31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Multiscale Data Reduction with Flexible Saliency Criterion for Biological Image Analysis

William J. Bosl *Member, IEEE*

Abstract—**Analysis of biomedical images requires attention to image features that represent a small fraction of the total image size. A rapid method for eliminating unnecessary detail, analogous to pre-attentive processing in biological vision, allows computational resources to be applied where most needed for higher-level analysis. In this report we describe a method for bottom up merging of pixels into larger units based on flexible saliency criteria using a method similar to structured adaptive grid methods used for solving differential equations on physical domains. While creating a multiscale quadtree representation of the image, a saliency test is applied to prune the tree to eliminate unneeded details, resulting in an image with adaptive resolution. This method may be used as a first step for image segmentation and analysis and is inherently parallel, enabling implementation on programmable hardware or distributed memory clusters.**

I. Introduction

Extracting relevant information from complex images is challenging in part because images must be processed at a pixel level, even when the scale of objects of interest is much larger. A key attribute of human vision is the ability to rapidly filter out large amounts of irrelevant data and extract useful information from a few salient parts of the whole visual scene [1]. Determination of what is important and what is clutter depends on the context. Biological vision is hierarchical process that involves a rapid, pre-attentive reduction of data volume followed by higher level perception based on conscious recognition of expected shapes [1,2]. The first step in this hierarchy must be fast and relatively simple, while still capturing the most important information. The algorithm presented here is intended as a structured digital analog to the low-level, data-driven pre-attentive processing that occurs in the retina before signals are sent through the optic nerve, where data volumes are reduced by up to four orders of magnitude[3].

Structured adaptive mesh refinement methods have been developed for large three-dimensional simulations of physical systems to focus computational resources only where high resolution is needed[4]. Since pixels in an image stack are arranged in a uniform grid structured adaptive mesh methods may be usefully adopted for neural image representation. A key issue for creating adaptively refined meshes is the criterion used to decide when to divide a block into smaller blocks. The refinement criterion determines which details will be retained and which are unimportant. A refinement strategy based on pixel brightness variation is described for the particular images shown here. However, the method is flexible so that alternative criteria may be defined that use image features or characteristics, such as measures based on texture or other derived quantities, may be implemented as alternative refinement criteria. The intent of this method is not to enhance images, but to provide a fast, efficient means for reliably reducing the amount of information necessary to represent the essential features in an image for rapid analysis, key feature extraction or storage and compression.

II. Method

A bottom-up approach to image decomposition is implemented. The quadtree is pruned using a saliency criterion while the tree is being constructed, thus requiring only one pass through the data. Saliency tests discussed here are based on colors, but relatively simple tests that use texture or other image characteristics are possible. The resulting structured multiscale image (SMI) significantly reduces the number of data points needed for to represent the essential features and appropriate detail in an image. The assumption is that this is needed because large numbers of images or image stacks must be processed and the results of processing stored.

A. Multiscale Image Decomposition

Without loss of generality, suppose that an image is 2^N x 2^N pixels. The entire image is defined to be a single node or superpixel of level zero. The color of a superpixel is the mean of all pixels contained therein. Level N in this case refers to the collection of all single pixels. Starting with the upper left corner of the image, sets of pixels may be combined to create superpixels of size 2x2, creating a level N-1 image. Similarly, sets of 4 nodes at level N-1 may be

Manuscript received April 24, 2009. This work was supported in part by the National Institutes of Health under Grants 1R21MH083606 and 5K01HG003418.
W. J. Bosl is with the Children's Hospital Boston Informatics Program at Harvard-MIT Division of Health Sciences and Technology, Boston, MA 02115 USA (corresponding author phone: 617-919-2517; fax: 617-730-0921-5555; e-mail: william.bosl@childrens.harvard.edu).

