CONTENT-BASED RETRIEVAL OF OPHTHALMOLOGICAL IMAGES

A. Gupta, S. Moezzi, A. Taylor, S. Chatterjee, R. Jain, M. Goldbaum, S. Burgess

University of California, San Diego
La Jolla CA 92093-0407
USA
{amarnath, moezzi, ataylor, shankar, jain}@vision.ucsd.edu
{mgoldbaum, sburgess}@ucsd.edu

ABSTRACT

This paper describes steps towards an information system for the storage and content-based retrieval of ocular fundus images. Based on the Virage Incorporated framework for defining similarity metrics, we have developed a number of primitives for the representation of ocular fundus images. A prototype Query By Pictorial Example (QBPE) system yields similarity rankings in approximate agreement with those of a human expert.

1. INTRODUCTION

More and more forms of media are moving from analog to digital technologies. The clear-cut advantages of managing textual data in a digital format have spurred this movement. A great deal of work is now being devoted to developing systems that manage diverse media with the fluidity that today's systems have brought to the management of textual data. In this paper, we describe a system which will eventually enable the content-based management of ophthalmological images of the human retina. This work is proceeding as part of the STARE (STructured Analysis of the REtina) Project at the University of California, San Diego [1]. The envisioned system will assist the user in a number of ways, allowing her to:

1. Store a large number of fundus images along with patient records, history and clinical findings in a single database.
2. Perform on-line annotations (such as labeling a bright region as an "retinal infarct") and region-of-interest (ROI) marking on the images (here is a region showing grape clusters).
3. Detect a set of visually apparent anatomical structures e.g., macular region, and abnormalities, e.g., vessel tortuosity and cotton-wool spots, using a set of built-in image processing algorithms.
4. Interactively or automatically analyze the features of any ROI (such as, compute area of an exudate to quantify its degree of severity). This would allow quantitative measurements of such things as vascular tortuosity and dilation.
5. Interactively search the database of labeled and classified images using patient records, location and severity of abnormality, and image properties (e.g., find other images with vessel narrowing similar to a given image).

The overall system is intended to be adaptable for research and clinical uses, such as repetitive image analysis tasks, quantitative measurements, image description, frequency analysis of manifestations in selected populations of diseases, and diagnostic screening.

2. CONTENT-BASED IMAGE RETRIEVAL FRAMEWORK

Content-based image retrieval can take many forms. The form we focus on in this paper is Query By Pictorial Example (QBPE). In a QBPE system, a user's query is simply an image, and the system's task is to return images that are 'similar' to the given image. The obvious problem is how to define the similarity between two images. If we assume that the notion of inter-image similarity can be represented by a *metric* defined on the space on images, then the problem of measuring image similarity can be reduced to the problem of formulating such a metric. Virage Incorporated has developed a framework for defining such 'similarity metrics,' and our system employs this framework [2].

What we desire is a function, m, from the set of all images, I, to the nonnegative real numbers, $\Re^+$. This defines the distance between two images, where images

Visiting Researcher from Virage Incorporated, 177 Bovet Road, Suite 520, San Mateo, CA 94402-3121

separated by a small distance are relatively 'similar', and images separated by a large distance are relatively dissimilar. The Virage framework inserts an intermediary layer between images and the distances between them, a layer of *primitives*. A primitive encompasses two things: First, a function from the space of all images, I, to some space of primitive values, call it P, and second, a metric defined on P, i.e. a function m_P from P to $\Re^+$ that happens to be a metric. For example, the basic Virage Engine includes four primitives: a color primitive, a composition primitive, a texture primitive, and a structure primitive. The color primitive, for instance, calculates a color primitive value that in some way summarizes the color information present in an image. Given the color primitive values for two images, the primitive calculates a metric function on the color primitive values that in some way reflects the (dis)similarity between the images as far as their color information is concerned.

Given that we have defined a set of n primitives, we can then calculate an n-dimensional 'primitive metric vector' for any two images. The question then arises as to how we combine these n quantities into a single nonnegative real number, which is the final output we require from our similarity metric. In theory, many different summaries of this data could be used as our final distance value. The Virage engine provides for weighted arithmetic means to be used for this final summary. The fact that the mean can be weighted is very helpful if one cares more about texture similarity, for instance, but still wants to take color information into account.

3. STARE PROJECT PRIMITIVES

In the Virage framework, the process of developing an image similarity metric for a particular class of images becomes the problem of defining suitable primitives for the domain. For the STARE Project, our primitives are based on the sorts of 'objects' that appear in fundus images, as well as the kind of general properties that can hold of this kind of image. For example, one type of object that may occur in a fundus image is called a 'flame hemorrhage.' This appears as a dark region with a brush-stroke-like shape. We therefore have a *flame-hemorrhage* primitive in our system, the values of which represent the location and severity of any flame hemorrhages present in the image.

