MODEL-BASED REGION-OF-INTEREST ESTIMATION FOR ADAPTIVE RESOURCE ALLOCATION IN MULTI-APERTURE IMAGING SYSTEMS

*Indranil Sinharoy**

Magnum Semiconductor
591 Yosemite Drive
Milpitas, California 95035 USA

Scott C. Douglas, Dinesh Rajan, and Marc P. Christensen†

Department of Electrical Engineering
Southern Methodist University
Dallas, Texas 75275 USA

ABSTRACT

Using intelligent resource allocation based on the information content in the imaging system's field-of-view for the successful design of a flat-profile multiplexed optical imaging system requires the use of adaptive techniques. This paper describes a model-based technique for determining regions of interest in aerial images using the 2D normalized power spectral density within Gilles' saliency map estimator. The proposed technique exploits the $1/f^{\alpha}$ spatial spectral shape of such natural imagery in a computationally-simple approach that is robust to additive noise. Application of the method to candidate aerial images shows its ability to identify consistent regions of interest for such data.

Index Terms— Optical imaging, image region analysis, information theory, multisensor systems, spectral analysis

1. INTRODUCTION

Recent work in computational imaging systems has yielded system designs that combine multi-aperature image collection with advanced image reconstruction algorithms in a lightweight flat form factor device [1, 2, 3, 4]. In these systems, multiaperture optical imaging systems built using state-of-the-art micro-optics technology are used for data collection, and multiple low-resolution images are combined using signal processing techniques to produce a final image with both a high angular resolution and a large field-of-view.

The performance of multiplexed imaging systems may be enhanced by optimizing imaging resource utilization through adaptive resource allocation based on the information content of the scene. Regions within a scene devoid of features are allocated fewer imaging resources, allowing more imaging resources to be devoted to regions of higher spatial information for improved spatial resolution. This design has been termed *PANOPTES* (processing arrays of Nyquist limited observations to produce a thin electro-optic sensor) [3], which employs steerable micro-mirror sub-imaging arrays to collect image information for reconstruction. Non-uniform spatial allocation of imaging resources matching the information content of the scene is critical for good performance in such a strategy. Techniques for estimating *saliency*, defined as the degree to which a portion of an image is pre-attentively distinct to the human eye, are therefore required. Regions with high saliency lead to immediate visual attention in the early stages of the human visual system [5].

The design of computational multi-aperture imaging systems require automatic allocation of resources to feature-rich regions within a scene. In [6], a novel approach for generating saliency images or maps was proposed that uses local estimates of the image power spectral density (PSD) to generate the saliency image as

$$J_{ij} = -\sum_{u}\sum_{v} p_{ij}(u,v)\log_2 p_{ij}(u,v). \qquad (1)$$

In this relation, $p_{ij}(u,v)$ is the *spatial frequency histogram* for the (i,j)*th* block that is computed by dividing the PSD by the sum of the powers in all of the frequency-domain bins, where u and v are discrete frequency-bin indices. Experimental results in [6] indicated that (1) tends to identify information-rich portions of the scene more readily and consistently than traditional entropy maps using histograms. The main drawback to (1) is its high computational complexity due to the number of terms and nonlinear $\log$ evaluations per saliency map pixel.

In this paper, we leverage the fact that the mean PSD of natural images follows an $1/f^{\alpha}$ power law distribution to develop a simple approximate procedure for calculating (1). Our procedure uses the relationship between the power-law exponent α and the saliency expression in (1). As the exponent α can be estimated by modeling the PSD of the local region as a power-law distribution, we can obtain a computationally-efficient implementation of (1) with another advantage: immunity to noise can be addressed through weighting of the DC spectral coefficients prior to calculating the saliency map. Examples illustrate the usefulness of the proposed method.

*The first author performed the work while at the Department of Electrical Engineering, Southern Methodist University, Dallas, Texas.

†This work was funded in part by DARPA and the U.S. Army.

Authorized licensed use limited to: SOUTHERN METHODIST UNIV. Downloaded on November 1, 2008 at 16:34 from IEEE Xplore. Restrictions apply.

