

# AUTOMATED DIAGNOSIS AND IMAGE UNDERSTANDING WITH OBJECT EXTRACTION, OBJECT CLASSIFICATION, AND INFERENCING IN RETINAL IMAGES

*Michael Goldbaum, Saied Moezzi, Adam Taylor, Shankar Chatterjee, Jeff Boyd, Edward Hunter, and Ramesh Jain*

Department of Ophthalmology and Department of Engineering and Computer Science  
University of California  
La Jolla, California 92093-0946  
USA  
E-mail: mgoldbaum@ucsd.edu

## ABSTRACT

Medical imaging is shifting from film to electronic images. The STARE (structured analysis of the retina) system is a sophisticated image management system that will automatically diagnose images, compare images, measure key features in images, annotate image contents, and search for images similar in content. We concentrate on automated diagnosis. The images are annotated by segmentation of objects of interest, classification of the extracted objects, and reasoning about the image contents. The inferencing is accomplished with Bayesian networks that learn from image examples of each disease.

This effort at image understanding in fundus images anticipates the future use of medical images. As these capabilities mature, we expect that ophthalmologists and physicians in other fields that rely in images will use a system like STARE to reduce repetitive work, to provide assistance to physicians in difficult diagnoses or with unfamiliar diseases, and to manage images in large image databases.

## 1. INTRODUCTION

### 1.1. Electronic vs. photographic film images

Ophthalmologists rely heavily on images of the eye in patient care and research. The most common method of acquisition and storage of color and fluorescein angiogram images of the retina and optic nerve is film-based. Today, inexpensive computers are able to handle electronic images large enough to contain important details. The user can manipulate electronic images in ways that are superior to film-based images. For example, the user will be able to obtain automated diagnosis from images, compare sequential images, measure important structures in an image, and aggregate images similar in content. The physician can thereby receive decision support, be relieved of repetitive actions, and routinely obtain useful measurements currently too difficult or arduous to obtain.

### 1.2. Image understanding and scene analysis

Image understanding in medicine is a complicated job because many steps (image preprocessing, segmentation, classification, registration, recognition of objects from arbitrary viewpoints, inferencing) are involved. This process requires comprehensive knowledge in many disciplines, such as signal processing, pattern recognition, database management, artificial neural networks, expert systems, and medical practice. The STARE (structured analysis of the retina) system is designed to achieve automated diagnosis from images, find changes in objects

in sequential images, make measurements of key objects, and search large image databases based on image content [1,2]. This manuscript concentrates on the automated diagnosis concept of the STARE system.

### 1.2.1. Medical applications

Due to the complexity of image understanding in medical images, many computer applications in medical imaging have been concerned with smaller tasks, such as image enhancement, analysis tailored to detect a specific object, or completion of a particular goal. Nevertheless, the integration of the steps necessary for image understanding in medical images is beginning to be addressed. [1,3]

### 1.2.2. Ophthalmologic use

The retina is a forward extension of the brain and its blood vessels. Images of the retina tell us about retinal, ophthalmic, and even systemic diseases. The ophthalmologist uses images to aid in diagnoses, to make measurements, to look for change in lesions or severity of disease, and as a medical record. For example, while screening images of the ocular fundus, the physician may suspect the presence of diabetes from a pattern of hemorrhages, exudates, (yellow deposits in the retina), and cotton-wool spots (microscopic loss of circulation).

It is a natural human desire to find ways to avoid repetitive or routine work and be left with interesting and challenging work. Also it is advantageous to make use of outside expertise at the moment it is needed. There is a need for an imaging system to provide physician assistance at any time and to relieve the physician of drudgery or repetitive work.

### 1.4. Computer vision system

The STARE computer vision system seeks to reproduce the capabilities of the human expert, who can extract useful information and make decisions about diagnosis or treatment from medical images, even if the images are degraded. The STARE system extracts objects of interest (lesions and anatomical structures) from the rest of the image of the ocular fundus, identifies and localizes the objects, and infers about the presence and location of abnormalities to make diagnoses or look for change in sequential images. This was the original paradigm conceived for the STARE project. Successful methods of image understanding in other applications do not necessarily work with medical images. For example, model-based object recognition gives superior performance in aerial reconnaissance, but the process is not suitable for segmenting some objects in medical images, because of the wide variety of presentations of model objects, such as retinal exudates.

