Interactive classification and content-based retrieval of tissue images

Selim Aksoy, Giovanni Marchisio, Carsten Tusk, Krzysztof Koperski

Insightful Corporation, 1700 Westlake Ave. N., Seattle, WA, 98109

ABSTRACT

We describe a system for interactive classification and retrieval of microscopic tissue images. Our system models tissues in pixel, region and image levels. Pixel level features are generated using unsupervised clustering of color and texture values. Region level features include shape information and statistics of pixel level feature values. Image level features include statistics and spatial relationships of regions. To reduce the gap between low-level features and high-level expert knowledge, we define the concept of prototype regions. The system learns the prototype regions in an image collection using model-based clustering and density estimation. Different tissue types are modeled using spatial relationships of these regions. Spatial relationships are represented by fuzzy membership functions. The system automatically selects significant relationships from training data and builds models which can also be updated using user relevance feedback. A Bayesian framework is used to classify tissues based on these models. Preliminary experiments show that the spatial relationship models we developed provide a flexible and powerful framework for classification and retrieval of tissue images.

Keywords: Content-based image retrieval, image classification, spatial relationships, fuzzy logic, tissue images.

1. INTRODUCTION

A challenging problem in medical imaging is automatic classification and retrieval in large image databases. Most of the proposed approaches use low-level features like color histograms and texture features to index images and then use distance measures in these feature spaces to find similarities between them. However, there is a large semantic gap between the low-level features and the high-level expert knowledge.

The VisiMine system¹ supports interactive classification and retrieval of tissue images by modeling them on pixel, region and image levels. Pixel level characterization includes color features and Gabor, co-occurrence and Laws texture features. After the features are computed for each pixel, an automatic region segmentation algorithm is used to compute an approximate polygon decomposition of each image. Then, region level features are computed using moments for shape and orientation information, and statistics of pixel features within and around individual regions.

Low-level features cannot always capture detailed knowledge in medical images. Traditional region or image level search algorithms assume that the regions or images consist of uniform pixel feature distributions. However, complex tissue images usually contain many pixels and regions that have different feature characteristics. Furthermore, two images with similar regions can have very different interpretations if the regions have different spatial arrangements. Therefore, we need a higher level visual grammar to describe these scenarios.

Previous approaches for region-based analysis and modeling of their spatial relationships² mostly include manual delineation by experts. Shyu *et al.*³ developed a content-based image retrieval system that used features locally computed from manually delineated regions. Neal *et al.*⁴ developed topology, part-of and spatial association networks to symbolically model partitive and spatial adjacency relationships of anatomical entities. Tang *et al.*^{5, 6} divided images into small sub-windows, and trained neural network classifiers using color and Gabor texture features computed from these sub-windows and the labels assigned to them by experts. These classifiers were used to assign labels to each sub-window in unknown images, and the labels were verified using a knowledge base of label spatial relationships. These rules were created by experts as some labels could not exist together

Further author information: Send correspondence to S.A. at saksoy@insightful.com.

with other labels. Petrakis and Faloutsos⁷ used attributed relational graphs to represent features of objects and their relationships in MRI images. They assumed that the graphs were already known for each image in the database and concentrated on developing fast search algorithms. Chu *et al.*⁸ described a knowledge-based semantic image model to represent image objects' characteristics. Graph models are powerful representations but are not usable due to the infeasibility of manual annotation in large databases. Different structures in complex medical images have different sizes so fixed sized grids cannot capture all structures either.

In this paper, we describe an automatic probabilistic framework that includes prototypes of primitive regions, their spatial relationships, and automatic and supervised algorithms to use them for content-based retrieval and classification. The rest of the paper is organized as follows. The concept of prototype regions is defined in Section 2. Spatial relationships of these prototype regions are described in Section 3. Algorithms for image retrieval and classification using the spatial relationship models are discussed in Sections 4 and 5, respectively. Conclusions are given in Section 6.

2. PROTOTYPE REGIONS

The first step to construct a visual grammar is to find meaningful regions in an image. Automatic extraction of regions is required to handle large amounts of data. To mimic the identification of regions by experts, we define the concept of prototype regions. A prototype region is a region that has a relatively uniform low-level pixel feature distribution and describes a simple tissue structure or part of a more complex structure. Color values or other pixel-level features like texture can be used for region segmentation. Ideally, a prototype is frequently found in a specific class of tissues and differentiates this class of tissues from others. In addition, using prototypes reduces the possible number of associations between regions and makes the combinatorial problem of region matching more tractable. (This will be discussed in detail in Section 5.)

