A Hierarchical SVG Image Abstraction Layer for Medical Imaging

Edward Kim[1], Xiaolei Huang[1], Gang Tan[1], L. Rodney Long[2], Sameer Antani[2]

[1]Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA;
[2]Communications Engineering Branch, National Library of Medicine, Bethesda, MD

ABSTRACT

As medical imaging rapidly expands, there is an increasing need to structure and organize image data for efficient analysis, storage and retrieval. In response, a large fraction of research in the areas of content-based image retrieval (CBIR) and picture archiving and communication systems (PACS) has focused on structuring information to bridge the "semantic gap", a disparity between machine and human image understanding. An additional consideration in medical images is the organization and integration of clinical diagnostic information. As a step towards bridging the semantic gap, we design and implement a hierarchical image abstraction layer using an XML based language, Scalable Vector Graphics (SVG). Our method encodes features from the raw image and clinical information into an extensible "layer" that can be stored in a SVG document and efficiently searched. Any feature extracted from the raw image including, color, texture, orientation, size, neighbor information, etc., can be combined in our abstraction with high level descriptions or classifications. And our representation can natively characterize an image in a hierarchical tree structure to support multiple levels of segmentation. Furthermore, being a world wide web consortium (W3C) standard, SVG is able to be displayed by most web browsers, interacted with by ECMAScript (standardized scripting language, e.g. JavaScript, JScript), and indexed and retrieved by XML databases and XQuery. Using these open source technologies enables straightforward integration into existing systems. From our results, we show that the flexibility and extensibility of our abstraction facilitates effective storage and retrieval of medical images.

Keywords: Scalable Vector Graphics, Abstraction Layer, Cervigram, Content Based Image Retrieval

1. INTRODUCTION

An important goal of medical image processing, like that of remote sensing, industrial visual inspection systems, and image-based biometrics, is to transform raw images into a quantifiable symbolic form for ease of indexing and retrieval. Content-based image retrieval has been less successful than text-search engines because it is generally difficult to separate accidental characteristics of an image from those relevant to the query. Unlike symbolic data, like text or tables of attribute values, image content is essentially unstructured. In the cervicographic photographs that we aim to organize and retrieve, as seen in Figure 1, "accidental characteristics" are caused by unpredictable illumination, variability in tissue texture, specular reflections from moist surfaces, shadows cast by folds of tissue, uncertain camera orientation, and normal anatomical differences among subjects. The relevant characteristics consist of subtle changes in the color and texture of tissue induced by pathological conditions. Therefore the goal of our research is the conversion, with as little human intervention as possible, of a collection of raw images into a searchable symbolic structure that encodes relevant image content.

In our research, we aim at developing a unified, hierarchical approach that combines low-level image segmentation, with high-level object recognition and human interaction. Further, we propose to structure visual information in a hierarchy and to query images based on the resulting symbolic representation. The center piece of the hierarchy is an image abstraction layer that codifies knowledge obtained from a variety of low-level, mid-level or high-level automated image understanding algorithms and from user annotations. Figure 2 illustrates the hierarchy. The middle abstraction layer describes image content on which indexing and retrieval applications can operate; it insulates the raw sensory data from query applications. Retrieval can be done on any content piece that is available in the abstraction layer—it can be a raw color feature, a polygon's shape, or an object's semantic label. The content pieces can come from various sources: low-level image processing, high-level computer vision algorithms, prior knowledge about the image or video, or user provided annotations.

E-mail: edk208@lehigh.edu, Telephone: 1 610 758 3065, Fax: 1 610 758 4096

(a) sample 1 (b) sample 2 (c) sample 3 (d) sample 4

Figure 1. *Sample cervigram images from our dataset which consists of 60,000 color uterine cervix images collected by the National Cancer Institute (NCI) and National Library of Medicine (NLM). Each image has significant variability in image and tissue characteristics that we decipher and encode in a structured format.*

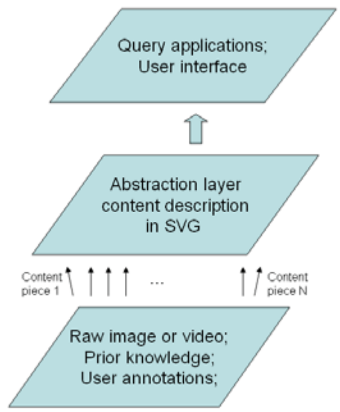


Figure 2. *Visual representation of what our abstraction layer represents. The abstraction layer is a structured knowledge representation above unstructured raw data that facilitates high level applications.*