2. ESTIMATION OF LOCAL SALIENCY

The statistical properties of the power spectral density (PSD) of natural images have been studied extensively in visual science [7, 8, 9, 10], whereby the average PSD $S(f)$ of natural image ensembles follows a power-law distribution in radial spatial frequency, or

$$S(f) = \frac{A}{f^{\alpha}}. \tag{2}$$

In this equation, f is the magnitude of the spatial frequency in cycles/pixel, α is the exponent of the power-law distribution, and A is an image contrast parameter. The PSDs of images containing a greater number of objects and edges tend to be broader than those with fewer features, such that α has a lower value for the former class of images [10]. This behavior of α coupled with the *scale invariance* property of power law distributions allows us to develop direct methods for estimating saliency using parametric modeling of local PSD estimates, where α is an intermediate variable. In what follows, we leverage the discrete-space nature of sampled images in this estimation task.

The power law, also called Zipf's law or the Pareto distribution, can model many natural and man-made phenomena. The power law for a discrete random variable *k* is

$$p_k = \frac{k^{-\alpha}}{\zeta(\alpha)}, \quad k \geq 1, \tag{3}$$

where α is the *exponent* of the power law and $\zeta(\alpha) = \sum_{k=1}^{\infty} k^{-\alpha}$ is the Riemann zeta function. Goldstein *et al.* [11] showed that maximum likelihood estimation (MLE) of α produces a robust and accurate estimate for fitting observed data $x = x_i, 1 \leq i \leq N$ to the power law distribution. The expression for the MLE estimate of α for the Pareto distribution is

$$\frac{\zeta'(\alpha)}{\zeta(\alpha)} = -\frac{1}{N}\sum_{i=1}^{N}\log(x_i) = -\sum_{k=1}^{\infty} p_k \log(k) \tag{4}$$

Since the two expressions on the right of (4) can be computed from data, the value of α can be determined by inverse interpolation using a table of $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$. A table relating $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ and α can be easily pre-computed [11].

Shannon defined the *entropy* $H(X)$ of a discrete random variable X with alphabet $\mathcal{X}$ as

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x) \tag{5}$$

where $p(x)$ is the *probability mass function* (pmf) of the discrete random variable. Then, the entropy associated with a random variable having a discrete power-law distribution is

$$\begin{aligned} H(\alpha) &= -\sum_{k=1}^{\infty} \frac{k^{-\alpha}}{\zeta(\alpha)} \log \frac{k^{-\alpha}}{\zeta(\alpha)} \\ &= \alpha \sum_{k=1}^{\infty} \frac{k^{-\alpha}}{\zeta(\alpha)} \log(k) + \frac{\log(\zeta(\alpha))}{\zeta(\alpha)} \sum_{k=1}^{\infty} k^{-\alpha} \\ &= \alpha E \log(k) + \log(\zeta(\alpha)) \end{aligned} \tag{6}$$

Thus, once the parameter α is estimated using (4), the entropy, $H(\alpha)$ as a function of α can be calculated using (6). When k in (6) represents discrete spatial frequency, p_k is the discrete PSD, and $H(\alpha)$ in (6) becomes the saliency $J(\alpha)$. Thus, once α has been estimated, we can compute $J(\alpha)$ as

$$J(\alpha) = \alpha \sum_{k=1}^{\infty} p_k \log(k) + \log(\zeta(\alpha)). \tag{7}$$

3. ALGORITHM IMPLEMENTATION ISSUES

The local saliency of an image can be calculated by estimating the parameter α over local regions via (4) and then using the estimated α in (7) to calculate the saliency. We now address several practical issues in this process.

The relation between α and the random variable k for a discrete power law distribution is as shown in (4). In the power-law model for the PSD of a natural image, the index k corresponds to the index of the k^{th} discrete frequency bin, and the value k^{α} corresponds to the amplitude within the k^{th} frequency bin. The discrete power-law distribution in (3) is defined for all $1 \leq k < \infty$, whereas the range of p_k available for processing is over $1 \leq k \leq \left(\frac{L}{2} + 1\right)$ which represents the discrete spatial frequency from DC to half the Nyquist frequency where L is the dimension of the data block. We may write (4) as

$$\begin{aligned} \frac{\zeta'(\alpha)}{\zeta(\alpha)} &= -\sum_{k=1}^{\infty} p_k \log(k) \\ &= -\sum_{k=1}^{\left(\frac{L}{2}+1\right)} p_k \log(k) - C(\alpha). \end{aligned} \tag{8}$$

$$\text{where } C(\alpha) = \sum_{\left(\frac{L}{2}+2\right)}^{\infty} \frac{k^{-\alpha}}{\zeta(\alpha)} \log(k).$$

