Foveated Image Formation through Compressive Sensing

Ronald Larcom and Thayne R. Coffman

21st Century Technologies
Austin, TX, USA
e-mail: rlarcom@21technologies.com

Abstract— **We describe two methods by which foveated (variable resolution) images can be created using the techniques of compressive sensing (CS). Foveated sampling (FS) combines a linear shift-variant foveation filter with the CS measurement operator. Foveated sampling and reconstruction (FSR) combines the foveation filter with the CS measurement operator and also with the sparse signal estimation algorithm used to reconstruct images. Both methods are shown to provide accurate reconstruction of foveated images at much higher compression levels than uniform resolution CS.**

Keywords-foveation; compressed sensing; Stagewise Orthogonal Matching Pursuit; image formation; human visual system

I. INTRODUCTION

Compression allows digital images to be used in a wide variety of applications by significantly reducing the space required for storage and bandwidth required for transmission. The recently introduced field of compressive sensing (CS) is a technique which not only reduces transmission bandwidth, but also measurement bandwidth [1]. CS reduces the number of measurements required to represent a discrete signal by exploiting an underlying sparse basis. CS can be applied to a wide variety of signal domains. The remainder of this paper focuses on its application to digital imaging [2].

The wavelet domain is a sparse basis for natural images [3]. Digital images of natural scenes can be represented to high fidelity with many fewer wavelet coefficients than the total number of pixels. CS can exploit this sparsity structure to enable a single pixel camera [4], which allows a single exotic photodetector to generate images at wavelengths outside the capabilities of focal plane arrays. In addition, CS sensors can transmit or store compressed images without the use of a compression algorithm after image acquisition. This can reduce the computational power required by the sensor. Image reconstruction is computationally intensive, but is deferred to a remote device with more powerful capabilities.

Many biological imaging systems use non-uniform bandwidth allocation to address bandwidth challenges. In the Human Visual System (HVS) photoreceptors in the eye are densest about the fovea [5]. Density decreases rapidly with increasing eccentricity. This creates images with space-variant spatial resolution, known as foveated images. Foveated vision provides clear benefits. High visual acuity in the fovea allows the performance of complex tasks and a wide field of view enhances situational awareness, all while bandwidth and processing requirements remain tractable.

The non-uniform resolution of the HVS stands in contrast to the uniform resolution of nearly all digital imaging systems.

This paper presents two strategies for creating digital foveated images by augmenting the techniques of compressive sensing. The approaches realize both the bandwidth reduction benefits of CS and the bandwidth allocation benefits of foveation.

Section II provides background on current approaches to compressive imaging and image foveation. Section III presents two methods for integrating the approaches, named Foveated Sampling (FS) and Foveated Sampling and Reconstruction (FSR). Section IV provides a qualitative comparison of FS, FSR, and a leading compressive imaging approach. Conclusions are given in Section V.

II. BACKGROUND

A. Compressive Sensing

Compressive sensing may be applied to digital imaging to sample and reconstruct a two dimensional image using a sampling rate lower than that suggested by the Nyquist-Shannon sampling theorem. This is achieved by exploiting the knowledge that an image of a natural scene has a sparse representation in the wavelet domain. An image I is represented as

$$I = \Psi x, \qquad (1)$$

where the columns of Ψ are an orthonormal wavelet basis, and x is the wavelet coefficient representation of I. The vector x is sparse. For an N-pixel image I, the M largest values in x contain most of the image energy. The remaining $N - M$ coefficients are zero or nearly zero, and $M \ll N$.

To compressively sense an image, a collection of samples y are generated using a measurement operator Φ that is incoherent with the sparsity basis Ψ, such that

$$y = \Phi x. \qquad (2)$$

A variety of techniques for generating Φ have been suggested [6]. This work uses a random matrix taken from the uniform spherical ensemble [1][7].

The sample vector y is length n, with $M < n \ll N$. Because the image being sensed does not occur naturally in the wavelet domain, the linear sensing operator includes the wavelet transform Ψ^*:

$$y = \Phi\Psi^*I. \tag{3}$$

The wavelet representation x is estimated from the sample vector y. Because $n << N$, (2) is underdetermined to recover x. Compressive sensing overcomes this limitation by exploiting the knowledge that x is sparse. One reconstruction approach is to find the solution to (2) with minimal ℓ_1 norm; optimizing the ℓ_1 norm enforces sparsity in the solution [1].

