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

Indranil Sinharoy and Scott C. Douglas

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

ABSTRACT

Successful design of a flat-profile multiplexed optical imaging system requires the use of adaptive techniques to make intelligent resource allocation based on the information content in the imaging system's field-of-view. This paper explores techniques for finding regions of interest in aerial images using local entropy as a descriptor. A novel method for identifying regions-of-interest in images is developed using the 2D normalized power spectral density within Gilles' saliency map estimator. Application of the method to candidate aerial images shows its ability to identify consistent regions of interest for such data for varying block sizes and under additive noise.

Index Terms— computational imaging systems, information theory, saliency, image reconstruction

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. In these systems, multi-aperture optical imaging systems built using state-of-the-art micro-optics technology are used for data collection [1] , 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. Because the systems use small aperture lenses, the depth-of-field of such devices is increased over traditional imaging systems, mitigating the need for adjustible lens focus.

The design of multi-aperture imaging systems requires an allocation strategy for the imaging resources across the field-of-view. For example, the TOMBO (Thin Observation Module by Bound Optics) [2] system design utilizes an array of micro-imaging systems with a fixed overlap of imaging resources across the region-of-interest. This system is optimized for a specific resolution and field-of-view [2],[3]. Spatial information is typically not distributed uniformly within a scene [3], and the TOMBO architecture does not take this non-uniformity of spatial information into account. 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. One such 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 [4].

This paper explores existing and novel methods for quantifying the amount of information in local regions within a scene for imaging resource allocation. Entropy-based methods employing both sample histograms and local power spectral density (PSD) estimates are first described. A novel criterion for local saliency estimation in images that uses local PSD estimates is proposed. Application of the method to typical aerial image data shows its ability to consistently identify similar regions-of-interest within such images for different processing strategies and when noise is present.

2. ENTROPY, SALIENCY, AND REGION-OF-INTEREST ESTIMATION

Methods for identifying image saliency are closely tied to entropy-based image coding techniques. Gilles defines saliency in an image in terms of local signal complexity or unpredictability and presents techniques for maximizing Shannon entropy computed from local intensity histogram patches for aerial image registration [5]. In this paper, we are concerned with identifying regions within a field of view that represent good candidates for increased image resource allocation. Such methods would be used within an adaptive procedure to allocate additional imaging resources within the multi-aperture optical system.

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

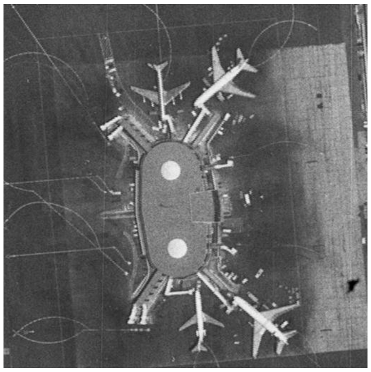


Fig. 1. Aerial image of an airport.

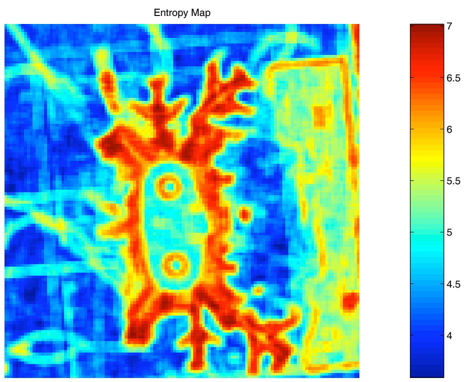


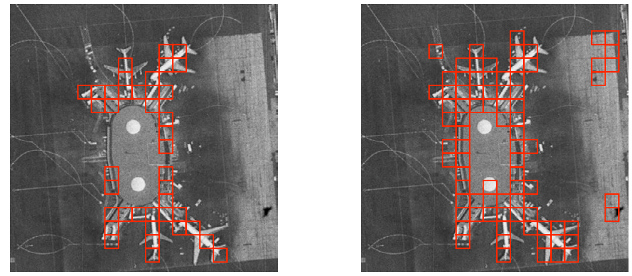
Fig. 2. Local entropy map generated from intensity histograms, 16×16 block size, 4-pixel offset.