## 2. STEPS IN STARE

### 2.1. Image acquisition

The images of the ocular fundus can be acquired with a fundus camera or a scanning laser ophthalmoscope. Soon, electronic detectors will have adequate spatial, intensity, and color resolution to substitute for film. We currently transfer photographs to digital medium with film scanners. Color and monochromatic images are quantized to 12 bits per color plane and down-sampled to 8 bits without losing relevant information. Most processing in the STARE system is done on images  $820^2$  or  $410^2$ , though resolution up to  $2160^2$  is possible. The angle of view can be  $10^\circ$  for detail to  $60^\circ$  for panorama. The complex image analysis is made easier by a single viewport, shadowless coaxial illumination, and a nearly 2-dimensional scene with minimal occlusion.

### 2.2. Segmentation

#### 2.2.1. Preprocessing

In our current setup, we do not need to do preprocessing due to the good quality of the acquired images. In the future we can investigate preprocessing to rectify distortions due to media decay (e.g. astigmatic blur, defocusing, color shift, uneven magnification, scratches, dust).

#### 2.2.2. Segmentation algorithms in general

Objects of interest may be separated from the background by seeking the object boundary or by using characteristics of the objects, such as texture, color, size, and shape. If you know what you are looking for, you will always find it. Most successful, if it can be done, is to use template-matching algorithms.

#### 2.2.3. Segmentation of images of the fundus oculi

**Type of objects:** The objects of interest in the ocular fundus are lesions and abnormalities of anatomical structures. All the objects can be organized into three superclasses (1-3) of objects and two specific (4,5) objects: 1) curvilinear objects (including blood vessels), 2) blobs brighter than blood vessels, 3) blobs darker than blood vessels, 4) the optic nerve, and 5) the fovea (central vision spot).

**Rotating matched filter for blood vessel-like objects:** We reduced a modeled profile of a blood vessel to the mathematical representation of a Gaussian curve. Because blood vessels are darker than the background, we make the curve negative. In the horizontal orientation,  $K(x, y) = -\exp(-x^2/2\sigma^2)$  for  $|y| = L/2$  where  $K(x, y)$  is the transect profile of the blood vessel,  $L$  is the length of the segment, and  $\sigma$  is the average blood vessel width [4]. The green plane is convolved with this template in 12 orientations over  $180^\circ$  with the output at each point being the maximum for the 12 orientations (figure1). This technique works in images with distorted blood vessels or confounding lesions surrounding or lying under the blood vessels. Thresholding yields blood vessels and other curvilinear objects, such as the edges of large objects.

**Blob detectors for bright and dark lesions:** The intensity of blood vessels is the most stable measurement. We compensate for image exposure by normalizing to the intensity of blood vessels. Bright objects are found in the green plane remapped between 1.2 times the mean blood vessel intensity and 255. We

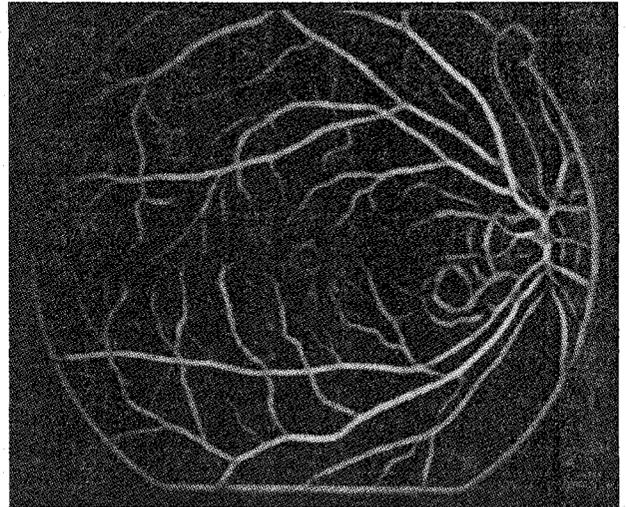


Figure 1: Output of blood vessel filter

convolve a flat circular bright object template at multiple scales for potential bright objects. Similarly, dark objects are extracted from images scaled between zero and 1.2 times blood vessel intensity. The borders of the gross blobs are refined by a histogram-thresholding technique to match the object border.