Therefore, we first propose a scheme to organize low level features and high level semantic information in an image abstraction layer where content pieces are represented in XML and SVG, and then describe a web based tool that visualizes, manipulates, and searches our abstraction layer. Our method has several advantages over other methods. First, because SVG is a W3C standard, this allows for native web rendering and interactivity. In other proprietary formats and also other XML variants, there are few ways to display and visualize the feature vector that is created by these systems. In contrast, our SVG layer allows for native display in a standard web browser or XML parser. Secondly, since SVG is built upon XML, our representation benefits from the extensibility and versatility that is associated with the XML language. Indeed, if desired, a user could extend the abstraction layer to include elements from Unified Medical Language System (UMLS), DICOM, or other PACS systems with little effort. Lastly, our abstraction layer supports and maintains important mutual relationships between image elements. Often times there are different possibilities of segmentation and grouping in an image, ranging in levels of specificity. For this reason, our proposed abstraction can maintain and organize multiple levels of segmentation.

2. BACKGROUND INFORMATION

Because codifying images manually requires much time and is often subjective, there has been longstanding interest in automated methods that bridge the semantic gap between raw sensory data and codifiable knowledge.[1,2] Considerable progress has been made in the past decades on content-based image retrieval;[2,3] however, there is still a gap, between the photographic recording of an object and its interpretation in a context that extends far beyond the object itself.

Only using the high level semantic interpretation or only the low level pixel representation has been shown to be insufficient for effective retrieval.[4] Complex taxonomies and feature vectors have been proposed,[4–6] and typically follow the architecture described in Muller et. al.[7] These systems extract visual features from the raw image data, and store these features in different ways for distance and similarity calculations. However, since each system is significantly different in infrastructure and organization, there still remain questions regarding ontology extensibility, and integration when interfacing to existing PACS systems. Other research has utilized open standards; specifically several have utilized a variant of XML[8,9] but do not utilize the inherent visual XML representation, SVG.

2.1 SVG Background

The SVG standard[10] is an XML based file format that describes two dimensional vector graphic shapes, images and text. Vector images and shapes are defined by mathematical instructions rather than traditional image formats based on individual pixels. Thus, vector objects are scalable and resolution independent. Different vector objects include lines, circles, shapes, and polygons that are alterable by spatial transformations, alpha masks, and other effects defined by the W3C standard. These shapes can be customized in color, fill, texture, and stroke style. The XML Document Object Model (DOM) enables the use of existing technology for efficient searching and modification of the SVG objects. Because of its compatibility with other web standards, interactivity and animation effects are able to be leveraged by scripting languages such as ECMAScript.

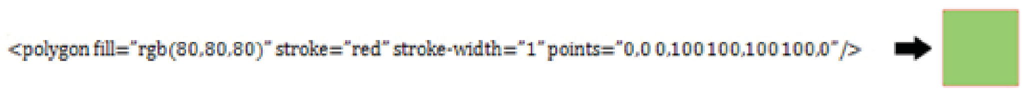


Figure 3. *Sample of the SVG code and rendering using the polygon element.*

When utilizing the SVG standard for our purposes, we decided to extend the SVG content directly without using the XML metadata languages, e.g. Resource Description Framework (RDF) or Web Ontology Language (OWL). The reason for this is because the SVG standard already allows for the inclusion of foreign namespaces and private data. The SVG agent is designed to include these extended elements into the DOM and ignore the rendering of these objects. The RDF/OWL description languages are not designed to be easily human readable, and further add complexity to our model.

3. METHODOLOGY

We consider an archive of 60,000 color uterine cervix images created by the National Library of Medicine (NLM) and the National Cancer Institute (NCI). Previous work on the raw image data set has yielded a wealth of feature information, including color, texture, and region boundary definitions. Also, for a significant subset of this data, we have obtained expert annotated ground truth with tissue classification and labeling. In this section, we first describe how to represent the raw image feature from the bottom up and then describe how to incorporate the high level semantic descriptions.

