

Map-Based Microscope Positioning

Grigory Begelman, Michael Lifshits, Ehud Rivlin
Technion, Computer Science Department,
Haifa, Israel
{gbeg,protezhe,ehudr}@cs.technion.ac.il

Abstract

In microscopy, regions of interest are usually much smaller than the whole slide area. Various microscopy related medical applications are liable to benefit greatly from microscope auto positioning in previously defined regions of interest. In this paper we present a method for image-based auto positioning on a microscope slide. The method is based on localization of a microscopic query image using a previously acquired slide map. It uses geometric hashing, a highly efficient technique drawn from the object recognition field. The algorithm exhibits high tolerance to possible variations in visual appearance due to slide rotations, scaling and illumination changes. Experimental results indicate high reliability of the algorithm.

1 Introduction

Microscopic digital imagery is increasingly seen as a powerful tool in many medical related fields. In diagnostic pathology, for instance, an examined specimen is mounted on a glass slide for microscopical examination. Under high magnification a typical specimen can provide tens of thousands of distinct images. Because of the limited field of view of microscope's eyepiece we only see a single image per view.

Various medical procedures that are related to microscopic tissue examination and utilize an automated microscope system (e.g. telepathology) would obviously be more efficient by the ability to find previously defined regions of interest on a slide.

Our algorithm introduces the ability to find the location of a particular region of interest on a slide in fractions of a second. The search is accomplished using geometric hashing, a well known object recognition technique. First, a map of geometric objects from the whole slide is efficiently stored in a hash look-up table. Then, accurate localization is achieved by the non-sequential search of the query image on the map, which is accelerated by indexing.

Several practical applications of the algorithm are obvious. The routine of pathological slide preparation includes slide pre-examination by a cytotechnician. The cytotechnician locates the regions of interest on a slide and marks them with a pen for further examination by an expert pathologist. The marks can be inaccurate, and moreover the procedure is invasive and time consuming. Moreover, it is impossible to mark features using a high magnification level due to the approach's very limited precision. The system described in this paper allows practitioners to put "virtual marks" on a slide, which are easily and automatically located at any magnification level.

Another possible application is the facilitation of computer aided diagnostics (CAD) systems. These systems provide diagnosis based on slide content, which is subsequently approved by an expert pathologist. While it is technically infeasible to store an image of the whole slide, individual images of key regions may be stored. A pathologist may wish to examine these key regions on the slide, previously marked by the CAD system as suspect. The ability to automatically locate their correct positions on the slide will greatly increase the efficiency of the pathologist's examination routine.

Our auto positioning system may be also applied in the area of medical education. An educational slide is examined by an expert and numerous descriptions associated with certain places of interest on the slide are recorded in the database. Later, students reviewing this slide may, by providing a software system with an image of a particular spot on the slide, get a full description of that spot.

Positioning accuracy, in our algorithm, is not affected by slide rotation and displacement, which are likely to happen as it is impossible to place the slide exactly in the same manner on another microscope. Moreover, the algorithm overcomes changes in lighting conditions or partial degradation of slide quality. The algorithm is inherently robust to rotations, scale changes and partial pattern obliteration. Likewise, the features we employ are insensitive to variations of light conditions.

The paper is organized as follows: Section 2 describes our approach to the positioning problem. In Section 3 the extraction of feature points is explained. Experimental results obtained using a slide positioning system based on the proposed algorithm are presented and analyzed in Section 4. Section 5 summarizes our work.

2 Proposed Approach for Positioning

Our approach for positioning on a slide is based on localization of small microscopic sub-images using a previously acquired slide map. The algorithm uses advanced image processing techniques and geometric hashing [4, 5, 6], a highly efficient technique drawn from the object recognition field.

2.1 Localization Problem as an Object Recognition Task

Object recognition is a known problem in the computer vision field. Recognition is achieved by finding the correspondence between a given object and a set of predefined objects. In the model-based object recognition approach, the descriptions of previously known objects are prepared in terms of various properties, such as shape, color, etc. These descriptions are referred to as "models". A given query object will be matched to one of these models.