3.1. Classes of STARE Project Primitives

The STARE Project primitives fall naturally into four classes:

1. Global One-Zero Primitives
2. Localized One-Zero Primitives
3. Global Severity Primitives
4. Localized Severity Primitives

3.1.1. Global One-Zero Primitives

Global one-zero primitives are the simplest of the four types. The value space for such a primitive is simply a set of k discrete members. They are called 'global' because they are based on some non-localized image property, such as whether the fovea displays a 'cherry red spot' pathology. This feature is certainly localized in the sense that it appears in the foveal region of the retina, but it is a pathology which occurs nowhere else, so all that need be said about a given image is whether it is absent or present, and in this sense it is global to the image. The 'one-zero' in the name of this class of primitives refers to the fact that the primitive metric is a one-zero metric: It takes the value zero if the two arguments are the same, the value one otherwise. This reflects the fact that the primitive space has no metric structure beyond the equality or non-equality of two elements. It happens that we have no non-binary global one-zero primitives in our current collection, but if it were possible to determine race from fundus images (it isn't), we might have a global one-zero *race* primitive, with possible values of African, Asian, Caucasian, Indian, etc.

3.1.2. Localized One-Zero Primitives

Localized one-zero primitives are used to describe features that can occur at a variety of different locations in an image, but have a one-zero metric structure modulo location information. An example of a localized one-zero primitive is the *artery-color* primitive, which, for a given location, can take on the values normal, copper, silver, sclerotic, focal inflammation, and sheathed, none of which have any clear relationship to one another besides equality or non-equality.

In the current implementation, the location of an object is represented as being one of 11 possible predefined regions of the retina. The structure of these regions, which we call the Fundus Coordinate System (FCS), is a codification of how domain experts think about location in the retina. Our 11 regions are defined relative to the center of the optic disk and the center of the fovea, and are shown in Figures 1 and 2. Thus the value space for a localized one-zero primitive will be a vector of 11 values, where each value is one of k discrete possibilities.

The primitive metric for localized one-zero primitives is rather involved, and we will not go into specifics

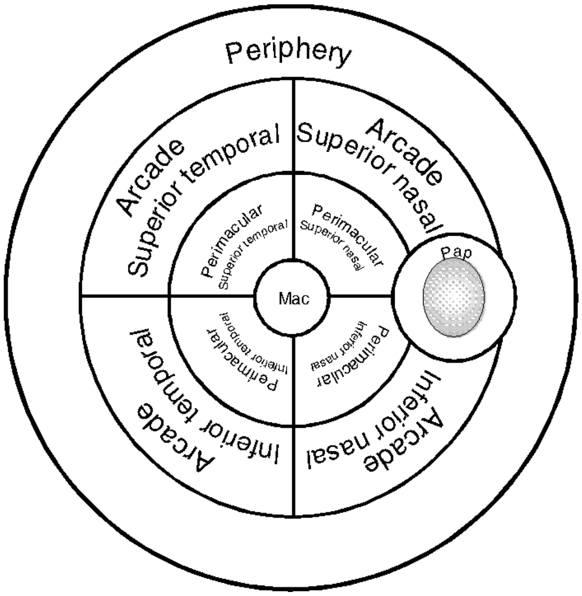

Figure 1: The Fundus Coordinate System (FCS)

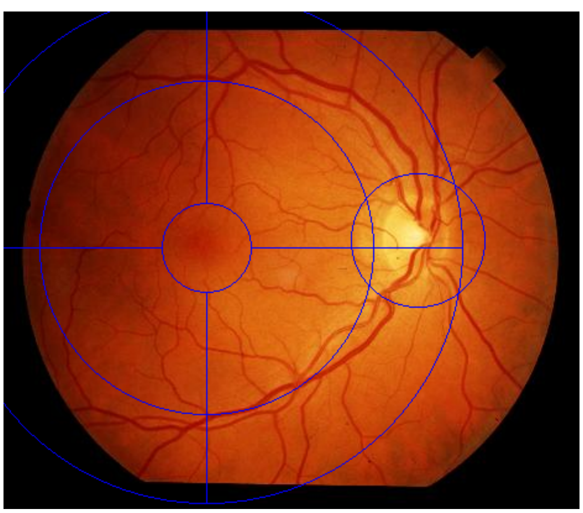

Figure 2: A fundus image overlaid with the FCS

here. It gives shorter distances for the same local primitive value at the same location, with longer distances for the same local value in neighboring locations, and even longer distances for disparate local values. The spatial relationships between locations in the FCS are represented via a simple adjacency graph.

3.1.3. Global Severity Primitives

Global severity primitives are like global one-zero primitives in that the primitive space consists of k discrete members, and that which of these values a given image maps to is based on some global image property, with 'global' understood in the correct sense. The difference, of course, is that the primitive metric is not the simple one-zero metric, but reflects a greater amount of metric structure among the primitive space members. These primitives are called severity primitives because they typically measure the severity of a given pathology. As such, it is sufficient to associate with each a point on the real line, and appropriate the usual Euclidean metric for use as our primitive metric.