3.1 Low Level Features

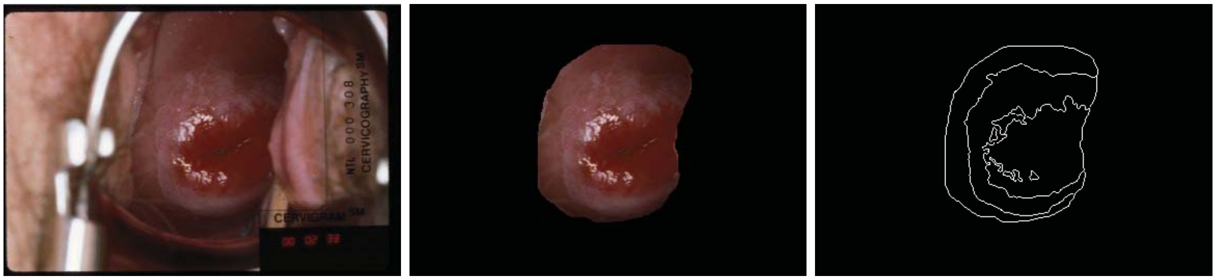
In SVG, there are several basic shape objects. For our purposes we choose polygons as our basis (path objects would also work) as described in Figure 3. The first step in encoding our images into our abstraction layer is the segmentation of the cervigram region of interest (ROI). For our application we used the EDISON mean shift algorithm[11,12] to identify edge enclosed regions as shown in Figure 4.

Given the edge result, we can then turn the resulting regions into SVG polygon points by tracing the boundary of each segmented region. For example, in Figure 5, we identify a region that has been segmented and then trace the contour of that region and return the point set. This point set becomes our "points" attribute for the polygon element.

We extend the basic instance of the polygon to incorporate other features extracted from the pixel level by adding elements to our polygon. We describe four features that we calculate for our SVG layer, *area, x centroid, y centroid,* and *eccentricity.*

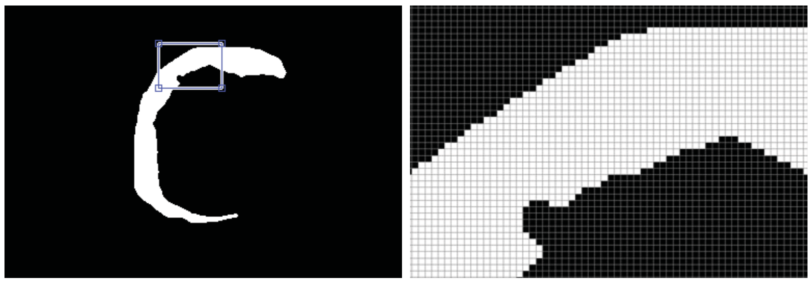
Area and centroid - The standard image moment calculation will provide us with the first three features, *area, x centroid,* and *y centroid*. The moments M_{ij} are calculated by,

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \tag{1}$$



(a) Raw image (b) ROI considered (c) Segmentation into region edges

Figure 4. *Sample cervigram image (a) with the identified ROI region (b). After using the EDISON mean shift method, we obtain the region edges of the segmentation result (c).*



(a) A sample filled region (b) Zoomed edge pixel values

Figure 5. *Given the binary filled region (a), we are able to trace the contour of the region and extract the boundary point set. (b) shows the zoomed image of the region.*

Where x and y are pixel locations and, I, is a binary region similar to that shown in Figure 5(a). The area is calculated by M_{00}, and the centroid is $(x_c, y_c) = (M_{10}/M_{00}, M_{01}/M_{00})$.