VisiMine uses unsupervised k-means and model-based clustering to automate the process of finding prototypes. Before unsupervised clustering, a piecewise-polynomial multiscale energy-based region growing segmentation algorithm⁹ is used to find regions in the image. K-means clustering¹⁰ partitions the input sample into k clusters by iteratively minimizing a squared-error criterion function. After clustering, prototypes are represented by the means of the feature vectors associated with each cluster. Then, Euclidean distance in the feature space is used to match regions to prototypes. In model-based clustering,¹⁰ clusters are represented by parametric density models. We use a Gaussian mixture model where prototypes correspond to the components of the mixture. We use the Expectation-Maximization algorithm¹¹ to estimate the parameters of the Gaussian components and their weights in the mixture. Then, the maximum a posteriori probability (MAP) rule is used to assign a prototype label to each region with the degree of match being the posterior probability of the prototype given the feature vector of that region. Figure 1 shows example prototype regions for different tissue images.

3. REGION RELATIONSHIPS

3.1. Second-order Region Relationships

Second-order region relationships consist of the relationships between region pairs. These pairs can occur in the image in many possible ways. However, the regions of interest are usually the ones that are close to each other. Representations of spatial relationships depend on the representations of regions. In VisiMine, regions are represented by their boundary pixels and moments. Other possible representations include minimum bounding rectangles, Fourier descriptors² and graph-based approaches.⁷

The spatial relationships between all region pairs in an image can be represented by a region relationship matrix. To find the relationship between a pair of regions represented by their boundary pixels and moments, we first compute

- perimeter of the first region, π_i
- perimeter of the second region, π_j





(b) Eosinophil and red blood cells in a peripheral blood tissue



(c) Mast cells in a connective tissue

(d) Anterior lobe acidophils in a pituitary gland tissue



(e) Acidophils and basophils in a pituitary gland tissue

(f) Interstitial cells of leydig in a male reproductive system tissue



(g) Plicae, villi and mucosa in a small intestine tissue

Figure 1. Example prototype regions for different tissue images. Images on the left are the original color images, images on the right contain pseudo-colored prototype labels for all regions.





- common perimeter between two regions, π_{ij}
- ratio of the common perimeter to the perimeter of the first region, $r_{ij} = \frac{\pi_{ij}}{\pi_i}$
- centroid of the first region, ν_i
- centroid of the second region, ν_i
- distance between the centroids, $d_{ij} = \|\nu_i \nu_j\|$
- angle between the horizontal (column) axis and the line joining the centroids, θ_{ij}

where $i, j \in \{1, ..., n\}$ and n is the number of regions in the image. Then, the $n \times n$ region relationship matrix is defined as

$$R = \{\{r_{ij}, d_{ij}, \theta_{ij}\} \mid i, j = 1, \dots, n, \ \forall i \neq j\}.$$
(1)

One way to define the spatial relationships between regions i and j is to quantize the r_{ij} , d_{ij} and θ_{ij} values. However, quantization gives crisp (Boolean) decisions and may have limited expressiveness. Another way is to define them as relationship classes. Each region pair can be assigned a degree of their spatial relationship using fuzzy class membership functions. Denote the class membership functions by Ω_c with $c \in \{\text{DIS}, \text{BOR}, \text{INV}, \text{SUR}, \text{NEAR}, \text{FAR}, \text{RIGHT}, \text{LEFT}, \text{ABOVE}, \text{BELOW}\}$ corresponding to disjoined, bordering, invaded_by, surrounded_by, near, far, right, left, above and below, respectively. Then, the value $\Omega_c(r_{ij}, d_{ij}, \theta_{ij})$ represents the degree of membership of regions i and j to class c.

Among the above, disjoined, bordering, invaded_by and surrounded_by are perimeter-class relationships, near and far are distance-class relationships, and right, left, above and below are orientation-class relationships. These relationships are divided into sub-groups because multiple relationships can be used to describe a region pair, e.g. invaded_by from left, bordering from above, and near and right, etc. See Figure 2 for illustrations.

For the perimeter-class relationships, we use the perimeter ratios r_{ij} with the trapezoid membership functions shown in Figure 3(a). The motivation for the choice of these functions is as follows. Two regions are disjoined when they are not touching each other. They are bordering each other when they have a common perimeter. When the common perimeter between two regions gets closer to 50%, the larger region starts invading the smaller one. When the common perimeter goes above 80%, the relationship is considered an almost complete invasion, i.e. surrounding. For the distance-class relationships, we use the perimeter ratios r_{ij} and centroid distances d_{ij} with the sigmoid membership functions shown in Figure 3(b). For the orientation-class relationships, we use the angles θ_{ij} with the truncated cosine membership functions shown in Figure 3(c).