The first term on the R.H.S. of (8) is a sum of the products of $\log(k)$ and the normalized PSD values p_k for $k = 1$ to $\left(\frac{L}{2} + 1\right)$. The normalized PSD p_k is a spatial frequency histogram derived from the two-dimensional local PSD. Using a radial spectral model is appropriate for our aerial imaging applications, as the local PSDs of such images exhibit an isotropic nature; thus, we obtain the one-dimensional PSD by approximate radial averaging of the 2D

Authorized licensed use limited to: SOUTHERN METHODIST UNIV. Downloaded on November 1, 2008 at 16:34 from IEEE Xplore. Restrictions apply.

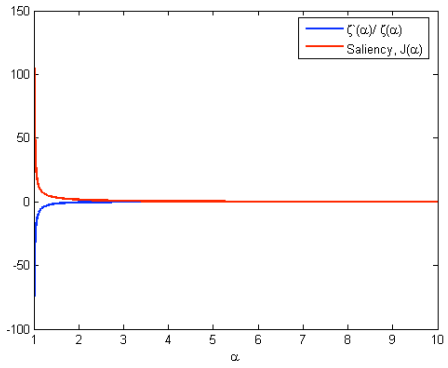

Fig. 1. Plot of the saliency $J(\alpha)$ and $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ versus α.

PSD [4]. The second term $C(\alpha)$ in (8) is dependent on α, and its value is typically much smaller than that of the first term. Thus, we can neglect $C(\alpha)$ in (8). An analysis of the truncation error resulting from neglecting $C(\alpha)$ is presented in [4]. The saliency maps that are produced with this approximation are comparable to the non-model based PSD-saliency maps proposed in [6].

Additional simplifications can be obtained by considering the shape of $J(\alpha)$ as it relates to $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$. A plot of $J(\alpha)$ and $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ versus α is shown in Fig. 1. The plot shows that $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ and $J(\alpha)$ varies almost inversely with α. A plot of $J(\alpha)$ versus $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ is shown in Fig. 2(a). It can be seen that the plot behaves nearly linearly suggesting that the saliency $J(\alpha)$ can be evaluated directly from $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ without calculating α. Furthermore, numerical experiments indicate that the local PSD-saliency values typically lie in the range $[0, 4]$ for the aerial imagery of interest. From Fig. 2(b), we see that this relationship is close to linear, and it is easily approximated using a non-uniform lookup table with fewer density of points for $\frac{\zeta'(\alpha)}{\zeta(\alpha)} \leq -1$. These considerations make the use of this procedure quite practical for real-time implementation.

The calculation of local saliency can be summarized in the following steps:

1. Calculate the 2D FFT to form the PSD vector p_k within each processing block.
2. Calculate $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ by taking the inner product between the p_k and the predefined vector $\log(k)$.
3. Estimate saliency J from a lookup table containing the values of $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ and the corresponding $J(\alpha)$ values.

The complexity of the above approach is now considered in terms of the numbers of multiplications, additions, and table lookups for the method. We exclude the cost of the 2D FFT calculation since it is common between both (1) and the proposed approach. For a data block side of $L \times L$ pixels,

