Face Recognition: A Critical Look at Biologically-Inspired Approaches

Gita Sukthankar

CMU-RI-TR-00-04

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

January 2000

Abstract

This paper analyzes the merits of two biologically-inspired face recognition models, eigenfaces and graph-matching, in the context of related neurophysiological and psychophysical data. Given the ambiguity of current biological evidence, a more promising direction for future face recognition research is in the development of models that conform more closely to human perception of facial similarity.

Introduction

For both neuroscientists and computer scientists, face recognition is a fascinating problem with important commercial applications such as mug shot matching, crowd surveillance, and witness face reconstruction. Physiological evidence indicates that the brain possesses specialized face recognition hardware in the form of face detector cells in the inferotemporal cortex and regions in the frontal right hemisphere; impairment in these areas leads to a syndrome known as *prosapagnosia*. Interestingly, prosapagnosics, although unable to recognize familiar faces, retain their ability to visually recognize non-face objects. In computer vision, some of the most popular face recognition algorithms (eigenfaces, graph-matching) have been biologically motivated. Using these models, researchers can quantify the similarity between faces; images whose projections are close in face space are likely to be from the same individual. We can compare the results of these models with human perceptions to determine whether distance in face space corresponds to the human notion of facial similarity. Although the biologically-inspired models are very useful for neuroscientists, ultimately, when building a commercial face recognition system, one should use the algorithm with the highest performance, regardless of biological relevance. However, for specialized applications, such as witness face reconstruction, in which human perception of similarity is relevant to the task, models developed using human psychophysical evidence might outperform other algorithms. This summary focuses on the following five papers:

"Face-Selective Cells in the Temporal Cortex of Monkeys": A critical analysis of the evidence for face neurons, focusing on neurophysiological experiments performed in monkeys [1].

"Eigenfaces for Recognition": One of the most popular biologically-inspired face recognition systems, based on the principal component analysis algorithm (PCA) [2].

"Face Recognition by Elastic Bunch Graph Matching": A different biologically-motivated algorithm which relies on features extracted from the intensity image, rather than a holistic analysis of the face [3].

"A Comparison of Two Computer-Based . . . Systems with Human Perceptions of Faces": A study comparing the performance of eigenfaces and the precursor to the elastic bunch graph matching system with human ratings of similarity [4].

"High-Performance Memory-based Face Recognition for Visitor Identification": A simple nearest neighbor algorithm for face recognition that outperforms traditional algorithms such as PCA, on the standard datasets [5].

Neurophysiological Evidence on Face Recognition

A selection of interesting research in the field of face recognition was published in a 1991 special issue of the *Journal of Cognitive Neuroscience*. In that issue, Robert Desimone presented a critical review of the neurophysiological evidence for face detector cells and discussed how these results relate to biological models of face detection [1]. The first face detection cells were discovered by Gross *et al.* [6] in a survey of the monkey inferior temporal cortex. Along with cells that seemed to respond best to faces, they reported a neuron tuned to detect hands, specifically monkey hands. This discovery reawakened interest in Konorski's gnostic units [7], more popularly known as "grandmother cells". In Konorski's theory of visual recognition, the outputs of simple edge and line detectors are used as the inputs to a pyramid of higher level cells, to eventually form complex cells that respond very selectively to a small range of inputs. At the apex of this pyramid could be a "grandmother cell" that would fire only when you saw your grandmother. By contrast, the prevailing theory for object recognition in the brain predicts that objects are encoded as changes in the firing pattern across a population of cells, and that these cells respond to a wide range of stimuli.