Stagewise Orthogonal Matching Pursuit (StOMP) is a significantly faster reconstruction approach [7]. StOMP uses a Greedy algorithm to estimate x by iteratively building an estimate of the location of its nonzero coefficients. New candidate coefficients are selected based on the response of the current residual error with a matched filter, Φ^T. A refined solution is estimated from a least-squares solution allowing only the selected coefficients to vary. The new solution is used to produce the error residual for the next iteration.

B. Foveation

Digital foveated imagery provides high resolution at points of interest and lower resolution in the periphery.

There are several ways to create digital foveated imagery. Some sensors can dynamically change pixel size directly on the focal plane array hardware to create 'super-pixels' in peripheral regions [8]. Foveation can also be achieved with specialized compression algorithms [9]. Those algorithms use wavelet coefficient quantization to vary the quantization of coefficients based on their perceptual value. Foveation filtering applies a bank of low-pass filters to the image and selects the value at each pixel according to the desired acuity [10]. This is the equivalent of using a linear shift-variant filter, and the resulting image can be more efficiently compressed by traditional compression algorithms [11].

In foveation filtering, the value at each pixel (u,v) is selected based on a desired local low-pass cutoff frequency $\omega_c(u,v)$. Typically, $\omega_c(u,v)$ is 1.0 in the fovea (representing no filtering) and decreases towards the periphery. Even though it is shift variant, foveation filtering can be represented as a matrix operation, F, because it is linear

$$I_f = FI, \tag{4}$$

$$F = \begin{bmatrix} L_{\omega_c(0,0)}(0,0) \\ L_{\omega_c(0,1)}(0,1) \\ ... \end{bmatrix}. \tag{5}$$

In (5), each row is a vector representation of a 2-D finite impulse response filter kernel (L) with cutoff frequency $\omega_c(u,v)$, padded and shifted to center it at pixel (u,v).

III. FOVEATED COMPRESSIVE SENSING

A. Foveated Sampling

Foveated sampling (FS) modifies the CS sampling process by incorporating foveation. During sampling, a linear operator is applied to the image occurring at the sensor focal plane. This operator ($\Phi\Psi^*$ in (3)), is the combination of a random projection matrix and the wavelet transform. The new sampling equation becomes

$$y = \Phi\Psi^*FI, \tag{6}$$

where F is the foveation matrix from (5).

The samples generated by (6) are used to reconstruct the sparse wavelet coefficients x_f of a foveated-filtered version I_F of the original image I, using a traditional CS reconstruction algorithm. The results below used StOMP reconstruction due to its demonstrated success and computational efficiency.

If the linear sampling matrix $\Phi\Psi^*F$ is pre-computed, FS places no additional computational requirements on the sensor. The benefit of FS is that foveation filtering increases image sparsity in the wavelet domain by eliminating unimportant or visually redundant image information in the periphery.

CS reconstruction error at a given sampling rate is directly related to underlying signal sparsity. Compared to CS reconstruction of I, foveated sampling creates a simpler reconstruction problem of finding I_F. Where CS *exploits* known signal sparsity, FS *enforces* a desired signal sparsity to redistribute error to visually unimportant fine wavelet coefficients in the image periphery.

B. Foveated Sampling and Reconstruction

Foveated sampling and reconstruction (FSR) modifies both the sampling and the reconstruction steps in CS.

FSR begins by generating samples with FS according to (6). This biases the measurement operator to provide more detail in the fovea at the expense of fine detail in the periphery. That bias places additional constraints on the structure of the sparse signal x_f being estimated during reconstruction.