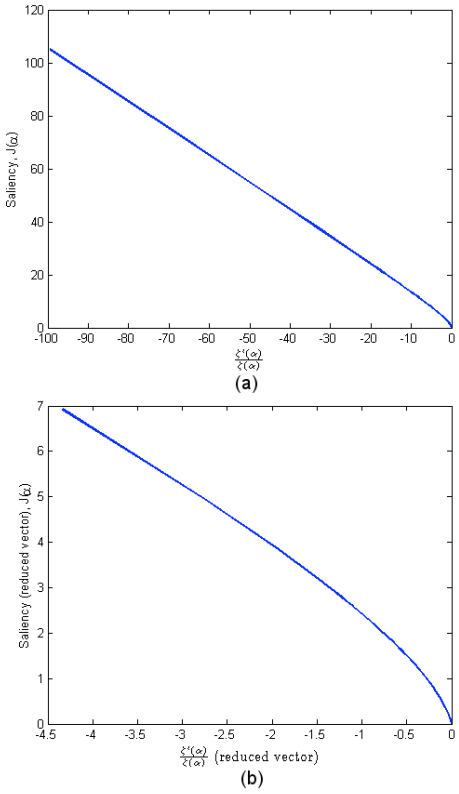

Fig. 2. Plot of $J(\alpha)$ versus $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$. (a) Plot of $J(\alpha)$ versus $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ for $1.01 \leq \alpha \leq 10$; (b) Plot of $J(\alpha)$ versus $\frac{\zeta'(\alpha)}{\zeta(\alpha)}$ for $0 \leq J(\alpha) \leq 7$. .

the method in (1) requires L^2 multiplies, $L^2 - 1$ adds, and L^2 table lookups, whereas the proposed method requires $L + 3$ multiplies, $\frac{L^2}{8}(L+2)+4$ adds, and 2 table lookups. Thus, the proposed method requires fewer multiplies and table lookups at the cost of more additions.

4. NUMERICAL EVALUATIONS

We now explore the performance of the proposed method via its application to aerial imagery. For comparison, we also include the original PSD-based method in (1).

Fig. 3 shows an aerial image of an airport. Both the non-parametric PSD-based method in (1) and the parametric method described in the last section are applied to this image using a (16×16) block size with a 4-pixel offset between blocks. Fig. 4 shows the results of both methods. It can be seen that the parametric PSD-based saliency map is very similar to that of the non-parametric method despite it being computationally-simpler.

Our method for calculating saliency maps does not consider noise effects directly; however, in real-world systems, noise is always present. Comparisons of the method in (1) and the parametric method described in the last section indicate

Authorized licensed use limited to: SOUTHERN METHODIST UNIV. Downloaded on November 1, 2008 at 16:34 from IEEE Xplore. Restrictions apply.

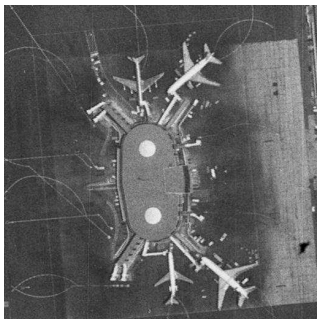

Fig. 3. Aerial image of an airport.

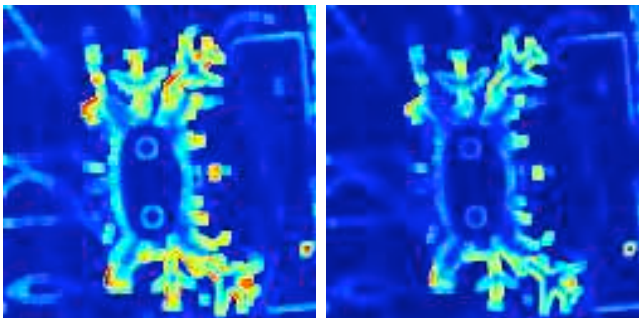

Fig. 4. (Left) Saliency map of the airport image for 16×16 block size and 4-pixel offset generated using (1). (Right) Saliency map using the proposed technique.

that the parametric method is more robust to low amounts of additive noise as compared to the non-parametric PSD method [4]. If the noise variance in the image is known or can be estimated, any performance degradation can alternatively be minimized by subtracting the average noise power in each bin from the PSD vector. This technique is valid under the assumption that the noise is white Gaussian and thus its power is uniformly spread across the entire bandwidth. Such a solution requires *a priori* knowledge.

Upon further study, however, the effects of noise were found to be easily mitigated with the following simple change to the PSD-based methods: *Use a constant DC power for all images that is proportional to the block size* L^2. Typically, the value L^2 for the DC component produced good saliency maps in the presence of noise. Our reasoning for this choice is as follows. The DC term of each block does not carry significant information about the image content, so such a substitution yields no loss of information. This change also leads to identical saliency maps for complemented images (*e.g.* image negatives). In addition, having a constant DC power value also renders the algorithm to be invariant to different lighting conditions. Fig. 5 shows the airport image with additive white Gaussian noise having a variance of 0.05 and the saliency map generated using the parametric PSD-saliency map estimator. As can be seen, the proposed method provides good immunity to noise effects.