We refer to a partial image of a slide (e.g. the current field of view of the microscope imaging system) as the "region of interest" (ROI). Localization on a slide is defined in the following manner: given an ROI on a slide, determine its exact coordinates on the slide map. Accordingly, map-based localization can be interpreted as model-based object recognition as follows. First, the slide map is constructed from partial images captured by a microscope imaging system moving over the surface of the slide. The slide map can be divided into adjacent parts that will be identified during ROI localization. The map parts correspond to a model set in the object recognition framework and the ROI plays the role of query object. Matching the current ROI to one of the previously constructed parts of

the slide map during localization is essentially the same as associating a query object to a known model in object recognition. The example of the ROI and corresponding part of the slide map is shown in Fig.1.

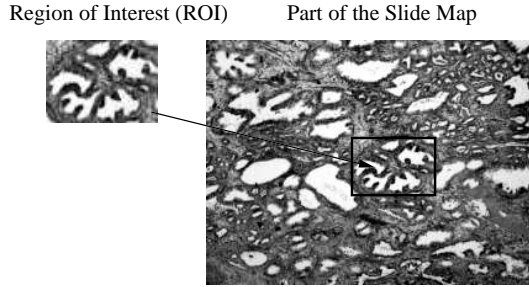


Figure 1: Example of the real region of interest (ROI) on a slide map.

2.2 The Localization Algorithm

In order to handle the enormous amount of geometric shapes contained in biological images, we choose to address the positioning problem using an efficient technique from the object recognition field, called geometric hashing [4, 5, 6]. ROI localization is performed by applying object recognition in the following way. We assume a set of predefined geometric models M_1, \dots, M_n , defining a slide map, and a query ROI image Q , formed from one of the model. The task is to find the model M_i , corresponding to the given query ROI Q .

It is assumed that the models are defined by a set of geometric features (e.g., object skeletons endpoints) and that the same features can be extracted from the query ROI image. A model can undergo similarity transformations to form the ROI: it can be rotated, translated and uniformly scaled. One way to make feature points invariant under this class of transformations is to represent them in the coordinate frame formed by the points themselves. For example, we may arbitrarily choose an ordered pair of model points to form a basis and describe the rest of the features in this coordinate frame. As there are multiple ways to choose a basis, we are faced with a combinatorial problem of finding the right one to match a model to the ROI.

The algorithm copes with this problem by shifting the computational burden to the off-line learning stage. Instead of going over all feasible ROI/model bases couples and trying to match them, all possible model representations are prepared in advance and stored in a hash table for efficient access. Thus, a query ROI projected onto an arbitrarily chosen basis has a matching model representation already stored in the hash table.

Assuming that the model M_i contains N_i feature points, there are $\binom{N_i}{2}$ different bases for that model. To form a transformation-invariant model representation, the invariant coordinates (α, β) are computed using each one of these bases $B_{\mu\nu} = \{\mathbf{m}_\mu, \mathbf{m}_\nu\}$ for every other model point. The corresponding entry $(M_i, B_{\mu\nu})$ is stored in the hash table with index (α, β) .

When analyzing the ROI during localization, the same invariant representation is used as an indexing key to access the hash table and vote for the possible model matches

(entries $(M_i, B_{\mu\nu})$ stored in the accessed bins). The model M_i accumulating a significant number of votes indicates the correspondence of the current ROI to that model. The outline of the localization process is shown in Fig. 3. This scheme demands low on-line complexity, which determines the actual time for localization. It depends linearly on the number of features contained in the ROI and is independent of the number of models stored in the system, thereby allowing fast positioning even on very large scale maps. The technique successfully deals with various possible visual transformations such as 2D rotations, translations and uniform scale.

2.3 Verification

The localization algorithm is completed by verification. Given a set of candidate models that have accumulated the highest number of votes, one has to determine which best matches the query ROI. To form the ROI, remember, the models have undergone a similarity transformation, which is a composition of translation, rotation and isotropic scaling. Thus, fitting a model to an ROI should be done by a similarity transformation estimation. The ROI is characterized in terms of a feature points set $\{\mathbf{x}'_i\}$ in \mathbb{P}^2 , and each of the candidate matching models is likewise described by its feature points $\{\mathbf{x}_i\}$, where \mathbb{P}^2 is a projective space. It is essential to find all $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ point correspondences to compute a similarity transformation H_S , which transforms a model into the ROI: $H_S \mathbf{x}_i = \mathbf{x}'_i$ for each i . Two correspondences are enough to fully constrain H_S , as the total number of degrees of freedom for similarity is four (one for the rotation, two for the translation and one more for scaling) and every correspondence gives rise to two independent equations in the entries of H_S . However, since the locations of points in the query ROI are not exact (due to noise), all of the correspondences should be used to determine the best transformation, given the data. Accordingly, H_S is calculated by finding the least-squares solution of the over-determined linear system.

