \rho \leq r_i, & i = 2, 3, \cdots \end{cases}$$

where (ρ,θ) is a point in the polar coordinate centered at p_c and the parameters r_i and R_{i-1} $(i = 1, 2, \cdots)$ are the radii and downsampling ratios of the circular regions in the nonuniform sampling pattern, respectively.

In order to demonstrate this method, some examples are done for which the following circular spatial warping algorithm is used for developing a computer program.

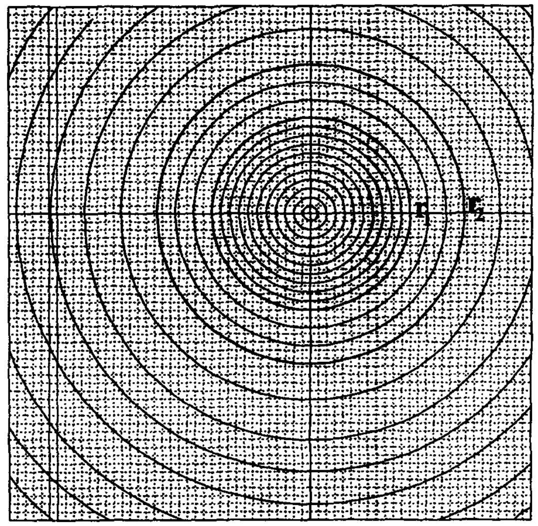

Figure 1: The image is partitioned into a central region (disc) surrounded by multiple annular regions. The central region retains all of its original samples, while each of the annular regions is downsampled to a different sampling density with rings more distant from the central region have lower sampling densities.

3.1. The Circular Spatial Warping Algorithm

The following explains the circular spatial warping algorithm that is used to map the desired nonuniform sampling pattern to a uniform sampling lattice centered around p_c.

Start with the point $(x, y) = (0, 0)$ in the left upper corner of the original image:

1. For (x, y), find out its associated r_i, R_{i-1} by computing the distance between (x, y) and p_c
2. Calculate the polar coordinate (ρ, θ) of (x, y) considering p_c as the origin
3. Calculate the Cartesian coordinate (x_n, y_n) of $(r_{i-1} + (\rho - r_{i-1}) \times R_{i-1}, \theta)$
4. Quantize (x_n, y_n) to obtain the point (x_q, y_q) in the uniform sampling lattice centered around p_c
5. If (x_q, y_q) is not previously assigned, assign it to (x, y)
6. Repeat the above steps for all the points in the original image

The above steps cause the given image to be encoded to a warped image. The decoding procedure would be the reconstruction of the original image from this warped image. Since the above algorithm should be run to generate a lookup table that contains the warping information, the parameters p_c, r_i, and R_{i-1} $(i = 1, 2, \cdots)$ are needed to be known for the decoding procedure. This lookup table is then used to unwarp/interpolate the warped image. Note that because of the quantization step 4, unwarping of the above algorithm would not be possible if step 5 were removed.

4. EXAMPLES

This section presents some examples that illustrate the proposed region-selective image compression method. The main purpose of doing these examples is to provide a preliminary illustration of the method's performance. An ideal model for warping function would be the one that is matched to the structure of the retina. Here, a spatial warping function such as the one explained above is used. Therefore future studies to find a better model for the warping function are needed.

The original image in Figure 2 is sampled first with a foveal pattern centered on the man's face ($R_1 = 2/3, R_2 = 1/2$) yielding the decimated image in Figure 3. The reconstructed image in Figure 4 shows good fidelity on the man's face, but some visible distortion away from this selected region. Another foveal pattern centered on the woman's face ($R_1 = 1/2, R_2 = 1/4$) yields the decimated image in Figure 5. In this case, decimation outside the foveal region is more severe and the resulting distortion outside the foveal region in the reconstructed image in Figure 6 is substantially more visible.