After Gross's initial finding, neurons responding preferentially to facial stimuli were reported by several other labs in both awake and anesthetized monkeys. Moreover these face detector cells did not respond to simple stimuli (e.g., textures, bars, gratings) nor to other complex objects, including those designed to evoke emotional responses. Some of these cells responded to simple line drawings of faces, but did not respond well to faces with rearranged facial features, even if all of the features were present. Face cells continued to respond to faces subjected to high-pass filtering, low-pass filtering, and size scaling; color shifting reduced but did not eliminate response. Considered in total, this body of evidence fairly conclusively substantiates the existence of face detector cells in the primate brain. Researchers have also reported cells tuned to isolated facial features, such as eyes, but it is more difficult to verify that these cells are specifically tuned to eye detection, rather than to a more general class of shapes, such as ellipses. Altering the placement of facial feature does affect the response of face cells, especially changing the interocular distance, the distance between eyes and mouth, or the amount of forehead hair. Although frontal face detector cells are the most common, certain cells respond preferentially to rotated or profiles views; interestingly, two of the computer face recognition systems discussed below also make use of rotated views to improve performance.[1]

Face detector cells have been reported in a wide variety of locations, spanning all regions of the inferior temporal gyrus and the superior temporal sulcus. Cells are most commonly found in the sulcus, which can be separated into two areas based on location and connectivity: a superior temporal polysensory area (STP) and the inferior temporal cortex (IT). These two distinct populations of face detector cells may contribute to different circuits, one honed for establishing identity and the other for recognizing facial expression.

Discussion

Desimone speculates that face recognition is an extremely important task for primates, who use facial expressions for communication. Accurate face recognition enables primates to identify important individuals in their social structures and subsumes the role of olfaction, which is a salient cue that other animals use for recognition. Thus, the existence of specialized biological "hardware" for face recognition is both plausible from an evolutionary point of view and substantiated by neurophysiological evidence.

Unfortunately, this collection of studies neither substantiates nor disproves any of the computer models for face recognition discussed here. Two of the systems, eigenfaces [2] and feature-based graph matching [3], have been specifically promoted as potential biological models. Some of the neurophysiological results directly contradict experiments performed in computer models; for instance, the eigenface method is very sensitive to face scaling whereas face cells have been demonstrated to be robust to scale changes in a 12 fold range. However, eigenfaces of different scales (see below) could be added to compensate for this weakness

[1] Turk & Pentland propose incorporating different views of the same face into their covariance matrix; also, ARENA makes use of additional, transformed exemplars.

in the original model. ARENA [5], a nearest-neighbor based face classifier, relies on synthetically generated facial views in different orientations to supplement training exemplars—do face detector cells tuned to non-frontal views perform the same role in the primate brain? The placement of internal facial features is critical for biological face detection which seems to suggest the validity of feature detection models, such as those present in the graph matching approach. Since so much of the brain is uncharted territory, assuming the existence of pre- or post-processing can be an *ad hoc* method to shore up the weaknesses in any biological model. Until we know enough about the brain to confirm or deny the existence of these other processing stages, the biological evidence clouds the issue of face recognition, as much as it illuminates it.

Eigenfaces Algorithm

Prior to Turk and Pentland's paper, "Eigenfaces for Recognition", many researchers focused on detecting individual facial features and categorizing different faces by the position, size, and relationship of these features. In contrast, Turk and Pentland take an information theoretic approach to the problem and derive their classification features with principal component analysis which extracts the optimal linear encoding of the dataset.[2] In a cluster of image vectors, the dimension with the greatest variance is captured by the eigenvector with the highest eigenvalue; thus when projecting a point into the eigenvector basis and examining the value of its coefficients, the most information about the point can be gained from the coefficient of the top eigenvector. Eigenvector methods rely on being able to classify images based solely on the coefficients of the top k eigenvectors. This technique is also exploited by Shree and Nayar for the general object recognition problem of classifying images of 3-D models with variable illumination and viewpoint. Turk and Pentland note that eigenvectors can be calculated using an $M \times M$ matrix, in which M is the number of images, rather than using the original covariance matrix which has dimensionality $N \times N$, where N is the number of pixels in the image. Typically $M \ll N$, so their technique of premultiplying the covariance matrix is computationally more efficient.