An important issue is how to efficiently find all of the $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ point correspondences. The voting stage of the algorithm provides one corresponding basis (two point-to-point correspondences) between the candidate model and the ROI. This allows us to approximate the desired transformation H_S by \widehat{H}_S and then, after applying \widehat{H}_S on the candidate model, every model point $\widehat{H}_S \mathbf{x}_i$ will correspond to the closest ROI feature \mathbf{x}'_i . Formally,

$$\mathbf{x}'_i = \arg \min_k d(\mathbf{x}'_k, \widehat{H}_S \mathbf{x}_i),$$

where subindex k indicates any ROI feature and $d(\mathbf{x}, \mathbf{y})$ are the Euclidian distance between two points \mathbf{x} and \mathbf{y} .

Thus, to compute all of the point correspondences we only need to check the distance of each point \mathbf{x}'_i to every transformed model point $\widehat{H}_S \mathbf{x}_i$. If the model contains m points and the ROI contains n points, these inter-set distances are computed in $O(mn)$ time. This computation can be accelerated by employing a *Voronoi tessellation* [7] for segmentation of the ROI image. Voronoi tessellation is partitioning of a plane with n points into n convex polygons such that each polygon contains exactly one point and every point in a given polygon is closer to its central point than to any other.

We start the verification by constructing the Voronoi tessellation from the points in the query ROI, which is done in $O(n \log(n))$ time [7] (see Fig. 2). This allows us to find the corresponding point of \mathbf{x}_i in $O(\log(n))$ by checking which polygon within the Voronoi

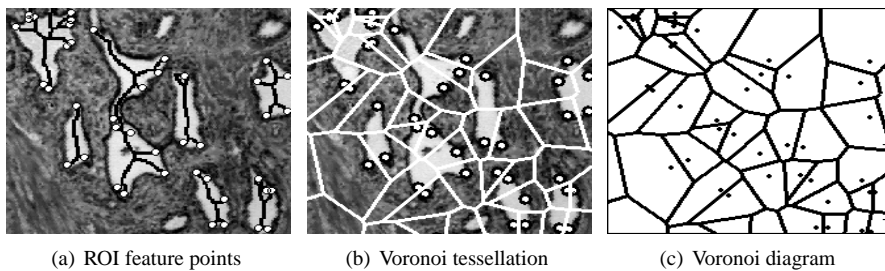


Figure 2: The process of constructing the Voronoi tessellation of the ROI for verification acceleration.

tessellation contains the transformed point $\hat{H}\mathbf{x}_i$ and choosing its center point. It follows that the time needed for point correspondences calculation is reduced from $O(mn)$ to $O(m\log(n))$.

In practice, the situation is complicated by the fact that some ROI feature $\{\mathbf{x}_i\}$ might be mistakenly reported and will not match any model point. The mismatched points, outliers, can severely disturb the estimated transformation, and consequently, should be identified. In order to make the verification robust to outliers, one has to obtain a big enough set of inliers from the presented correspondences so that the transformation can be re-estimated in an optimal manner. This is done by the RANSAC algorithm [8].