FSR augments StOMP (or other algorithms) to exploit that additional structure during reconstruction. The initial estimate $x_{f,0}$ of the sparse wavelet coefficients is set to zero, and the set of active wavelet coefficient locations I_0 is empty. At each iteration s, the residual error r_s is calculated as

$$r_s = y - \Phi x_{f,s-1}. \tag{7}$$

FSR then generates a vector of residual correlations with a matched filter and reweights them using the foveation operator to form the weighted correlation vector

$$c_s = \Psi^*F\Psi\Phi^Tr_s. \tag{8}$$

Weighting is used to bias the reconstruction towards those coefficients least impacted by the foveation operator. Significant residual correlations are identified by testing against a threshold, T, which forms a set of newly selected candidate coefficients J_s. J_s is merged with the set of active wavelet coefficients to create the new support set I_s

$$J_s = \{j : |c_s(j)| > T\}, \tag{9}$$

$$I_s = I_{s-1} \cup J_s. \tag{10}$$

The submatrix Φ_I is formed by selecting the columns of Φ indexed by I. This is used to find the least squares estimate of x_f using only the active coefficient set

$$x_{f,s} = (\Phi_{I_s}^T \Phi_{I_s})^{-1} \Phi_{I_s}^T y . \qquad (11)$$

If StOMP termination conditions are met, the estimated solution x_f is set to the current estimate, $x_{f,s}$. The reconstruction algorithm for FSR only differs from StOMP in (8), which now incorporates the foveation operator.

The wavelet domain foveation operator, $\Psi*F\Psi$, impacts the formation of the active set, biasing it away from coefficients that were suppressed by foveation filtering.

C. Experimental Results

FS and FSR performance was quantified on a set of ten test images representing a variety of subjects and imaging conditions. Results were compared to an unmodified StOMP implementation based on SparseLab (available online at http://sparselab.stanford.edu). All tests used the multiscale CS extension described in [6], which directly samples the 1024 coarsest wavelet coefficients using the "symmlet8" wavelet. In addition, the effects of quantization were considered by rounding all samples to an eight bit range during processing. The test images were originally 512 x 512 pixels, in eight bit grayscale.

The StOMP algorithm allows a variety of threshold selection methods. Its performance is very sensitive to parameter selection when using either the Constant False Discovery Rate (CFDR) or Constant False Alarm Rate (CFAR) rule. Instead, we adopted the simpler heuristic of setting the threshold to a fixed percentage of the peak correlation value. This provided good results for StOMP, FS, and FSR while allowing tests on a diverse set of test images.

Figure 1 shows a comparison of the approaches as applied to an outdoor bridge scene and to the Lena image, all with sampling ratio of 0.1. FS and FSR produce foveated reconstructions of the original image with significantly fewer artifacts than StOMP at this sampling ratio.

The differences between the reconstructions produced by FS and FSR are subtle and are primarily within the fovea. Because of its threshold bias, FSR includes more fine scale wavelet coefficients in its support estimate for the center of the image. This yields crisper edges and finer detail in the fovea, but also increased high frequency noise.

For quantitative comparisons, the test images were reconstructed with each approach at sampling ratios of 0.05, 0.1, 0.25 and 0.5. Reconstruction accuracy was measured with Mean Squared Error (MSE). MSE was measured against a foveation filtered version of the test image for FS and FSR, and against the original image for StOMP.

Results are shown in Figure 2. FS and FSR both outperform StOMP at low sampling rates. They provide more accurate reconstructions of the foveated image than StOMP provides of the full resolution image. FS and FSR achieve comparable reconstruction error, with a slight edge to FS at higher sampling ratios. The error in FSR is primarily

Figure 1.   Reconstruction of bridge and Lena images StOMP (left), FS (right) and FSR (right). Reconstructed 512 x 512 images generated using 25000 eight-bit samples.
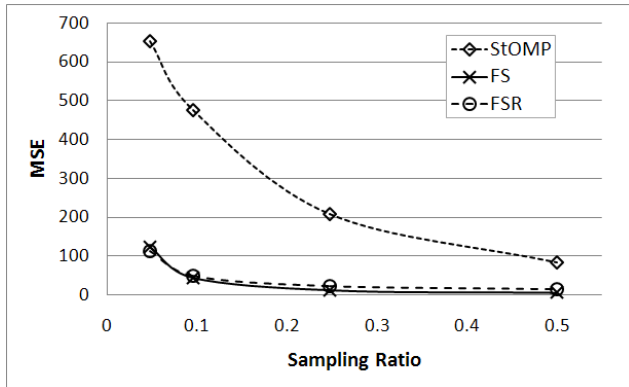
Figure 2. Comparison of FS, FSR, and StOMP error at different sampling ratios.

high frequency noise near the fovea. FSR may outperform FS under a more robust perceptual error index such as SSIM [12], which embodies the noise masking principle, or using a foveation-weighted image quality index [13].