5. CONCLUSION

This paper has proposed a new method for region-selective image compression that is based on the use of the nonuniform sampling theorem of J.J. Clark. This method is motivated by the foveated property of the human eye and it provides a means for reconstruction of an image from a nonuniform sampling pattern that matches the structure of the retina. Some examples were presented as a preliminary demonstration of the proposed region-selective image compression method.

Implementation of the proposed method in practice requires future performance and complexity studies, as well as comparison study with the other existing methods. In particular, combining this method with the existing standard image encoder/decoder structures needs

Figure 2: Original image.

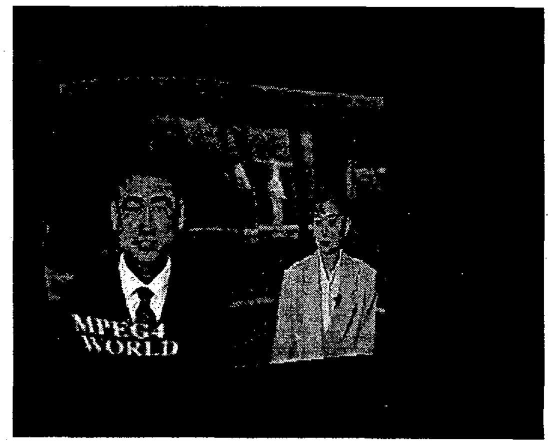

Figure 3: Image after foveal sampling centered in the man's face ($R_1 = 2/3, R_2 = 1/2$).

Figure 4: The resulting reconstructed image for Figure 3.

Figure 5: Image after foveal sampling centered in the woman's face ($R_1 = 1/2, R_2 = 1/4$).

Figure 6: The resulting reconstructed image for Figure 5. Note the more aggressive decimation in this case and the resulting higher distortion outside the foveal region.

to be investigated. Also, additional investigation is needed to obtain a better model for the warping function that would be matched to the structure of the retina.

6. REFERENCES

[1] J.J. Clark, M.R. Palmer, and P.D. Lawrence, "A transformation method for the reconstruction of functions from nonuniformly spaced samples," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-33, pp. 1151–1165, October 1985.

[2] W.S. Geisler and J.S. Perry, "A real-time foveated multiresolution system for low-bandwidth video communication," in B. Rogowitz and J. Allebach (Eds.) *Human Vision and Electronic Imaging*, vol. 3299, pp. 294–305, Proceedings of the SPIE International Society for Optical Engineering, 1998.

[3] W.S. Geisler and J.S. Perry, "Variable-resolution displays for visual communication and simulation," *The Society for Information Display*, vol. 30, pp. 420–423, 1999.

[4] P.T. Kortum and W.S. Geisler, "Implementation of a foveated image coding system for bandwidth reduction of video images," *In B. Rogowitz and J. Allebach (Eds.) Human Vision and Electronic Imaging.* 1996, vol. 2657, pp. 350–360, Proceedings of the SPIE International Society for Optical Engineering.

[5] T. Kuyel, W.S. Geisler, and J. Ghosh, "Retinally reconstructed images: Digital images having a resolution match with the human eye," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 29, no. 2, pp. 235–243, March 1999.

[6] F. Robert and E. Dinet, "An image filtering process based on foveal mechanism simulation," *Proceedings of 31th Asilomar Conference on Signals, Systems, and Computers*, pp. 1725–1729, Monterey, CA, 1997.

[7] D. Cochran and J.J. Clark, "On the sampling and reconstruction of time warped band-limited signals," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Albuquerque, USA, April 1990.

[8] S. Azizi, D. Cochran, and J.N. McDonald, "On the preservation of bandlimitedness under non-affine time warping," *Proceedings of the 1999 International Workshop on Sampling Theory and Applications*, Loen, Norway, August 1999.

[9] Y.Y. Zeevi and E. Shlomot, "Nonuniform sampling and antialiasing in image representation," *IEEE Transactions on Signal Processing*, vol. SP-41(3), pp. 1223–1236, March 1993.

[10] D.E. Dudgeon and R.M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1984.