Eccentricity - The eccentricity of a region can be calculated using the central moments μ_{pq},

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y) \tag{2}$$

And using the eigenvalues λ_i, of the covariance matrix,

$$cov[I(x, y)] = \begin{bmatrix} \mu'_{20} & \mu'_{11} \\ \mu'_{11} & \mu'_{02} \end{bmatrix} \tag{3}$$

Where,

$$\mu'_{20} = \mu_{20}/\mu_{00} = M_{20}/M_{00} - x_c^2 \tag{4}$$

$$\mu'_{11} = \mu_{11}/\mu_{00} = M_{11}/M_{00} - x_c y_c \tag{5}$$

$$\mu'_{02} = \mu_{02}/\mu_{00} = M_{02}/M_{00} - y_c^2 \tag{6}$$

We can compute the eccentricity as,

$$\sqrt{1 - \frac{\lambda_2}{\lambda_1}} \tag{7}$$

Encoding - The encoding of these attributes can be seen in Figure 6. Additional features such as texture information and color histograms are added in a similar manner. In general, any low level feature extracted from the raw image can be represented here as an attribute in alphanumeric form. Regions of the image that belong to the same group are encapsulated in the SVG file with the <g>element. The <g>element is a container construct that associates graphics elements together. Because the <g>element can also contain other nested

<g>elements to an arbitrary depth, we can utilize these nodes to represent hierarchical relationships. At the top layer of our hierarchy we store a link to the raw image file. Advantageously, the SVG standard allows for the rendering of raster objects in conjunction with vector objects. We can combine all these features to encapsulate the raw image data with different levels of specificity as seen in Figure 6.

Figure 6. *Illustration of the SVG encoding of different regions in a group of elements. Characteristics such as the area, centroid position, and eccentricity are added as children to the polygon element. The increasing specificity of region information is naturally encoded by a hierarchy of group elements.*

3.2 High Level Semantics

In Zhao et. al,[4] effective retrieval techniques require both high level semantic information and machine level feature matching. Our abstraction layer naturally extends to incorporate region information and classification. Here, we describe a customized high level ontology for our cervigram image dataset using several elements defined in the SVG DTD (document type definition) for the group element, <g>, and several other custom defined nodes.

<title>, <desc>- These are SVG defined elements that provide information about the semantics.

<classification>- Given several ground truth segmentations and classifications, we are able to label certain tissue areas according to expert classifications. Typically, for cervigram image analysis, one of the most important observations is the Acetowhite region which is caused by the whitening of potentially malignant regions of the cervix epithelium. Some other observations are Squamous Epithelium, Columnar Epithelium, blood, polyps, cyst, Punctation, Atypical Vessel, among others.

<grade>- Areas of the cervigram containing grades of lesions from normal, CIN1, CIN2, to cancerous.

<annotations>- As seen in Muller et. al,[7] it is important for CBIR systems to support relevance feedback and user annotations. Thus, our abstraction layer defines annotation elements that can be dynamically updated by users through standard XML DOM modifications.

For different applications, semantic elements can be easily added or extended. For example, additional description elements can be added to the encoding including DICOM image information or UMLS definitions.