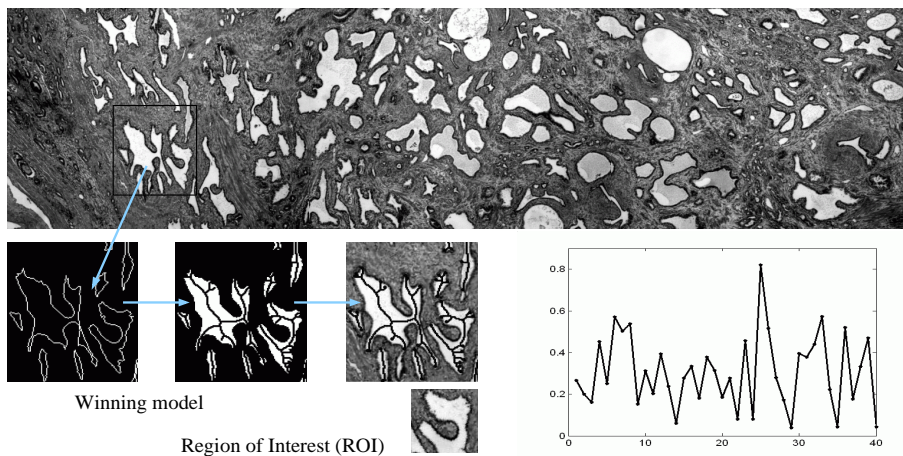


Figure 3: Outline of the localization process. A part of the slide map that is constructed from 40 model images is shown on the top. A histogram of the voting results for the query ROI, shown on the bottom in the middle are plotted on the right. The enlarged image of the winning model (25) and the process of its feature points formation is presented on the left. First, glands are found with the robust segmentation algorithm (explained in the following section) and skeletons are constructed from their shape. Then skeleton endpoints are used as the feature points for localization.

3 Feature Points Extraction

We experimented with microscopic histological prostate slides captured at zoom $\times 20$ (see Fig.1 for an example of the captured images). The images contain rich color, texture and shape information. However, due to possible variations in the lightning conditions and orientation of the slides, we are limited in selection of features to the features that are not affected by these changes. For the current zoom level we use the glands' shape information for positioning.

In order to extract the shape of the glands from the microscopic images of prostate tissue we perform segmentation. This is a challenging problem for the data set we use due to nonuniform illumination, sensor noise and various artifacts. Moreover, there is color variations between slides as a result of color fading in the stained tissues. In order to segment the images automatically, we model the distribution of colors as a mixture of three Gaussians:

$$f(x|\xi) = \sum_{i=1}^3 \alpha_i f_i(x|\xi_i),$$

where x is a feature vector, α_i represents the mixing weights ($\sum_{i=1}^3 \alpha_i = 1$), ξ represents the collection of parameters ($\alpha_1, \alpha_2, \alpha_3, \xi_1, \xi_2, \xi_3$), and f_i is a multivariate Gaussian density parameterized by ξ_i . Each component of the mixture corresponds to one of three object types: nuclei, glands, and stroma.

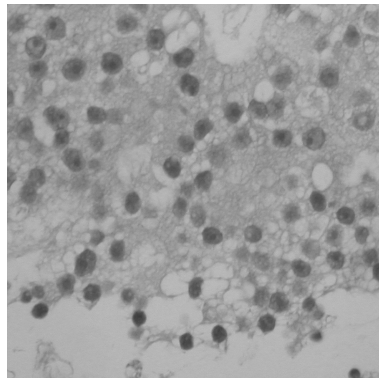
We utilize prior information about object colors in order to identify the components of the Gaussian mixture. This information is based on the hematoxylin and eosin slide staining process: the stroma areas are colored pink, the nuclei are colored blue, and the glands remain unstained (transparent).

We use the Expectation-Maximization algorithm [10] to determine the maximum likelihood parameters estimates of f in the feature space (we suppose that there are three kinds of objects in the images: stroma, glands and nuclei). Having estimated the model's parameters, we assign each point to one of three clusters according to the estimate of the point's membership based on its color features. At this point a morphological cleaning algorithm is applied to the segmented image. As a result we get three types of areas: stroma, glands and nuclei (for an example of segmentation, see figure 4).

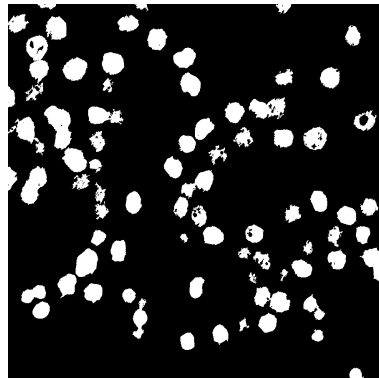
The next step is to extract shape information from the detected glands. There are many approaches to extracting shape representation (see [11] for a comprehensive survey). An important and essential criterion for shape representation is invariance to translation, scaling and rotation. Therefore, we use topology skeletons [12] as shape descriptors of the glands.