3.1.4. Localized Severity Primitives

Localized severity primitives are the predictable extension of global severity metrics. The primitive space is the space of 11 dimension vectors where each vector element is taken from one of k discrete severity values.

The metric for these primitives takes advantage of the mapping from each of the k severity values to the real line, and accounts for both differences in severity as well as differences in location. We omit the details.

4. PROTOTYPE QUERY INTERFACE

We have implemented a prototype version of the described system. A set of about 150 images has been manually annotated, and we have used this set of annotated images to test the utility of our system. The query interface is implemented as a set of World-Wide Web documents. A user can browse the images and select a query image. This image is then compared against all other images in the system, and a subset of these images is returned, ranked in order of decreasing similarity. This is shown in Figure 3. The rankings returned are found to agree closely with an expert's notion of which images are similar to the query image.

5. FUTURE WORK

It should be pointed out that our demonstration system only implements half of what makes QBPE systems so potentially interesting. We have implemented a means

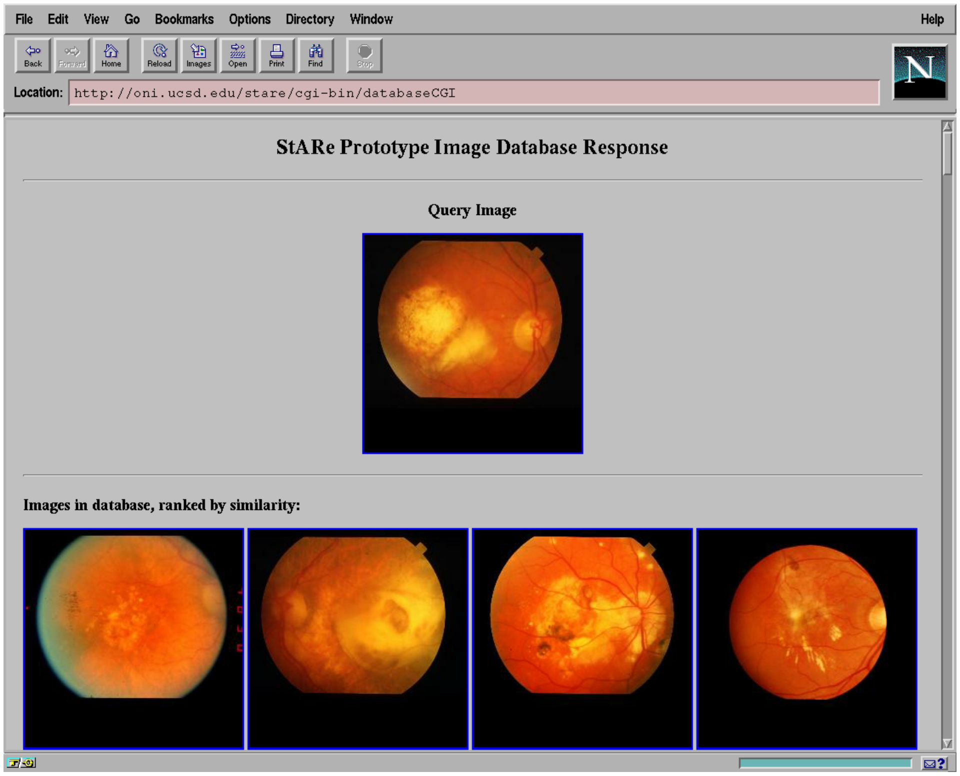

Figure 3: A ranked list of similar images

for comparing primitive values calculated from images, but we have only partially implemented the segment of the system that actually calculates these primitives from image data [3]. This portion of the system is based on segmentation and classification algorithms developed in earlier phases of the STARE Project, and we expect to be able to incorporate some amount of 'automatic annotation' very soon. We also plan to augment our current tool for performing manual image annotation so that it becomes an 'annotation workbench,' allowing the user to integrate manual and automatic image annotations.

Another avenue which we have plans to explore is 'value queries.' We imagine that an ideal fundus image information system would support both QBPE as well as more specific queries, such as "Retrieve all images showing both a large amount of non-circinate retinal exudate in the superior temporal medial quadrant as well as stellate maculopathy." We plan to support this type of query with some form of a fuzzy relational calculus, perhaps along the lines of [4].

6. REFERENCES

[1] M. H. Goldbaum, N. P. Katz, S. Chaudhuri, M. Nelson, and P. Kube. Digital image processing for ocular fundus images. *Ophthalmological Clinics of North America*, 3:447–466, 1990.

[2] Amarnath Gupta. Visual information retrieval technology, a Virage perspective. Available at http://www.virage.com/wpaper/, 1995.

[3] M. H. Goldbaum, S. Moezzi, A. Taylor, S. Chatterjee, E. Hunter, and R. Jain. Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images. In *IEEE International Conference on Image Processing*, November 1996.

[4] P. Bosc and O. Pivert. SQLf: A relational database language for fuzzy querying. *IEEE Transactions on Fuzzy Systems*, 3(1):1–17, February 1995.