Note that the pairwise relationships are not always symmetric, i.e. $\Omega_c(r_{ij}, d_{ij}, \theta_{ij})$ is not necessarily equal to $\Omega_c(r_{ji}, d_{ji}, \theta_{ji})$. An implementation should keep track of the direction of asymmetric relationships. Furthermore, some relationships are stronger than others. For example, *surrounded_by* is stronger than *invaded_by*, and *invaded_by* is stronger than *bordering*, e.g. the relationship "small region *invaded_by* large region" is preferred over the relationship "large region *bordering* small region". The class membership functions are chosen so that only one of them is the largest for a given set of measurements $r_{ij}, d_{ij}, \theta_{ij}$. We label a region pair as having the perimeter-class, distance-class and orientation-class relationships

$$c_{ij}^{1} = \underset{c \in \{\mathsf{DIS},\mathsf{BOR},\mathsf{INV},\mathsf{SUR}\}}{\arg\max} \Omega_{c}(r_{ij}, d_{ij}, \theta_{ij})$$

$$c_{ij}^{2} = \underset{c \in \{\mathsf{NEAR},\mathsf{FAR}\}}{\arg\max} \Omega_{c}(r_{ij}, d_{ij}, \theta_{ij})$$

$$c_{ij}^{3} = \underset{c \in \{\mathsf{RIGHT},\mathsf{LEFT},\mathsf{ABOVE},\mathsf{BELOW}\}}{\arg\max} \Omega_{c}(r_{ij}, d_{ij}, \theta_{ij})$$
(2)

with the corresponding degrees

$$\rho_{ij}^t = \Omega_{c_{ij}^t}(r_{ij}, d_{ij}, \theta_{ij}), \quad t = 1, 2, 3.$$
(3)













Brown DISJOINED with blue Brown BORDERING blue

Brown INVADED BY blue Brown SURROUNDED BY blue

Green NEAR blue



from blue

(a) Perimeter-class relationships: disjoined, bordering, invaded_by and surrounded_by

(b) Distance-class relationships: near and far



(c) Orientation-class relationships: right, left, above and below

Figure 2. Spatial relationships of region pairs.



 ${\bf Figure \ 3.}\ {\rm Fuzzy\ membership\ functions\ for\ pairwise\ spatial\ relationships.}$

3.2. Higher-order Region Relationships

Higher-order region relationships can be decomposed into multiple second-order relationships. Therefore, the measures defined in the previous section can be used for each of the pairwise relationships and can be combined to measure the combined relationship. The equivalent of the Boolean "and" operation in fuzzy logic is the "min" operation. For a combination of k regions, there are $\binom{k}{2} = \frac{k(k-1)}{2}$ pairwise relationships. Therefore, the relationship between these k regions can be represented as lists of $\binom{k}{2}$ pairwise relationships using Equation (2) as

$$c_{1\dots k}^{t} = \{c_{ij}^{t} \mid i, j = 1, \dots, k, \ \forall i < j\}, \quad t = 1, 2, 3$$

$$\tag{4}$$

with the corresponding degrees computed using Equation (3) as

$$\rho_{1\dots k}^{t} = \min_{\substack{i,j=1,\dots,k\\i < j}} \rho_{ij}^{t}, \quad t = 1, 2, 3.$$
(5)

See Figures 4 and 5 for example decompositions.

4. IMAGE RETRIEVAL

Users can compose queries for complex tissue structures by giving a set of example regions or by selecting an area of interest. VisiMine encodes and searches for the query using the proposed visual grammar as follows:

- 1. Let k be the number of regions selected by the user. Find the prototype label for each of the k regions.
- 2. Find the perimeter ratio, centroid distance and orientation for each of the $\binom{k}{2}$ possible region pairs.
- 3. Find the spatial relationship and its degree for these k regions using Equations (4) and (5). Denote them by $\bar{c}^t = \{\bar{c}_{ij}^t \mid i, j = 1, \dots, k, \forall i < j\}, t = 1, 2, 3 \text{ and } \bar{\rho}^t, t = 1, 2, 3$, respectively.
- 4. For each image in the database,
 - (a) For each query region, find the list of regions with the same prototype label as itself. Denote these lists by $U_i, i = 1, ..., k$.
 - (b) Rank region groups $(u_1, u_2, \ldots, u_k) \in U_1 \times U_2 \times \cdots \times U_k$ according to the distance

$$\left| \min_{t=1,2,3} \bar{\rho}^t - \min_{\substack{t=1,2,3 \ i,j=1,\dots,k \\ i < j}} \Omega_{\bar{c}_{ij}^t}(r_{u_i u_j}, d_{u_i u_j}, \theta_{u_i u_j}) \right| \tag{6}$$

or alternatively according to

$$\max_{t=1,2,3} \max_{\substack{i,j=1,\dots,k\\i< j}} \left| \bar{\rho}_{ij}^t - \Omega_{\bar{c}_{ij}^t}(r_{u_i u_j}, d_{u_i u_j}, \theta_{u_i u_j}) \right|.$$
(7)