3.3 Query Application

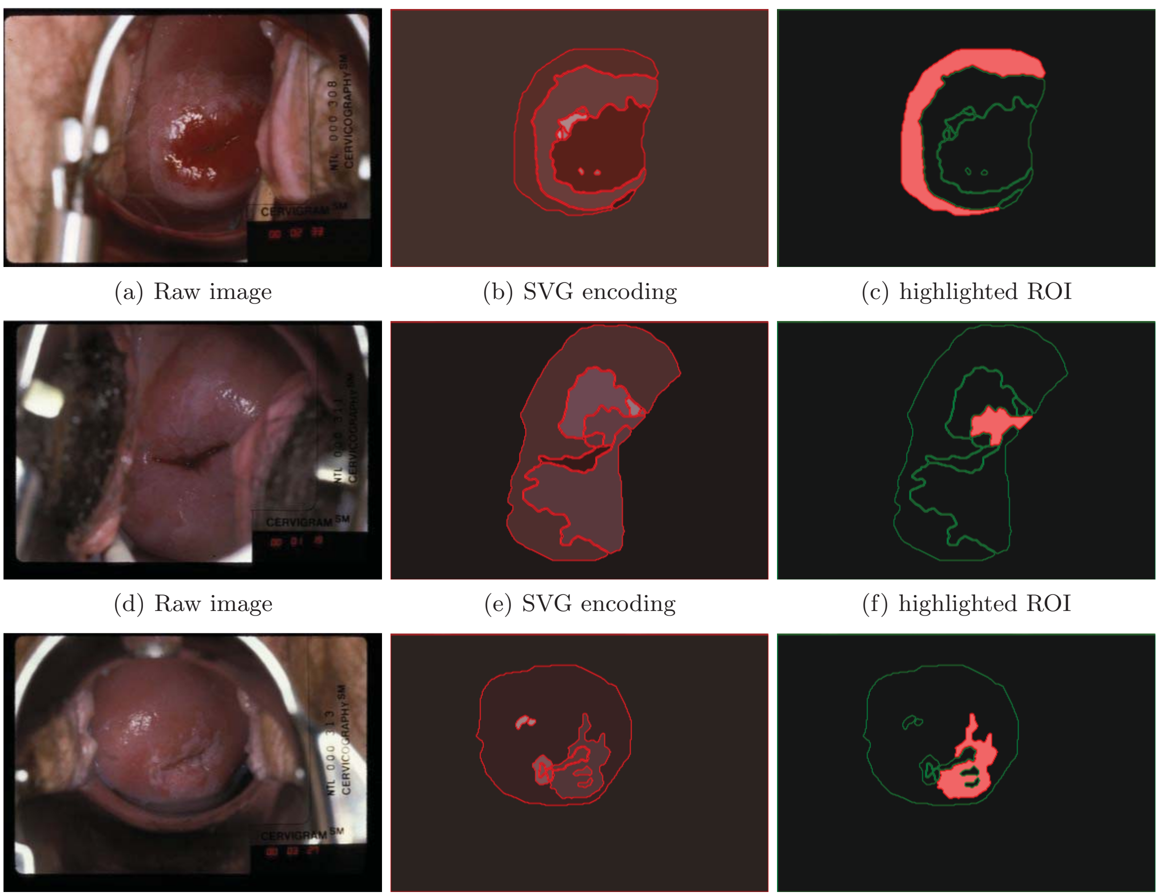
Our abstraction is stored and indexed in an XML database that is searchable using the XPath and XQuery language. Basic supported searches include, finding a region with a similar low level feature, or matching a high level semantic element. For example, we can search for all images with the high level classification, “acetowhite” and return the regions within the SVG documents as well as their specific region properties, e.g. color, size, eccentricity, etc. For a more advanced search, a combination of these low level and high level features could be used.

4. RESULTS

For our results we automatically segment and encode 30 cervigram images and build a web interface to visualize and interact with our abstraction. In our examples, visualization and interaction is performed by the Mozilla Firefox browser, version 3.

4.1 Visualization

A key advantage to our hierarchical layer is the visualization of image features. We are able to view the raw image, the segmentation results in different levels of specificity, and user annotated tissue classifications. Further, using the SVG capabilities, we are able to manipulate the abstraction layer to better visualize boundaries and ROIs (regions of interest). As noted before, viewing and editing the XML can occur through a standard web browser. Several different visualization effects can be seen in Figure 7.



(a) Raw image (b) SVG encoding (c) highlighted ROI

(d) Raw image (e) SVG encoding (f) highlighted ROI

(g) Raw image (h) SVG encoding (i) highlighted ROI

Figure 7. *Sample cervigram images (a)(d)(g) and their corresponding SVG visualizations (b)(e)(h). The ROIs, (c)(f)(i) can be selected and highlighted by altering the polygon “fill” and “stroke” attributes.*

4.2 Interaction

By utilizing existing scripting technologies such as Javscript, we can interact with any part of the XML DOM defined in our abstraction layer. The scripts enable the web tool to alter the visual effects of the regions or add or delete elements in the SVG definition. To annotate medical images, we select a region in our abstraction and add user level information via AJAX, HTML forms, and XQuery. An example can be seen in Figure 8, where we allow the user to provide an annotation to the ROI.

We build a complete user interface that allows the user to traverse the XML database, view, and manipulate the SVG documents. Several example screen shots that illustrate the full UI can be seen in Figure 10.