**Optic nerve:** We use three properties of the optic nerve in order to locate it: 1) the optic nerve normally appears as a bright disk approximately  $1500\mu\text{m}$  in diameter, 2) large vertical blood vessels enter the nerve from above and below, and 3) blood vessels in the retina tend to converge at the nerve. The nerve location algorithm forms three images, each indicating the strength of one of the three properties, and computes a weighted average of the three images. The position of the pixel of maximum intensity in the average image indicates the position of the optic nerve. Averaging ameliorates the impact of confounding factors that occasionally make one of the properties unreliable [5].

**Fovea:** The fovea can be identified in the blue plane image. The fovea is located 4.5mm temporal to the optic nerve and is marked by yellow pigment in the retina, which shows as a dark spot in the blue plane image.

### 2.3. Classification and location of objects of interest

#### 2.3.1. Classifier components

The input vector was comprised of a feature set of mathematical properties of objects and object measurements meaningful to ophthalmologists tailored to the superclass of objects. The original feature set was reduced by genetic algorithms [6]. We tested linear discriminant function, quadratic discriminant function, logit classifier, and back propagation artificial neural networks. For each superclass of objects, we chose a classifier that balanced accuracy and computation cost. The learning was supervised. The accuracy was tested with cross validation.

#### 2.3.2. Curvilinear objects

We divided the curvilinear objects into blood vessel segments and non-blood vessels. The input vector for each curvilinear object was the mean color of the center line of the

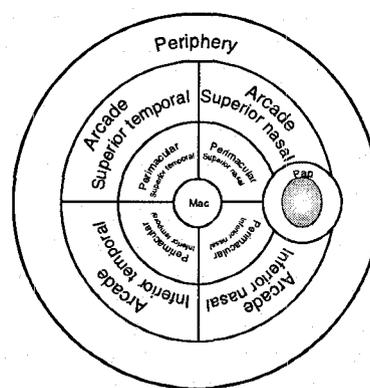
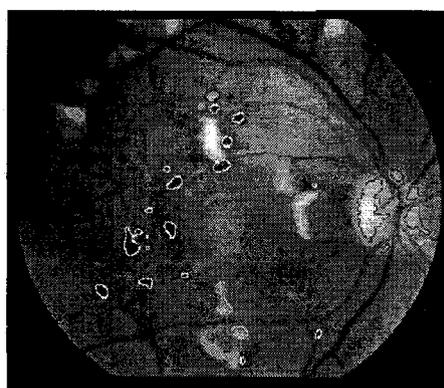
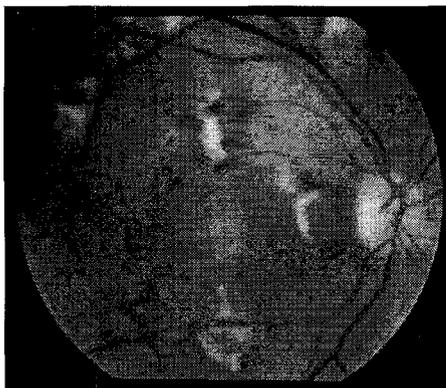


Figure 2: Cotton-wool spots, hemorrhage

Figure 3: Objects segmented and identified

Figure 4: Fundus coordinate system

object in the original image, the standard deviation of the color of the center line in the original image, the mean intensity of the center line for an object in the blood-vessel filter image, the length of the object, turns per object length, and the relative brightness of the original image on either side of the object. This last feature, the Duda Road Operator, had the highest utility [7]. A linear classifier has been sufficiently accurate, yielding accuracy of 84% compared to the human expert.

### 2.3.3. Bright objects

The list of bright objects includes exudates, cotton-wool spots, drusen, photocoagulation scars, subretinal fibrosis, and false objects. Useful features included object color, border color, texture measures, compactness, area, edge gradient, and turns per length of the border [8]. For color and brightness measures, the image was normalized for image exposure by average background color obtained after removing all objects. The logit classifier was best, with an accuracy of 89% (figure 2,3) [9].

### 2.3.4. Dark objects

The initial set of dark objects included hemorrhages, retinal blood vessels, pigment, and false objects. With the same features, the logit classifier provides accuracy of 78%.

### 2.3.5. Fundus coordinate system

Ophthalmologists tend to use clock hours centered on the fovea (center of macular region) and the distance from the fovea to describe positions on drawings or images of the retina. They also refer to regions and quadrants. From these concepts, we devised a fundus coordinate system, based on the polar coordinate system, that discretizes the retina into 11 regions meaningful to ophthalmologists and useful for inferencing (figure 4).