(c) The equivalent of the Boolean "or" operation in fuzzy logic is the "max" operation. To rank image tiles, use the distance

$$\left| \min_{t=1,2,3} \bar{\rho}^t - \max_{\substack{(u_1, u_2, \dots, u_k) \in \\ U_1 \times U_2 \times \dots \times U_k}} \left\{ \min_{\substack{t=1,2,3 \ i, j=1, \dots, k \\ i < j}} \Omega_{\bar{c}_{ij}^t}(r_{u_i u_j}, d_{u_i u_j}, \theta_{u_i u_j}) \right\} \right|$$
(8)

or alternatively the distance

$$\min_{\substack{(u_1, u_2, \dots, u_k) \in \\ U_1 \times U_2 \times \dots \times U_k}} \left\{ \max_{\substack{t=1,2,3 \ i, j=1, \dots, k \\ i < j}} \left| \bar{\rho}_{ij}^t - \Omega_{\bar{c}_{ij}^t}(r_{u_i u_j}, d_{u_i u_j}, \theta_{u_i u_j}) \right| \right\}.$$
(9)

In some cases, some of the spatial relationships (e.g. *above*, *right*) can be too restrictive. Relevance feedback can be used to find the most important relationship class (perimeter, distance or orientation) for a particular query. Example queries are given in Figures 6 and 7. These images come from a dataset of 615 images obtained from a histology course¹² at the University of Delaware. The dataset contains a single example for most of the tissue types so precision is expected to be low in query results.



Figure 4. Spatial relationships of an eosinophil, a red blood cell and the surrounding plasma in a peripheral blood tissue decomposed into pairwise relationships.



Figure 5. Spatial relationships of a mast cell, nuclei and the surrounding interstitial space in a connective tissue decomposed into pairwise relationships.



Figure 6. Top 8 search results for the red blood cell described in Figure 4.



Figure 7. Top 8 search results for the mucosa, villi and the surrounding cavity in a small intestine tissue.

5. IMAGE CLASSIFICATION

The visual grammar can also be used to classify images in a Bayesian framework. The input to the system is a set of example images for each class defined by the user. Let s be the number of classes, m be the number of relationships defined for region pairs, k be the number of regions in a region group, and t be a threshold for the number of region groups that will be used in the classifier. Denote the classes by w_1, \ldots, w_s . VisiMine automatically builds classifiers from the training data as follows:

- 1. Count the number of times each possible region group with a particular spatial relationship is found in the set of training images for each class. This is a combinatorial problem because the total number of region groups (unordered arrangements without replacement) in an image with n regions is $\binom{n}{k}$ and the total number of possible relationships (unordered arrangements with replacement) in a region group is $\binom{m+\binom{k}{2}-1}{\binom{k}{2}}$. A region group of interest is the one that is frequently found in a particular class of images but rarely exists in other classes. For each region group, this can be measured using class separability which is computed in terms of within-class and between-class variances of the counts.
- 2. Select the top t region groups with the largest class separability values. Let x_1, \ldots, x_t be Bernoulli random variables for these region groups, where $x_j = T$ if the region group x_j is found in an image and $x_j = F$ otherwise. Let $p(x_j = T) = \theta_j$. Then, the number of times x_j is found in images from class w_i has a Binomial (v_i, θ_j) distribution where v_i is the number of training images for class w_i . Using a Beta(1, 1) distribution as the conjugate prior, the Bayes estimate for θ_j is computed as

$$p(x_j = T | w_i) = \frac{v_{ij} + 1}{v_i + 2} \tag{10}$$

where v_{ij} is the number of training images for w_i that contain x_j . Using a similar procedure, the Bayes estimate for an image belonging to class w_i is computed as

$$p(w_i) = \frac{v_i + 1}{\sum_{i=1}^s v_i + s}.$$
(11)

In other words, discrete probability tables are constructed using v_i and v_{ij} , i = 1, ..., s, j = 1, ..., t, and conjugate priors are used to update them when new images become available via relevance feedback.