A straightforward implementation of Gilles' saliency model to region-of-interest estimation is the entropy map

$$H_{ij} \quad = \quad -\sum_{x \in \mathcal{X}} p_{ij}(x) \log_2 p_{ij}(x) \tag{1}$$

where $p_{ij}(x)$ is the normalized intensity histogram taken in a region about pixel (i, j) in the image. As an example, an aerial image of an airport is shown in Fig. 1. The entropy map calculated using (1) for local regions of size 16×16 pixels and an offset of 4 pixels is shown in Fig. 2. Fig. 3 shows two versions of the airport image with superimposed non-overlapping squares representing 40 and 80 regions having the highest entropy values in this image. While a non-overlapping fixed grid for resource allocation is shown, the methods presented are not restricted to this choice. The local regions identified in Figs. 2 and 3 appear to correlate well with our visual attention in this image.

Computing local entropy using intensity histograms suffers from a number of drawbacks. (i) It is computationally-burdensome for large block sizes. (ii) Its ability to identify regions of interest is sensitive to noise in the image. (iii) Its abilities vary widely with choice of block size and offset [6]. Figs. 4 and 5 illustrate some of the difficulties with this approach for an entropy map using 8×8 block sizes with



(a) (b)

Fig. 3. Resource allocation maps generated from the image in Fig. 2: (a) 40 regions, and (b) 80 regions.

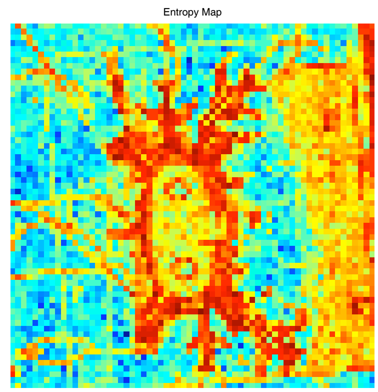


Fig. 4. Local entropy map generated from intensity histograms, 8×8 block size, no overlap.

no overlap. As can be seen in Figs. 5(a) and (b), regions of largely constant intensity are identified to the right of the airport, causing resources to be mis-allocated and likely leading to an overall lower-quality reconstruction.

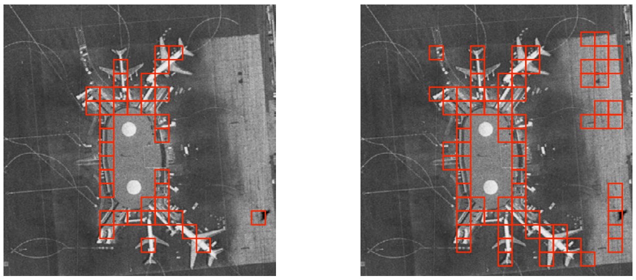
For these reasons, we desire alternative representations of local entropy for saliency estimation. One such alternative, developed by Huck *et al* [7], is the information rate defined as

$$\mathcal{H} = \frac{1}{2}\iint_{\hat{B}} \log \hat{\Phi}_L(u, v)\, du\, dv - \frac{1}{2}\iint_{\hat{B}} \log \hat{\epsilon}^2(u, v; k)\, du\, dv, \tag{2}$$

where $\hat{\Phi}_L(u, v)$ is the power spectral density (PSD) of the radiance-field within the sampling passband $\hat{B}$ and $\hat{\epsilon}^2(u, v; k)$ is the minimum mean-square restoration-error PSD that accounts for the loss of information that the perturbations of the image-gathering process cause within the sampling passband. Eq. (2) has been derived assuming that the radiance field is wide-sense stationary and Gaussian-distributed. A discrete approximation to (2) is

$$\widehat{H}_{ij} = \sum_u \sum_v \log P_{ij}(u, v) - K, \tag{3}$$

where $P_{ij}(u, v)$ is the discrete 2D PSD of the signal in the local region centered at pixel (i, j), u and v are discrete fre-



(a) (b)

Fig. 5. Resource allocation maps generated from the image in Fig. 4: (a) 40 regions, and (b) 80 regions.