combined to create larger superpixels. The four pixels that make up a node at the next level are children of that node. This process is used to create a quadtree structure. Each block of four pixels is tested for salience and a parent node created. Color is the property of interest in the examples discussed here. The color of each parent node is the mean color of all of its children. As soon as a parent node is created, the child nodes are tested for salience. Hereafter 'salience' will refer to blocks that pass the refinement criterion. That is, salient blocks or pixels are determined to contain important details and should not be merged into larger blocks. Figure 1 illustrates how pixels are merged to create larger blocks of uniform color. Leaves that are higher up in the quadtree represent larger blocks. We note that pruning a quadtree representation of an image is conceptually equivalent to pruning a decision tree. In both cases, the goal is to efficiently reduce the amount of data while retaining the most important information. The process is entirely bottom-up, requiring a single depth-first traversal of the tree as it is constructed. Numbers in figure 1 indicate the order of traversal in this example.

The depth of the quadtree can be predetermined if single pixel resolution is not required for the objects in the image or is not required for the analysis to be done. The number of levels allowed for decomposing the image determines the amount of detail retained. In some cases, restricting the decomposition to fewer levels is justified and will greatly reduce the size of the image data structure. If the maximum depth of the quadtree is set to be N-1 in an image that is a $2^N x 2^N$ pixels, then blocks of 4 pixels are averaged uniformly throughout the image initially.

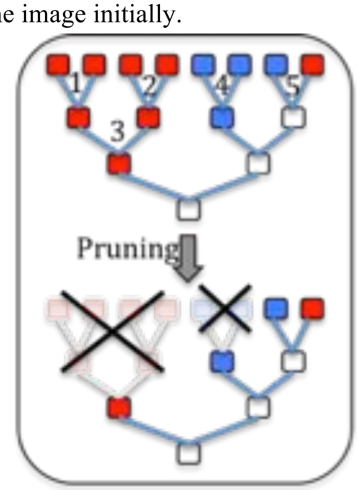


Figure 1. Pruning a quadtree. The first two pixels are both red, so the parent node is also colored red. This information is used to determine whether to prune the children or not. Numbers indicate the order of traversal in this example.

B. *Pruning and Saliency Criterion*

Fast elimination of unnecessary detail in an image requires a simple, robust test that effectively distinguishes where fine detail is important and where it is irrelevant. As with biological vision, relevance depends on the immediate image analysis needs. Figure 2 shows histology sections where the pathologist is primarily interested in the small dark blue objects in this image. The dark blue pixels in this image have mean RGB values near (125,75,150). These values are distinctly different from the pink and white regions in the rest of the image. A simple criterion is then to classify the color of each child node of a given node in succession as being either dark-bluish-purple or not.

The actual test used for this figure was to classify every node as belonging to a dark region of interest or not by computing a least-squares color distance. If any two child nodes belong to different classes, that particular node contains salient detail and all of the child nodes are labeled as leaves, unless a node is already defined to be a branch with refined or child nodes of interest. Although some detail is lost in this process, the primary reason for executing this process on an image is to discard details that are deemed irrelevant for a particular the purpose. All details concerning edges between dark blue regions and background are retained in this case because the saliency criterion requires it. Other details that are lost are considered irrelevant for the purpose of interest.