IV. CONCLUSIONS

FS and FSR allow the sampling and reconstruction of foveated images within the framework of compressive sensing. FS and FSR produce high-quality foveated images at a sampling ratio of 0.1. In contrast, most traditional CS imaging approaches require sampling ratios of 0.25 or higher to provide acceptable quality. FS and FSR outperform traditional CS techniques at low sampling ratios.

FS and FSR were evaluated with respect to a foveated reference image, where StOMP was evaluated with respect to the full resolution image. This does not address the distortion introduced by foveation itself (i.e., the lack of detail in the periphery). Our work is therefore most relevant to applications where foveation is appropriate and the fovea is correctly placed on areas of interest within the image.

Many tractable fovea placement strategies exist, making the results and conclusions broadly applicable. One simple strategy (from an algorithmic standpoint) is to fix the fovea and require an operator or external control system to steer the field of view. Alternatively, fovea placement can be selected based on image content by using 'smashed' filters [4] to identify salient points in the compressed foveated domain. Multiple fovea configurations can be incorporated with zero computational cost at runtime by precomputing relevant quantities for alternative foveation operators $F_1 \ldots F_K$.

This work showed how a foveation operator can be incorporated into the StOMP reconstruction algorithm to achieve superior performance at low sampling rates. We are also developing techniques to augment other reconstruction algorithms by exploiting foveation structure. One alternative approach under study exploits the statistics of natural images to modify prior probabilities during reconstruction [14].

ACKNOWLEDGMENT

REFERENCES

[1] D. Donoho, "Compressed sensing," IEEE Trans. Info. Theory, vol. 52, no. 4, pp. 1289-1306, 2006.

[2] J. Romberg, "Imaging via compressive sampling," IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 14-20, 2008.

[3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Processing, vol. 1, no. 2, pp. 205-220, 1992.

[4] M. Duarte, et al., "Single-pixel imaging via compressive sampling," IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 83-91, 2008.

[5] L. Cormack, "Compuational model of early human vision," in Handbook of Image and Video Processing, A. Bovik, Ed. New York: Academic, 2000.

[6] Y. Tsaig and D. Donoho, "Extensions of compressed sensing," Signal Processing, vol. 86, no. 3, pp. 549-571, 2006.

[7] D. Donoho, Y. Tsaig, I. Drori, and J. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," Tech. Rep. 2006-02, Dept. of Statistics, Stanford University, Stanford, CA, 2006. (software available at http://sparselab.stanford.edu/).

[8] P. McCarley, M. Massie, and J. Curzan, "Large format variable spatial acuity superpixel imaging: visible and infrared systems applications". SPIE Proc., Infrared Technology and Applications XXX, vol. 5406, Orlando, FL, USA, 2004.

[9] Z. Wang, and A. Bovik, "Embedded foveation image coding," IEEE Trans. Image Processing, vol. 10, no. 10, pp. 1397-1410, 2001.

[10] S. Lee, A. Bovik, and B. Evans, "Efficient implementation of foveation filtering," Proc. Texas Instruments DSP Educators Conference, Houston, TX, USA, August 1999.

[11] H. Sheikh, S. Liu, B. Evans, and A. Bovik, "Real-time foveation techniques for H.263 video encoding in software," Proc. IEEE ICASSSP, vol. 3, Salt Lake City, UT, USA, May 2001.

[12] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncellie, "Image quality assessment: from error visibility to structural similarity," IEEE Tran. Image Processing, vol. 13, no. 4, pp. 600-612, 2004.

[13] S. Lee, M. Pattichis, and A. Bovik, "Foveated image quality assessment," IEEE Tran. on Multimedia, vol. 4, no. 1, pp. 129-132, 2002.

[14] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," IEEE Trans. Signal Processing, vol. 56, no. 6, pp. 2346-2356, 2008.