Skeletonization is a process for reducing foreground regions in a binary image to a skeletal remnant that largely preserves the extent and connectivity of the original region while throwing away most of the original foreground pixels. The skeletons can be produced in the following way. First, calculate the distance transform of the image; then, find all the pixels lying along the singularities (i.e., curvature discontinuities) in the distance transform. Examples of the calculated skeletons for prostate glands are shown in Fig. 5.

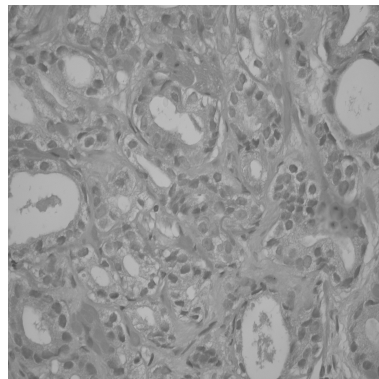
We select the endpoints of the computed skeletons as feature points for our experiments (see examples of segmented glands, skeletons of the glands and their extracted feature points in Fig. 2).



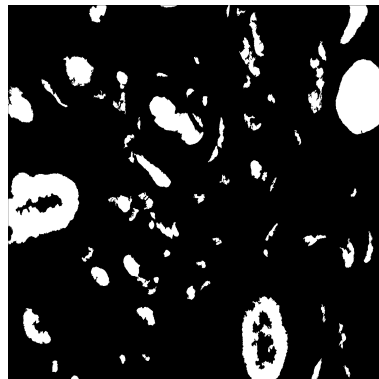
(a) A sample of prostate tissue



(b) Segmented nuclei



(c) A sample of prostate tissue

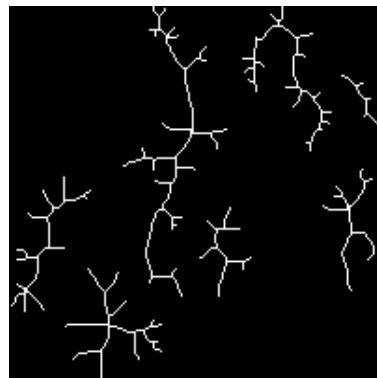


(d) Segmented glands

Figure 4: Examples of prostate image segmentations for feature extraction



(a) Segmented glands



(b) Calculated skeletons

Figure 5: Examples of skeletons calculated for glands.

4 Experiments

We now show the capabilities of the proposed localization algorithm and provide an evaluation of its performance and efficiency. We performed tests on the map covering an area of $10\text{mm} \times 4\text{mm}$ on the surface of a prostate slide. We used adjacent microscopic images obtained on Nikon Eclipse E600 microscope (zoom $\times 20$) with a Point Grey CCD camera to create the slide map. In order to index the hash table during localization and vote for the correct model, an invariant description of the query ROI is calculated. This description is based on a pair of features that form a basis, as illustrated in Section 2.2. In practice, it is possible that one of the points used to form a basis will be reported by mistake and, as a result, not match any model point. Moreover, possible inaccuracies in the basis point locations (induced by the noise) may have a similar effect. In order to avoid this, many multiple attempts should be made using different bases (e.g. different descriptions), to ensure with sufficiently high probability that at least one of them is free of extremely noisy points or outliers.

To obtain a statistically meaningful measure of the algorithm performance we tested it on a total of 10^4 different ROI localization tasks. We varied the number of different ROI feature bases being used in voting, and evaluated the localization performance with different levels of added Gaussian noise. Each time we selected a random ROI and then, if the correct location on the slide map was reported by the algorithm (ground truth was available due to the nature of data set formation), the result was considered to be true positive (TP). The result was correspondingly considered a miss if an incorrect or no location was detected.