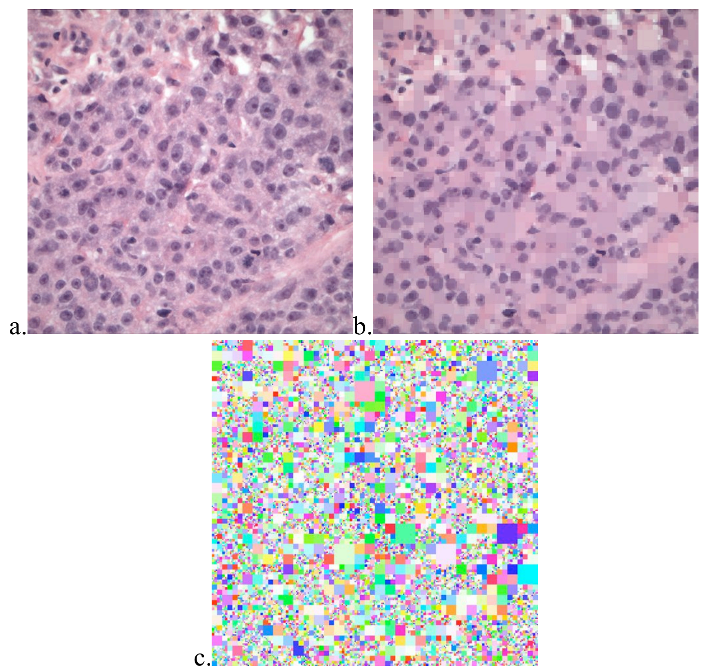


Figure 2. Histology section. a. Original image. b. Multiscale image. c. False color is used to show the superpixels used in the multiscale image. The multiscale image has 60:1 compression, yet salient fine detail is retained. Each node of the reconstructed image has the average color of all the original pixels in its region.

III. RESULTS

Another example of reducing data volume is shown in Figure 3. In this histology section the amount of data is reduced by a factor of 60 and still retains image features that are considered important, which, in this case, is to determine the fractional volume of the darker blue regions. Segmentation and analysis of the image in 3.c is considerably easier than in the original image.

The process of pruning pixels in the quadtree is analogous to pruning a neural arbor or a binary decision tree. Initially all connections are made as every pixel is connected to a parent node and all the details are retained. Pruning proceeds to eliminate unnecessary branches that contain little new information. Thus, if four child nodes are the same color, they contain no more information than the single parent cell, and the cost of maintaining that extra information is not justified. The resulting structure is a more efficient representation of the information content of the image. The process entailed here is similar to the backward pruning used in building an efficient decision tree [5].

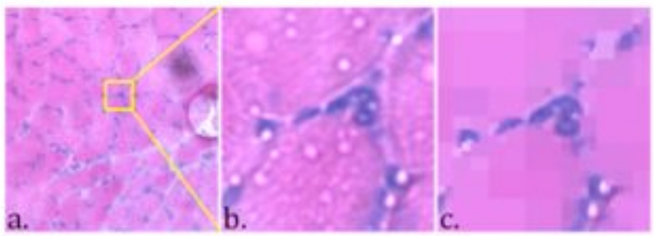
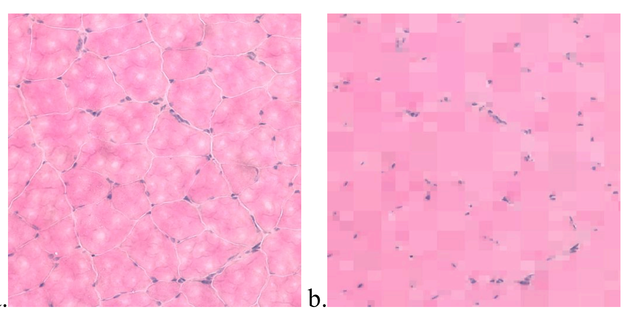


Figure 3. Histology section. a. Original image. b. Expanded view of original image in region outlined by yellow line. c. Multiscale image after 60:1 compression showing fine detail is retained. Each node of the reconstructed image has the average color of all the original pixels in its region.

The merge criterion determines what details are significant and must be kept and what variation is irrelevant and can be merged with neighboring blocks. In the examples shown here, only color or brightness variation among child nodes has been used to determine pruning in the examples shown here, although this can be generalized to use texture or other derived characteristics[6].

When the salient features in an image contain a relatively small fraction of the total area of the pixels, the data compression is correspondingly larger. The salient features in figure 4.a are the total area and number of small dark regions. Using the same pruning criterion as in previous images, the total number of superpixels required to represent the image is reduced by a factor of 400, shown in 4.b. Edge details near the dark regions are preserved at single pixel resolution, while unessential detail in the pink regions is merged into larger blocks.