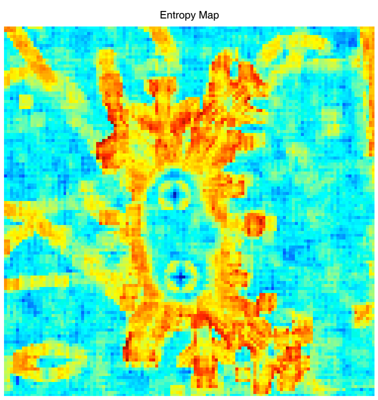


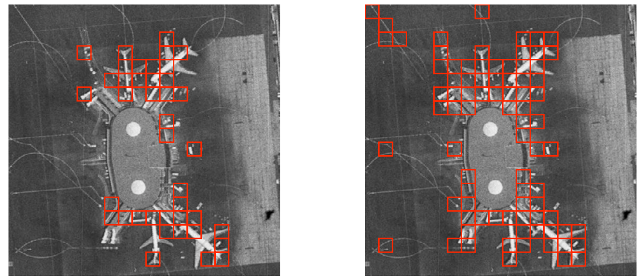
Fig. 6. Local entropy map generated using (3), 16×16 block size, 4-pixel offset.

quency bin indices, and K is constant due to quantization noise.

Figs. 6 and 7 show entropy maps generated using (3) and the resulting resource allocation maps for 16×16 local pixel regions and a 4-pixel offset. Investigation of this particular formula for region-of-interest calculation indicates that it is unsuitable for identifying salient regions within a scene for aerial images for the following reason: The value of the local entropy maps is biased by the total power of the image within each local region, causing brighter portions of the image to be identified as having higher saliency. Hence, another approach for saliency map estimation is needed.

3. SALIENCY MAP ESTIMATION USING 2D POWER SPECTRAL DENSITY

Although (3) is not appropriate for finding salient regions within image data in the chosen application, it provides an important insight: the 2D power spectral density can be used as a descriptor of local saliency within a scene. We propose to use the *normalized PSD* in place of the intensity histogram within Gilles' saliency map algorithm, such that the local PSD is treated like a probability distribution. This *spatial frequency histogram* for (i, j)th block is found by dividing the PSD by the sum of the powers in all of the bins in the frequency domain, or



(a) (b)

Fig. 7. Resource allocation maps generated from the image in Fig. 6: (a) 40 regions, and (b) 80 regions.

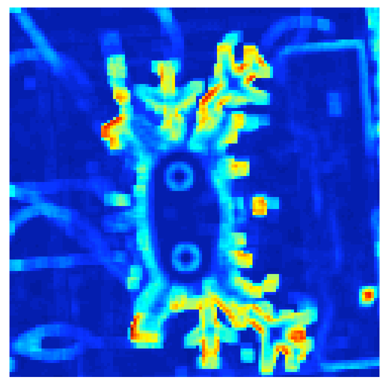


Fig. 8. Local entropy map generated using (5), 16×16 block size, 4-pixel offset.

$$p_{ij}(u,v) = \frac{P_{ij}(u,v)}{\sum_y \sum_z P_{ij}(y,z)} \quad (4)$$

Note that $p_{ij}(u, v)$ does not represent any randomness, and thus the proposed measure is not equal to the entropy of the data. The resulting saliency map estimator is

$$J_{ij} = -\sum_u \sum_v p_{ij}(u,v) \log_2 p_{ij}(u,v). \quad (5)$$

The proposed saliency map estimator has several interesting properties. Firstly, it is maximized when the 2D spectrum within a local block is flat, corresponding to regions containing energy at all spatial frequencies. It is orientation-independent, such that rotations and translations of image content do not change its value. Finally, regions within a scene that contain spatial features such as objects and edges will be identified with a higher saliency than other regions due to the broader frequency spectrum of the former region types. These attributes suggest that (5) can be used to estimate regions-of-interest within a scene.

Local saliency maps and resource-allocation maps for the airport image using (5) are shown in Figs. 8 and 9 for a block

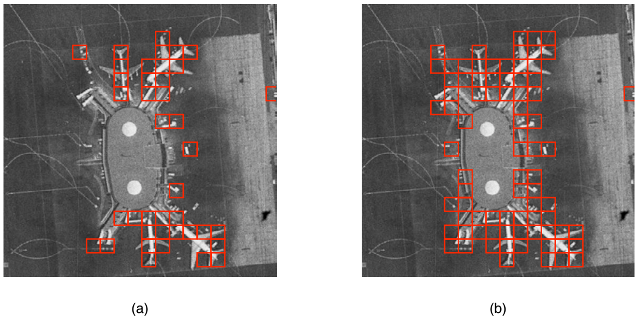


Fig. 9. Resource allocation maps generated from the image in Fig. 8: (a) 40 regions, and (b) 80 regions.