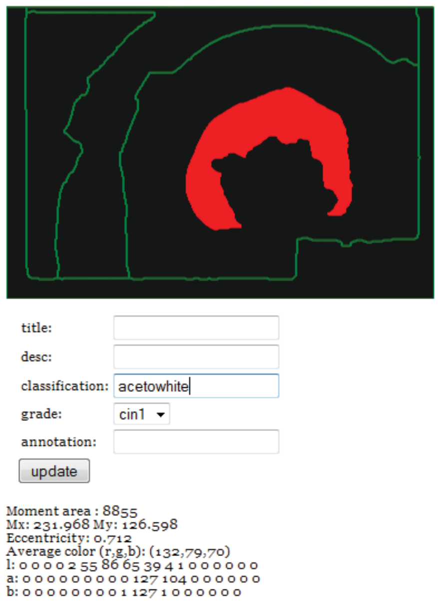


Figure 8. *Annotation interface example where a user can dynamically update the semantic description of a region.*

4.3 Basic Query

Using our database and the XQuery language, we demonstrate a sample search. For this search, we perform a high level semantic search on the element, "classification", where the value of the tissue classification is, "acetowhite". Figure 9 shows the first ten results of this query.

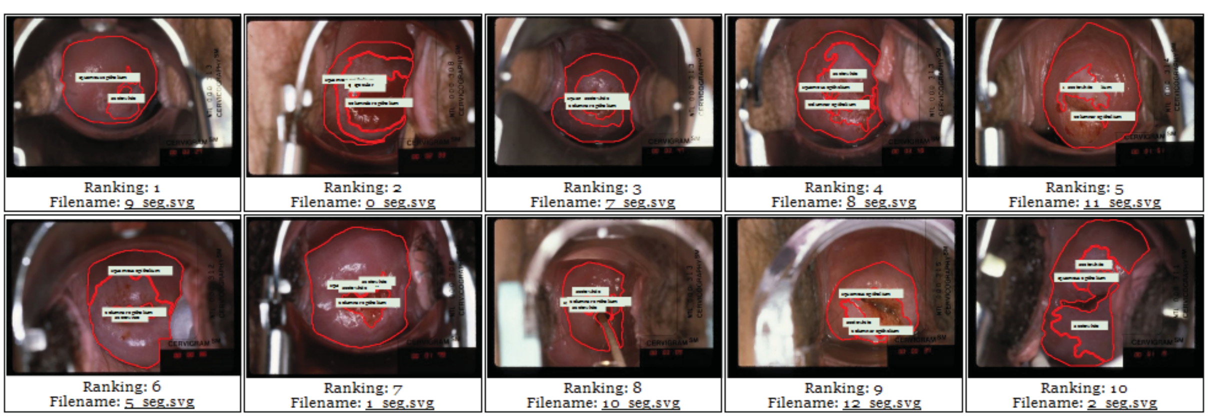
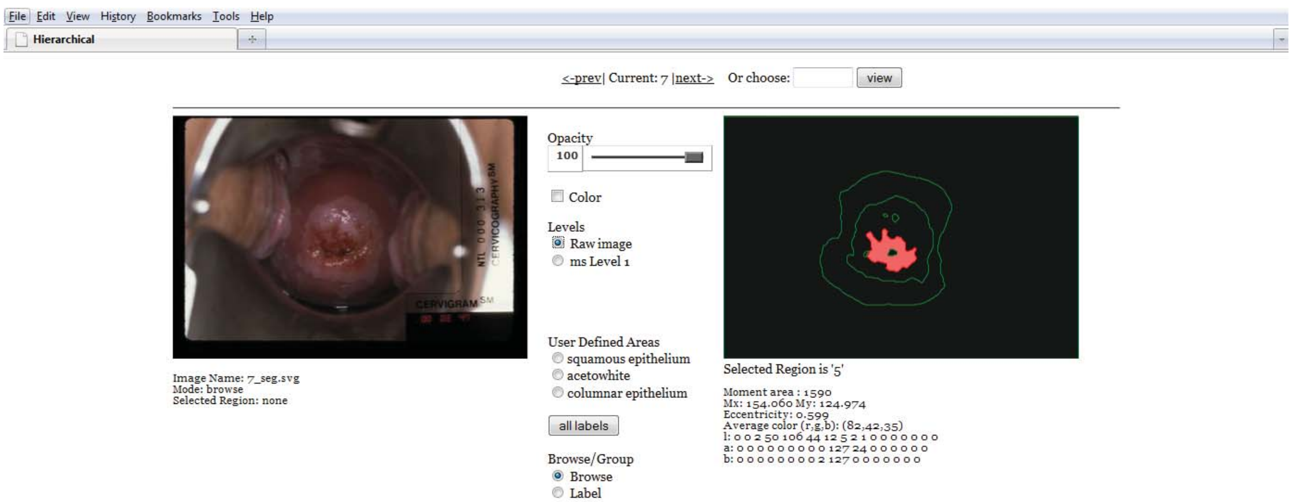