Turk and Pentland use the following technique to initialize their system. After acquiring an initial training set of images, they calculate the eigenfaces (eigenvectors) for the matrix and use those to define a "face space". Each individual is represented in this face space by averaging the coefficients extracted from labeled exemplars of his/her face.[3] Once the system is initialized, images can be classified by projecting them into the eigenvector basis, using only the top k vectors with the highest eigenvalues. The projection is achieved simply by taking the dot product of the image with each eigenvector; this computation can be performed with a three layer network in which weights are the eigenvectors and the hidden layer's outputs are the coefficients. The final output of the network is an image, quite similar to the original image, that has been reconstructed from the information stored in the coefficients of the top k eigenvectors. Non-face images are not represented well by this basis set so the final image will be noticeably different from the original image, whereas a face passed through the network will have almost identical inputs and outputs.[4] After non face images are discarded, the test image is assigned the label of the closest class mean. If the Euclidean distance between the test image and the closest mean is above a certain threshold it is classified as an unknown face which can be added to the dataset later. New faces can be used to modify the eigenfaces, which unfortunately involves recalculating all the eigenvectors and the coefficient means.

[2] The encoding is optimal for minimizing the sum square error of a reconstructed image.

[3] This is the original method proposed by Turk & Pentland in [2]. Variations on this approach are discussed later.

[4] Note that this dissimilarity arises because only the coefficients of the top eigenvectors are considered; if all the non-zero eigenvectors are used even a non face image will be reconstructed perfectly.

Turk and Pentland test how their algorithm performs in changing conditions, by varying illumination, size, and orientation of the faces. They found that that their system had the most trouble with faces scaled larger or smaller than the original dataset. To ameliorate this problem, they suggest using a multi-resolution method in which faces are compared to eigenfaces of varying sizes to compute the best match. Also they note that image background can have a significant effect on performance, which they minimize by multiplying input images with a 2-D Gaussian to diminish the contribution of the background and highlight the central facial features. Their system performs face recognition in real time and they also use their method, along with motion cues, to segment faces out of images by discarding squares that are classified as non-face images.

Discussion

Turk and Pentland's paper was very seminal in the field of face recognition and their method is still quite popular due to its ease of implementation. Also their algorithm can be plausibly implemented as a biological network, and possesses some of the characteristics of human face recognition such as high recognition speed and robustness to occlusion. Although their system can be fooled by short term changes, such as variations in facial hair or hairstyle, long term facial changes due to aging can be handled if one assumes that new training images are periodically added.

However, there are some important fallacies about eigenvector methods that this paper does not address. Turk and Pentland suggest that projecting facial images into the eigenvector basis makes the image classes more separable because the eigenvectors represent the dimensions of maximum variance. Although this projection efficiently spreads the points, it does not consider how the points are assigned to classes. Ultimately, we want to project the images such that the *classes* can be separated rather than the points; even if the points are spread along an axis, one can have the undesirable situation of a point on one end of the cluster belonging to the same class as points on the other end of the cluster. Averaging the coefficients of these points will give a misleading class mean in the middle of the cluster. In another variant of PCA, the coefficients of all the exemplars are retained, instead of just the mean coefficients. This variant is more successful at handling non convex cases such as the one described in the previous example. Methods such as Fisher's linear discriminant analysis take class labels into account when re-projecting the points and have been successfully applied to the face recognition problem (Fisher Faces [8]). Although PCA provides a linear projection that may work for some recognition problems, there is no reason to believe that it is the best projection for recognition merely because it works well for encoding.

Feature Based Recognition: Elastic Bunch Graph Matching

The other dominant trend in face recognition has been feature matching: deriving distance and position features from the placement of internal facial elements. Kanade [9] developed one of the earliest face recognition algorithms based on automatic feature detection. By localizing the corners of the eyes, nostrils, etc. in frontal views, the system computed parameters for each face, which were compared (using a Euclidean metric) against the parameters for known faces. A more recent feature-based system, based on elastic bunch graph matching, was developed by Wiskott *et al.* [3] as an extension to their original graph matching system (see next section). Although their system isn't strongly biologically grounded, they detect features using Gabor kernels, which are believed to simulate the output of simple cells in the V1 cortical area.