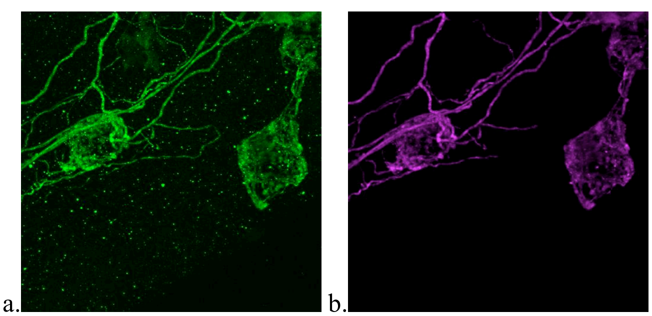
a. b.

Figure 4. Histology section. a. original image. b. multiscale image, 400:1 compression. When the salient information is sparse, compression is large while information loss is small.

Parallel implementation on clusters is relatively straightforward algorithmically. Image data is processed in a layered fashion. For a $2^N x 2^N$ image, N-2 steps are required to fully decompose the image. It is assumed that level 1, with four $2^{N-1} x 2^{N-1}$ subregions is the final step. As many as 2^{N-1} processors can be utilized to perform the averaging and pruning operation on level N-1, for which the child nodes are the original image pixels. Processing on level N-2 can begin as soon as each set of four child nodes on level N-1 are completed. On each level, an independent process will involve four additions and a division to compute the mean value of the parent node, computations to classify each of the child nodes, and a test to decide whether to prune or retain the child nodes. Once a decision has been made to not prune, all ancestors of that node will be retained and no testing or averaging is needed. Note that computations on each level involve only values at the level immediately below, thus the original pixel values are used only once.

IV. CONCLUSION

The algorithm presented here differs from filtering with a specified kernel and standard morphological operations such as opening and closing in that the filtering or merging is applied selectively in the image. The merge or refinement criterion determines where important image details are, such as edges or fine structures. The refinement criterion identifies regions where merging will not result in loss of essential information. The key here is to determine how identify regions in the image where detail must be retained. The refinement criterion discussed and used for examples in this paper works quite well for the images shown and may be appropriate for a wide range of medical images. Once that criterion is set, the algorithm may be used for compression to enable fast analysis of the image or efficient storage of results. Careful attention to the saliency or refinement criterion must be given to adapt this method for specific applications, as this determines what details or information in the image is essential and must be retained.



a. b.

Figure 5. a. original image. b. multiscale image, A maximum projection compression of 85 images from a stack of sub-mandibular ganglion (SMG) neuron images is shown. The images in the original stack are each 1024x1024 pixels, resulting in approximately 89 million pixels in the image. Structured multiscale adaptive meshing was used to reduce the stack to 73,551 uniformly colored superpixels. Note that fine edge detail is retained, while noise reduction is automatic. Brightness enhanced in original to emphasize noise.

The efficiency of adaptive meshing depends upon how much essential information is contained in the image volume. That is, large data volume reduction is possible when large relatively homogenous regions exist in the image stack that can be summarized by a single uniform block. Figures 2 and 3 contain a great deal of information on relatively small scales, but larger than single pixels for most features. In figure 5, a projection from a 3D stack of images of reconstructed neurons is shown. A statistical salience criterion was used that takes into account the node size as well as variability, as in Nock and Nielson (2004)[7]. Data compression is rather large, greater than 1000:1, and noise filtering is an automatic consequence of the pruning process. Random variation in small blocks was not statistically significant where the noise was scattered.