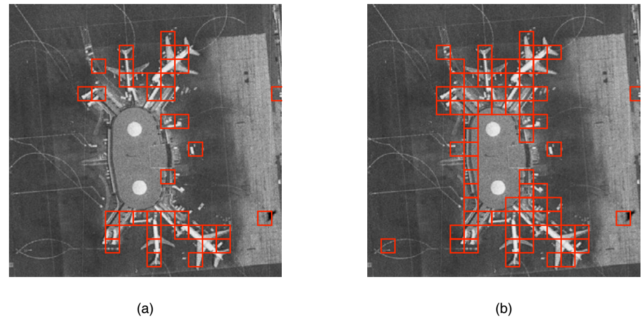


Fig. 11. Resource allocation maps generated from the image in Fig. 10: (a) 40 regions, and (b) 80 regions.

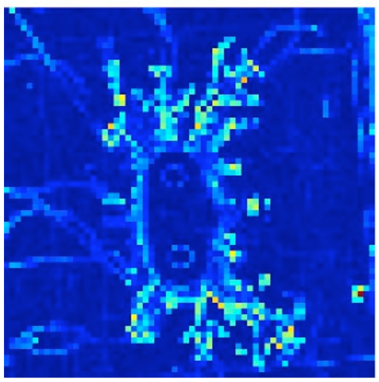


Fig. 10. Local entropy map generated using (5), 8×8 block size, no overlap.

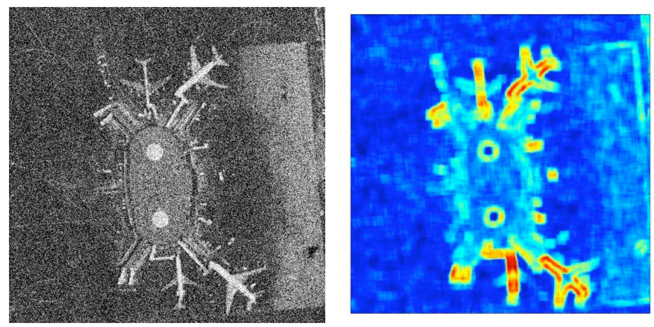


Fig. 12. (left) Image of a airport with a noise variance of 0.05 in normalized scale. (right) Saliency map of the noisy airport image using DC compensation of $8L^2$, for $L \times L$, $L = 16$ block size and a 2-pixel offset.

size of 16×16 pixels and a 4-pixel offset. Figs. 10 and 11 show the local saliency maps using (5) for a block size of 8×8 pixels and no overlap. From this and other numerical evaluations performed on aerial image data, the saliency maps generated using the spatial frequency histogram appear to identify regions of interest more often and with a higher accuracy than saliency maps generated from intensity histograms, and the behavior of (5) remains well-behaved even for small block sizes. They also have another distinct advantage over intensity-histogram-based methods: the effects of additive noise can be easily compensated. Fig. 12 shows a noisy version of the airport image along with the local saliency map computed using the spatial frequency histogram, in which the DC power of each local block has been replaced by a constant value proportional to the size of the block. As can be seen, the regions around the airport terminal remain highlighted in the saliency map despite the significant noise present. Details regarding these properties, as well as a complete description of how Fig. 12 was calculated, can be found in [6].

4. CONCLUSIONS

In designing computational imaging systems, allocating imaging resources to regions within a scene can be automated using methods for estimating local image saliency. In this paper, we explore two existing and one new techniques for determining regions of interest that employ intensity and spatial frequency histograms as descriptors of local content. Our novel method employs the local power spectral density within Gilles' saliency map estimator. Application to aerial images indicates the usefulness of the method for region-of-interest estimation. The proposed technique can be further simplified when combined with parametric models for PSD estimation of natural images; additional details regarding these parametric methods are in [6].

5. REFERENCES

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

[2] 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.

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

[5] S. Gilles, "Robust description and matching of images." Ph.D. Dissertation, University of Oxford, Oxford, 1998.

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

[7] F. O. Huck, C. L. Fales, and Z. ur Rahman, *Visual Communication: An Information Theory Approach* (Boston: Kluwer, 1997).