Figure 9. *Query for the semantic classification label, "acetowhite". The results show the matching files and their region classifications.*

5. CONCLUSION AND FUTURE WORK

In summary, we present a SVG based abstraction layer that is able to organize both low level machine features and high level semantics for efficient storage and searching of medical images. Additionally, we built a web system to visualize and interact with our abstraction. Several advantages of our method include interactivity, visualization, and extensibility.

For our future work, we plan to utilize our layer for interesting and complex content queries. For our low and high level similarity measurements, recent research has been done on what distance methods are most suitable for the SVG medium. As specified in Jiang et. al,[13] Hamming distance, Hausdorff distance and Edit distance similarity measures work well for SVG point sets. Combining these measures with traditional text, low level extracted features, and high level semantics defined in our abstraction layer could provide for an interesting and effective content based image retrieval system.



(a) User interface in a Firefox web browser. This image shows the raw image (left) and a selected ROI (right).

(b) User interface in a Firefox web browser. This image shows the segmentation result (left) and the semantic classification labels (right).

Figure 10. *User interface screen shots of our web tool used for visualization and interaction with our SVG dataset.*

ACKNOWLEDGMENTS

This work is supported in part by the DHHS contract Grant NIHNLM- HHSN276200900722P and the NSF grant IIS-0812120. The authors would like to thank the National Library of Medicine Communications Engineering Branch and the National Cancer Institute for providing the data used in this work, and also acknowledge stimulating discussions with Dr. Daniel Lopresti (Lehigh) and Dr. George Nagy (Rensselaer Polytechnic Institute).

REFERENCES

[1] Deserno, T., Antani, S., and Long, R., "Ontology of gaps in content-based image retrieval," *Journal of Digital Imaging* **22**(2), 202–215 (2009).

[2] Smeulders, A., Worring, M., Santini, S., Gupta, A., and Jain, R., "Content-based image retrieval at the end of the early years," *IEEE Transactions on pattern analysis and machine intelligence* **22**(12), 1349–1380 (2000).

[3] Datta, R., Joshi, D., Li, J., and Wang, J., "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys* **40**(2), 1–60 (2008).

[4] Zhao, R. and Grosky, W., "Bridging the semantic gap in image retrieval," in [*Distributed multimedia databases: Techniques and applications*], 14–36, Idea Group Publishing (2001).

[5] Lehmann, T., Gold, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., and Wein, B., "Content-based image retrieval in medical applications," *Methods of Information in Medicine* **43**(4), 354–361 (2004).

[6] Xue, Z., Long, L., Antani, S., Jeronimo, J., and Thoma, G., "A Web-accessible content-based cervicographic image retrieval system," *Proceedings of SPIE* **6919** (2008).

[7] Muller, H., Michoux, N., Bandon, D., and Geissbuhler, A., "A review of content-based image retrieval systems in medical applicationsclinical benefits and future directions," *International journal of medical informatics* **73**(1), 1–23 (2004).

[8] Tahmoush, D. and Samet, H., "A web collaboration system for content-based image retrieval of medical images," *Proceedings of SPIE–Medical Imaging* **6516** (2007).

[9] Eidenberger, H. and Breiteneder, C., "Semantic feature layers in content-based image retrieval: implementation of human world features," *Proceedings of International Conference on Control, Automation, Robotic and Vision* **1**, 174– 179 (2002).

[10] "Scalable vector graphics (svg)," *http://www.w3.org/Graphics/SVG/* (2003).

[11] Comaniciu, D. and Meer, P., "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence* **24**(5), 603–619 (2002).

[12] Meer, P. and Georgescu, B., "Edge detection with embedded confidence," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(12), 1351–1365 (2001).

[13] Jiang, K., Fang, Z., Ge, Y., and Zhou, Y., "Information Retrieval through SVG-based Vector Images Using an Original Method," *Proceedings of IEEE International Conference on e-Business Engineering (ICEBE)* , 183–188 (2007).