After pruning, simple thresholding was applied to dark nodes in figure 5 to and they were all set to black. Since this image is monocolor, segmentation is relatively simple after the pruning and thresholding operations. Relatively small amounts of information are sufficient to represent the image in figure 5 and the pruning process has eliminated unnecessary detail creating a more efficient representation of the essential details of the neural arbor in the original image stack. The amount of detail required in an image depends to a large extend on the questions to be answered. For many fundamental questions in neural imaging, the precise size and outline of axons and cell bodies is not critical.

When single pixel resolution is not required, the maximum number of levels to use in the initial octree decomposition can be limited. Figure 6 shows segmented images that use minimum block sizes of 1^3, 2^3, and 4^3 respectively when SMI is applied to the segmented image stack in figure 4. Visually, the images for the first two compression levels are nearly indistinguishable. Some detail appears to be lost in the lowest resolution image, but a compression ratio of greater than 10,000:1 was attained.

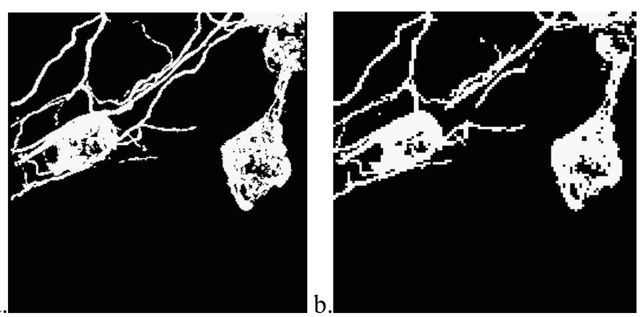


Figure 6. The SMI image from figure 5 (b) was further processed by restricting the smallest superpixel to 2x2 (a) and 4x4 (right). Total image compression from the original is 3000:1 and 10,000:1 respectively.

The SMI algorithm presented is primarily used as a rapid pre-processing step for reducing the amount of information in an image before more complex and computationally intensive segmentation or feature classification operations. The salience test may be defined appropriately for the particular application. For the structured multiscale process to be significantly faster than later processing, the salience test must be relatively simple and fast. With appropriate salience tests, noise elimination and edge detection are automatic. The latter is achieved when boundaries are defined by a well-defined difference in color, brightness, or other easily calculated property in neighboring pixels or superpixels.

The SMI algorithm may be run on a mono color or segmented image to provide an optimal data reduction. Each region in the structured decomposition will be represented by as large a superpixel as possible. For real-time applications, such as medical devices that process images, the algorithm may be implemented in programmable hardware to enable very fast reduction of image information. An ImageJ plugin will be made available that implements the SMI algorithm and will also provide methods for computing with structured multiscale images, such as nearest neighbor searching and extracting images with specified levels of resolution to enable further image analysis to be carried out much more efficiently than working with the original image pixels.

ACKNOWLEDGMENT

Thanks to Kamila Naxerova for histology images and to members of the Lichtman lab who supplied neural image data for this study. Thanks to Jeff Lichtman, Ryan Draft, Bobby Kashmuri, Juan Carlos Tapia, and Tamily Weisman for helpful suggestions.

REFERENCES

1 D. Marr and H. K. Nishihara, *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character* **200** (1140), 269 (1978).

2 L. Itti and C. Koch, *Nature reviews* **2** (3), 194 (2001).

3 P. Verghese and D. G. Pelli, *Vision research* **32** (5), 983 (1992).

4 R. D. Hornung, A. M. Wissink, and S. R. Kohn, *Engineering with Computers* (2006).

5 I. H. Witten and Eibe Frank, *Data mining : practical machine learning tools and techniques*, 2nd ed. (Morgan Kaufman, Amsterdam ; Boston, MA, 2005).

6 Rafael C. Gonzalez and Richard E. Woods, *Digital image processing*, 3rd ed. (Pearson/Prentice Hall, Upper Saddle River, NJ, 2008).

7 R. Nock and F. Nielsen, *IEEE Trans Pattern Anal Mach Intelligence* **26** (11), 1452 (2004).