Wiskott *et al.* present a general in-class recognition method for classifying members of a known class of objects. Faces are represented as graphs, with nodes positioned at fiducial points, and edges labeled with 2-D distance vectors. Each node contains a set of 40 complex Gabor wavelet coefficients, including both phase and magnitude, known as a *jet*. Wavelet coefficients are extracted using a family of Gabor kernels with 5 different spatial frequencies and 8 orientations; all kernels are normalized to be zero mean. To identify a new face, an object-adapted graph has to be positioned on the face using elastic bunch graph matching. This can be performed automatically if the face bunch graph (FBG) has already been initialized with enough faces (approximately 70). A face bunch graph (FBG) consists of a collection of individual face model graphs combined into a stack-like structure, in which each node contains the jets of all previously initialized faces from the database. To position the grid on a new face, the graph similarity between the image graph and the existing FBG is maximized. Graph similarity is defined as the average of the best possible jet match between the new image and any face stored within the FBG minus a topography term, which accounts for distortion between the image grid and the FBG. After the grid has been positioned on the new face, the face is identified by comparing the similarity between that face and every face stored in the FBG, ignoring the phase term.[5]

Discussion

The authors report that the initial grid matching phase requires about 30 seconds on a SPARCstation 10-512, and the actual recognition process completes in less than 1 second. On the FERET dataset, the algorithm performs impressively well on the frontal images, correctly ranking the image first 98%. For half-rotated and profile images, performance degrades to 57% and 84% correct; however since this is a difficult case for face detection and recognition systems, these results are comparatively good. The authors claim that their graph-matching technique correctly handles cases where the nodes are occluded, but do not specify details.

One disadvantage of this technique is that the tedious grid placement must be performed manually for the first 70 images, before the automatic graph matching becomes sufficiently reliable. Also, the authors did not test the algorithm on faces with variable illumination and background. Intuitively, it seems that this technique should be resilient to intensity fluctuations, compared to PCA based methods (see next section).

Human vs. Computer Facial Similarity Ratings

For applications such as criminal identification, it would be useful to have a computer system that understood human conceptions of facial similarity; for instance, one can imagine wanting a database program that could retrieve similar faces from a comprehensive mug shot database after a witness selected one face from a small initial grouping.[6] Another interesting application would be the construction of "electronic lineups": displaying a photo of the real criminal along with a group of automatically-selected distractor photos to test the witness's recollection. For such applications to be successful, the program has to be able to encode the similarity metrics used by humans to group faces. Given this problem, one might wonder how successful the current biological models are at capturing humans' understanding of facial similarity. In their paper, Hancock *et al.* [4] compare the performance of eigenfaces and graph-matching face recognition methods

[5] Although the phase is useful for finely positioning the grid, omitting it proves to be beneficial in the final classification.

[6] In essence, this is a more specialized version of IBM's QBIC (Query by Image Content) [10]

with human ratings of similarity and accuracy on memory tasks. Hancock *et al*. believe that perception of the *face* should not be confounded with the memory of the *image*; thus an important theme of their experiment was to have subjects match the same faces across different images. This type of matching is slightly more difficult for PCA, which mainly captures low-order statistical regularities in the data.

Based on psychophysical evidence, Hancock *et al.* observe that it is likely that face recognition is based on low-level image properties, rather than on an abstract representation of the face. Certain image transformations, such as intensity negation, strange viewpoint changes, and changes in lighting direction can severely disrupt human face recognition. If the human visual system was actually maintaining an abstract 3-D representation of the face, it isn't clear why these transformations would be disruptive. One possibility is that a shape from shading operation is required to extract the 3-D representation; this operation could be disrupted by intensity changes. Given that the human visual system might be performing image-based face recognition, it seems reasonable to compare human face recognition to image-based biological models of face recognition. The two models examined in the paper are: (1) a modified version of PCA (eigenfaces) which uses morphed faces in which all internal features are aligned by image warping; the original positions of the features are encoded by an additional shape vector added to the image; (2) a graph matching face recognition system which relies on features extracted with Gabor-type wavelets at different scales and orientations. For the PCA system, Hancock *et al.* examined the effects of using different groups of components (those with the highest eigenvalues vs. least eigenvalues).[7] The highest correlation with human data was obtained by using the top ten components. Results from the two computer systems were compared with human data in three experiments: (1) similarity; (2) rated distinctiveness; (3) actual memory.