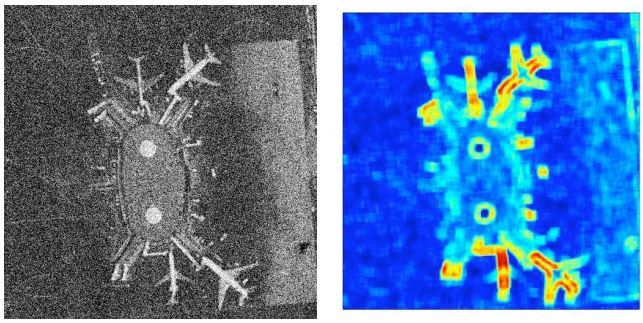

Fig. 5. (left) Image of a airport with a noise variance of 0.05 in normalized scale. (right) Saliency map of the noisy airport image using DC compensation of $8L^2$, for $L \times L$, $L = 16$ block size and a 2-pixel offset.

5. CONCLUSIONS

Performance of multi-aperture computational imaging systems can be improved by using adaptive allocation of imaging resources within a scene. A novel method to automatically detect feature-rich regions within a scene using spacial frequency histograms instead of the conventional intensity histograms was proposed in [6]. The method generated saliency maps of scenes using the local power spectral density within Gilles' saliency map estimator. In this paper, we derive a parametric approach to determine the PSD based saliency maps by modeling the 2D PSD of images as power law distributions. The model-based approach was shown to be computationally much simpler at the same time producing comparable saliency maps as the non-model based methods. Detailed analysis of the parametric method can be found in [4].

6. REFERENCES

[1] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka, "Thin observation module by bound optics (TOMBO): Concept and experimental verification," *Applied Optics Inform. Proc.*, vol. 40, pp. 1806-1813, Apr. 2001.

[2] J. W. Duparré and F. C. Wippermann, "Micro-optical artificial compound eyes," *Bioinspiration & Biomemetics* 1, R1-R16, May 2006.

[3] M. P. Christensen, V. Bhakta, D. Rajan, T. Mirani, S. C. Douglas, S. L. Wood, and M. W. Haney, "Adaptive flat multiresolution multiplexed computational imaging architecture utilizing micromirror arrays to steer subimager fields of view," *Applied Optics*, vol. 45, pp. 2884-2892, May 2006.

[4] I. Sinharoy, "Region-of-interest estimation in multi-aperture imaging systems," Master's thesis, Southern Methodist University, Dallas, Texas, Dec. 2006.

[5] T. Kadir and M. Brady, "Saliency, scale and image description," *Int. J. Comp. Vis.,* vol. 45, pp. 83–105, 2001.

[6] I. Sinharoy and S. C. Douglas, "Region-of-interest estimation for adaptive resource allocation in multi-aperture imaging systems," *IEEE Int. Conf. Acoust., Speech, Signal Processing,* Honolulu, HI, vol. 2, pp. 597-600, Apr. 2007.

[7] A. Torralba and A. Oliva, "Statistics of natural image categories," *Network: Computation in Neural Systems,* vol. 14, pp. 391–412, 2003.

[8] R. M. Balboa and N. M. Grzywacz, "Power spectra and distribution of contrasts of natural images from different habitants," *Vision Research,* vol. 43, pp. 2527–2537, 2003.

[9] D. L. Ruderman, "Origins of scaling in natural images," *Vision Research,* vol. 37, pp. 3385–3398, 1997.

[10] R. -P. Millane, S. Alzaidi and W. -H. Hsizo, "Scaling and power spectra of natural images," *Image and Vision Computing,* pp. 148–153, 2003.

[11] M. -L. Goldstein, S. A. Morris and G. G. Yen, "Problems with fitting to the power-law distribution," *Eur. Phys. J.,* , pp. 255–258, 2004.

Authorized licensed use limited to: SOUTHERN METHODIST UNIV. Downloaded on November 1, 2008 at 16:34 from IEEE Xplore. Restrictions apply.