3. For an unknown image, search for each of the t region groups (determine whether $x_j = T$ or $x_j = F, \forall j$) and compute the probability for each class using the conditional independence assumption as

$$p(w_i|x_1, \dots, x_t) = \frac{p(w_i, x_1, \dots, x_t)}{p(x_1, \dots, x_t)}$$

= $\frac{p(w_i)p(x_1, \dots, x_t|w_i)}{p(x_1, \dots, x_t)}$
= $\frac{p(w_i)\prod_{j=1}^t p(x_j|w_i)}{p(x_1, \dots, x_t)}$. (12)

Assign that image to the best matching class using the MAP rule as

$$w^{*} = \arg \max_{w_{i}} p(w_{i}|x_{1}, \dots, x_{t})$$

= $\arg \max_{w_{i}} p(w_{i}) \prod_{j=1}^{t} p(x_{j}|w_{i}).$ (13)

Classification examples are given in Figures 8 and 9.



Figure 8. A distinguishing region group for a male reproductive system tissue (marked in the top-left image) and other images that contain a similar region group.



Figure 9. A distinguishing region group for a pituitary gland tissue (marked in the top-left image) and other images that contain a similar region group.

6. CONCLUSIONS

In this paper we described a probabilistic framework to automatically analyze complex tissue images using spatial relationships of regions and described algorithms to use them for content-based image classification and retrieval. The proposed approach includes identification of prototype regions in images, fuzzy modeling of region spatial relationships, and Bayesian classifiers to learn image classes based on automatic selection of distinguishing relations between regions. Future work includes using supervised methods to update prototype models, automatic generation of metadata for very large databases, and natural language search support.

ACKNOWLEDGMENTS

This work is supported by the NIH grant R44 LM06520-02. The authors would like to thank Dr. Roger C. Wagner from the Department of Biological Sciences, University of Delaware and Dr. Fred E. Hossler from the Department of Anatomy and Cell Biology, East Tennessee State University for their permission to use the tissue images.

REFERENCES

- 1. K. Koperski, G. Marchisio, S. Aksoy, and C. Tusk, "VisiMine: Interactive mining in image databases," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, (Toronto, Canada), 2002.
- 2. S. Santini, Exploratory Image Databases: Content-Based Retrieval, Academic Press, 2001.
- C.-R. Shyu, C. E. Brodley, A. C. Kak, and A. Kosaka, "ASSERT: A physician-in-the-loop content-based retrieval system for hrct image databases," *Computer Vision and Image Understanding, Special Issue on Content-Based Access of Image and Video Libraries* 75, pp. 111–132, July/August 1999.
- 4. P. J. Neal, L. G. Shapiro, and C. Rosse, "The digital anatomist structual abstraction: A scheme for the spatial description of anatomical entities," in *Proceedings of American Medical Informatics Association Annual Symposium*, (Lake Buena Vista, FL), November 1998.
- L. H. Tang, R. Hanka, H. H. S. Ip, and R. Lam, "Extraction of semantic features of histological images for content-based retrieval of images," in *Proceedings of SPIE Medical Imaging*, 3662, pp. 360–368, (San Diego, CA), February 1999.
- L. H. Tang, R. Hanka, H. H. S. Ip, and R. Lam, "Semantic query processing and annotation generation for content-based retrieval of histological images," in *Proceedings of SPIE Medical Imaging*, (San Diego, CA), February 2000.
- E. G. M. Petrakis and C. Faloutsos, "Similarity searching in medical image databases," *IEEE Transactions* on Knowledge and Data Engineering 9, pp. 435–447, May/June 1997.
- W. W. Chu, C.-C. Hsu, A. F. Cardenas, and R. K. Taira, "Knowledge-based image retrieval with spatial and temporal constructs," *IEEE Transactions on Knowledge and Data Engineering* 10, pp. 872–888, November/December 1998.
- G. Koepfler, C. Lopez, and J. M. Morel, "A multiscale algorithm for image segmentation by variational method," SIAM Journal of Numerical Analysis 31, pp. 282–299, 1994.
- 10. R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, John Wiley & Sons, Inc., 2000.
- 11. G. J. McLachlan and T. Krishnan, The EM Algorithm and Extensions, John Wiley & Sons, Inc., 1997.
- 12. R. C. Wagner and F. E. Hossler, "Mammalian Histology B408, Department of Biological Sciences, University of Delaware." http://www.udel.edu/Biology/Wags/histopage/histopage.htm.