Hancock *et al.* presented human subjects with faces and asked them to rate how distinctive they were on a scale of 1–10; participants were asked the question, "suppose you had to meet the person at a station, how easy would it be to pick them out?". The initial faces were either chosen from the neutral (N) condition (no expression) or the expressing (E) condition (faces expressing an emotion, such as happiness or anger). Then after a short period, the subjects were presented with a new group of faces and asked whether they had seen these faces earlier. The subjects' answers were expressed as a confidence level between 1–10, with 10 signifying great confidence in having seen the face before. The different test conditions were as follows: (1) subjects initially shown neutral faces and asked to recognize expressing faces (NE); (2) subjects initially shown neutral faces and asked to recognize neutral faces (NN); (3) subjects initially shown expressing faces and asked to recognize neutral faces (EN). This section of the experiment was designed to test distinctiveness ratings and memory. In tests with the computer systems, distinctiveness was considered equivalent to confidence rating, or in the case of the PCA system, reconstruction error, which is a measure of how well the face is represented by the initial basis set. In the other section, subjects were asked to group faces based on perceived similarity. Subjects could choose to separate the faces into as many subsets as desired. Half of the subjects were given standard faces, and the others were asked to sort faces with the hair removed. For all the tests, the computer systems were initialized with the original neutral faces with hair. Over the experiments, the authors noted the following results.

1. The results of both computer systems correlated equally well with human ratings of similarity in the condition with hair, whereas the graph matching system did significantly better in the hair-free condition.

2. Both systems' confidence measures agreed with human distinctiveness ratings.

[7] In some variants of PCA, some of the components with the largest eigenvalues are discarded, because these supposedly encode gross differences between images, such as lighting variations, rather than face-specific information.

3. There was a strong correlation between human memory performance and the PCA system in the NN condition, which drops dramatically in the NE condition. The graph matching system correlates about equally well in both conditions.

Discussion

The authors conclude by noting that PCA is much more sensitive to deviations from the initial training conditions. They mention that the PCA system completely failed to recognize one of the faces, and later they realized that the image had been mirror-reversed during scanning. Neither the humans nor the graph matching system noticed this error and both were able to recognize the face. Neither system's similarity results correlated very strongly with the humans, and the authors hypothesize the existence of other metrics used by the humans which were not captured by either of the computer systems. For the standard conditions (with hair and on neutral faces), the rankings of both computer systems were very highly correlated so either of them would serve equally well for grouping images in criminal identification applications.

In hindsight, it seems fairly obvious that PCA would be more severely affected by relatively trivial changes to the image. Removing hair from the photos was very distracting for the PCA system, presumably because a large part of the image area is composed of hair pixels, whereas the graph-matching system derives more features from internal facial contours. PCA performance correlated best with human performance in the NN case, where humans were performing recognition tasks on the same image, but in this case, face and image recognition is somewhat confounded.

Visitor Identification: An Application of Face Recognition

The final paper I examine describes a real-world application of face recognition techniques—an automated visitor identification system that notifies lab researchers about guests. Building visitors are detected using an image differencing technique as they approach the front door. Faces are extracted from the image using a neural-network based face detector [11] and classified with a memory-based face recognition algorithm called ARENA [5]. ARENA matches a reduced resolution version of the image against a database of previously collected exemplars using an L_0^* similarity metric. The system is implemented as a group of cross-platform agents and includes an interface for online training that enables users to correct the recognition system whenever it misidentifies a visitor; thus, new labeled exemplars can immediately be incorporated into the database for future use. This system operates in an outdoor environment with substantial lighting variation; however, much of the background clutter and facial pose variation is eliminated by the face detection system.

The face recognition algorithm, ARENA, uses reduced-resolution versions of the images created by averaging over non-overlapping rectangular regions in the image. In experiments on the ORL dataset, algorithm performance plateaus at the 16x16 resolution. Similarity is judged using the L_0^* norm—this is intuitively equivalent to counting the number of image components whose magnitudes differ by more than a threshold value. The L_0^* norm provides better performance than the standard L_2 norm, which overly penalizes data outliers. ARENA's performance also improves when the training set is augmented with additional, synthetically generated face images created using simple geometric transformations.