The summary of the obtained results is presented in Fig. 6. The hit rate $\left(\frac{\#TP}{\#Tests}\right)$ was

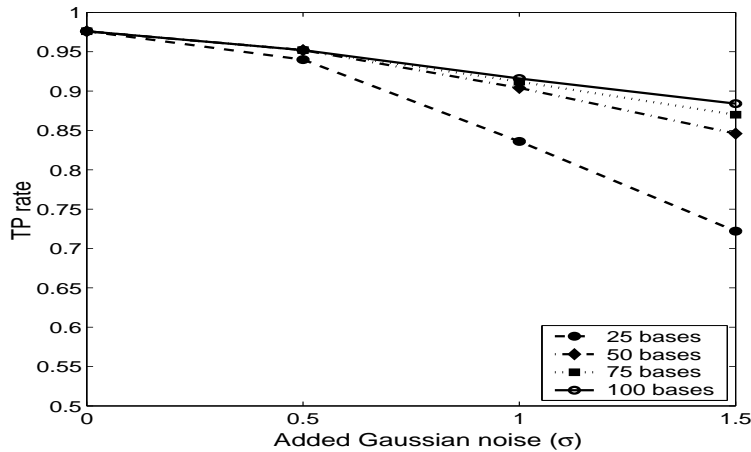


Figure 6: System behavior with a different number of bases used in voting and different levels of added Gaussian noise.

above 90% for a moderately noisy level ($\sigma < 1$) when fifty bases are used.

We formulated the accuracy of the localization result as follows. Assuming the ROI features are measured with a Gaussian error of standard deviation σ , it can be shown that [9]: the root mean square (RMS) distance of the estimated point location from its true

value is: $\sigma(2/n)^{1/2}$, where n is the number of correspondences used. Therefore, for 50 sample points the algorithm estimation error is 0.2 pixels, taking $\sigma = 1$. As a result if the ROI image of size 200×200 pixels captures an area of 40×40 micron, our algorithm provides a positioning accuracy of $0.04 \mu\text{m}$.

5 Summary

We presented a novel method for image-based positioning on microscopic slides, based on geometric hashing. Precise positioning is achieved using invariant features extracted from the specimen representative objects. It can greatly increase the efficiency of the pathologist's remote examination routine as well as assist in computer-aided diagnosis. We showed how verification can be significantly accelerated by applying a Voronoi tessellation of the ROI image. Experimental analysis demonstrates the high reliability of the proposed method.

Acknowledgements. The authors would like to thank Michael Pechuk for suggestions on this work.

References

- [1] S. S. Cross, T. Dennis, and R. D. Start, Telepathology: current status and future prospects in diagnostic histopathology, *Histopathology*, 41:91-109, 2002.
- [2] K. Glatz-Krieger, D. Glatz, and M. J. Mihatsch, Virtual slides: high-quality demand, physical limitations, and affordability, *Hum. Pathol.*, 34(10):968-74, 2003.
- [3] B. Wiley, L. Matz, Use of the AutoPap as a Primary Automated Cervical Cancer Screening System, *Medical Journal Australia*, 174:3:151-2, 2001.
- [4] J. Schwartz and M. Sharir, Identification of partially obscured objects in two and three dimensions by matching noisy characteristic curves, *Int. Journal of Robotics Research*, vol. 6, no. 2, pp. 20-44, 1986.
- [5] Y. Lamdan and H. J. Wolfson, Geometric hashing: a general and efficient model-based recognition scheme, Proc. *IEEE Int. Conf. on Computer Vision*, Tampa, FL, USA, pp. 238-249, June 1988.
- [6] Y. Lamdan, J. T. Schwartz, and H. J. Wolfson, Affine invariant model-based object recognition, *IEEE Trans. Robotics and Automation*, vol. 6, no. 5, pp. 578-589, October 1990.
- [7] M. V. Kreveld, M. Overmars, O. Schwarzkopf, and M. De Berg, *Computational Geometry*, 2nd ed. Berlin: Springer Verlag, 2000, ch. Voronoi Diagrams: The Post Office Problem, pp. 147-163.

- [8] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [9] R. I. Hartley and A. Zisserman *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.
- [10] A. Dempster, N. Laird and D. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. Royal Statistical Soc.*, vol. 39, no. 1, pp. 1-38, 1977.
- [11] S. Loncaric, A survey of shape analysis techniques, *Pattern Recognition*, vol. 31, no. 8, pp. 983-1001, 1998.
- [12] P. Golland, W. Eric and L. Grimson, Fixed Topology Skeletons, *In Proc. CVPR*, pp. 10-17, 2000