With these refinements to the simple nearest-neighbor algorithm, ARENA achieves impressive performance on the standard ORL and FERET face recognition datasets: significantly better than three variants of PCA, hidden Markov models, and a self organizing map combined with a convolutional network classifier. In the more demanding task of visitor identification, the authors report accuracies of 55% for an image set containing 50 individuals and over 1000 training images; many of the faces photographed display significant out of plane rotation, occlusion, and extreme lighting conditions.

Discussion

Although no attempt has been made to provide biological motivations for ARENA, one can envision the algorithm as being one possible implementation of the gnostic unit or grandmother cell recognition theory. A stored exemplar functions as a sort of grandmother cell that only activates for one particular face; each face is actually encoded by multiple exemplars, but only one need activate for a match to be successful. The nearest neighbor method handles the non-convex domain much better than the means based classifier used in Turk and Pentland's original eigenfaces paper. Seriously speaking, there is no real biological evidence supporting the existence of an ARENA type algorithm; the storage requirements for a nearest-neighbor biological system could become quite daunting, especially if multiple exemplars of each face are stored. However, this is also a valid argument against the nearest-neighbor PCA variants. Like the elastic bunch graph matching technique, ARENA is not a face-specific recognition algorithm and could be successfully applied to other classes of objects as well.

Conclusion

Face recognition, especially in a cluttered dynamic environment, is a difficult problem; most of the published results have been obtained on static, high-quality, frontal facial images. Better algorithms are needed to overcome the problems of out-of-plane facial rotation, lighting variations, occlusion, and viewpoint changes. Many researchers have derived inspiration from the biological study of face recognition, but it is unclear whether these techniques succeed either as physiological models or as effective algorithms:

- The neurophysiological evidence is sufficiently ambiguous to permit several plausible models; with the addition of pre- or post-processing steps, almost any model can be adjusted to fit the available evidence.
- Although PCA based techniques appear computationally elegant, they suffer from the flawed assumption that reprojecting images into an eigenvector basis will improve the separability of the image *classes*.
- One of the most promising areas for computer based face recognition algorithms is the development of systems that correlate well with human ratings of similarity. Current computer algorithms, such as PCA and graph-matching, correlate well with each other, but are less good as predictors of human perception.
- Standard face image datasets are typically inadequate for measuring the true performance of algorithms, since they lack illumination and background variation. As shown by ARENA, even relatively

simple approaches, such as nearest-neighbor classifiers, can excel on such a testset.

It is interesting to compare current face recognition research with the experience of early scientists studying aviation. Some of those pioneers were inspired by watching bird flight and built their vehicles with mobile wings. Although a single underlying principle, the Bernoulli effect, explains both biological and man-made flight, we note that no modern aircraft has flapping wings. Thus, it is important that researchers look to the deeper truths without becoming distracted by the surface similarities between biological and computer models—when we study flight, we want to make sure that we're not just flapping our arms.

References

[1] R. Desimone. Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, 3(1), 1998.

[2] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991.

[3] L. Wiskott, J. Fellous, N. Krüger, and C. von der Malsbnurg. Face recognition by elastic bunch graph matching. *IEEE Trans. PAMI*, 19(7), 1997.

[4] P. Hancock, V. Bruce, and M. Burton. A comparison of two computer-based face identication systems with human perceptions of faces. *Vision Research*, 38(15–16), 1998.

[5] T. Sim, R. Sukthankar, M. Mullin, and S. Baluja. Memory-based face recognition for visitor identification. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, 2000.

[6] C. Gross, C. Rocha-Miranda, and D. Bender. Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, 35, 1972.

[7] J. Konorski. *Integrative Activity of the Brain. An Interdisciplinary Approach.* University of Chicago Press, 1967.

[8] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on PAMI*, 19(7), 1997.

[9] T. Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto University, 1973.

[10] J. Ashley, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic. The query by image content (QBIC) system. In *Proceedings of SIGMOD*, 1995.

[11] H. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1), 1998.

[12] R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5), 1995.

[13] T. Fromherz. Face recognition: A summary of 1995–1997. Technical Report ICSI TR-98-027, International Computer Science Institute, Berkeley, 1998.