### 2.3.6 Annotated image

The objects are valued in severity in each of the 11 regions. For example, in the superior temporal perimacular region, exudates can be valued as absent, low percentage of region, high percentage of region. An image thus annotated can be used for inferencing about diagnosis.

## 2.4. Inferencing

### 2.4.1 Diagnosis as a large classification problem

Depending on how comprehensive the list of diagnoses, the number of ophthalmologic diagnoses can be several hundred.

Likewise the set of features and their values can be several hundred. Such large classification tasks become tractable if compromises and assumptions are made. The computational cost of obtaining the values of a full set of image features is large. Sequential ascertainment of data is profitable if the utility of each feature obtained is high [10,11]. Artificial neural networks do not require the features to be independent, but sequential data collection is not practical, and the reasoning steps are not available. An expert system can use sequential input of features based on utility and provides an audit trail. The data should be independent. Acquiring expert knowledge and improving system performance are arduous. Changing the classification or feature set requires repetition of this time-consuming and skill-intensive work. An expert system based on learning Bayesian networks reduces the need for such skilled labor, produces an audit trail of reasoning, and can incorporate beliefs or frequency analysis from the literature [11].

### 2.4.2 Learning Bayesian probabilistic expert system

A pilot inferencing system for fundus images in the STARE system has been developed. We present the format for the inferencing system as it is designed to be completed.

**Knowledge engineering:** The evolution of the expert system involves selecting a set of diagnoses and features, defining a causal probabilistic structure over the set of diagnoses and features, quantizing the features into discrete values, assembling a representative set of images, annotating the images, teaching the network, filling out the feature set with beliefs and frequency values, improving performance, and validating the results of the process.

The initial expert system is directed at a set of 43 vascular diseases of the retina. Among the types of diseases are diabetic retinopathy, branch and central retinal artery occlusion, branch and central retinal vein occlusion, arteriovenous anomalies, Coat's disease, sickle retinopathy, and hyperviscosity syndromes. The values of each feature (image manifestation) is correlated to each disease with conditional probabilities,  $p[M_i|D_j]$ . Such features include retinal exudates, cotton-wool spots, retinal hemorrhages, microaneurysms, venous caliber, arterial caliber, tortuosity, etc. An example of mutually exclusive and exhaustive values for a feature is absent, low severity, or high severity for cotton-wool spots in selected regions of the retina.

**Structure:** We model the decisions with influence diagrams, which are directed acyclic graphs containing nodes representing diagnoses or disease manifestations and arcs representing probabilistic dependence between the nodes (figure 5) [12].

**Knowledge acquisition:** We insert the relative incidence of each disease in four age groups, 0 to 6 months, 6 months to 2 years, 2 years to 60 years, greater than 60 years. We use annotated examples of each disease to propagate probabilities in the Bayesian network. The computational cost of obtaining each manifestation from an image is entered to help in deriving the utility of the next best feature. If necessary, probabilities are adjusted to optimize classification accuracy.

**Embedded expert system:** Hypotheticodeductive reasoning is applied. For any disease group, the values of specified features are always presented to the expert system. Based on the ranking of the diagnoses with the current set of features entered, the feature with the maximum utility from the remaining set is used to update the disease ranking. When the probability of a diagnosis reaches threshold probability, the diagnosis is accepted.

### 3. DISCUSSION

We consider this system as a pioneering effort for an imaging understanding system. We feel that in coming decades, capabilities being addressed in this system will be mature enough to become an integral part of the practice of medicine in disciplines that rely on images.

### 4. REFERENCES

[1] M.H Goldbaum, N.P Katz, S. Chaudhuri, M. Nelson, "Image Understanding for Automated Retinal Diagnosis," Proceedings: The Thirteenth Annual Symposium on Computer Applications in Medical Care, pp. 756-760, 1989.  
 [2] A. Gupta, S. Moezzi, A. Taylor, S. Chatterjee, R. Jain, S. Burgess, M. Goldbaum, "Content-Based Retrieval of Ophthalmological Images," IEEE International Conference on Image Processing, vol. ICIP-96, 1996.  
 [3] W.E.L. Grimson, "Medical Applications of Image Understanding," IEEE Expert, vol.10, pp. 18-28, 1995.  
 [4] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, "Detection of Blood Vessels in Retinal Images Using Two-Dimensional Matched Filters," IEEE Transactions on Medical Imaging, vol. 8, pp. 263-269, 1989.

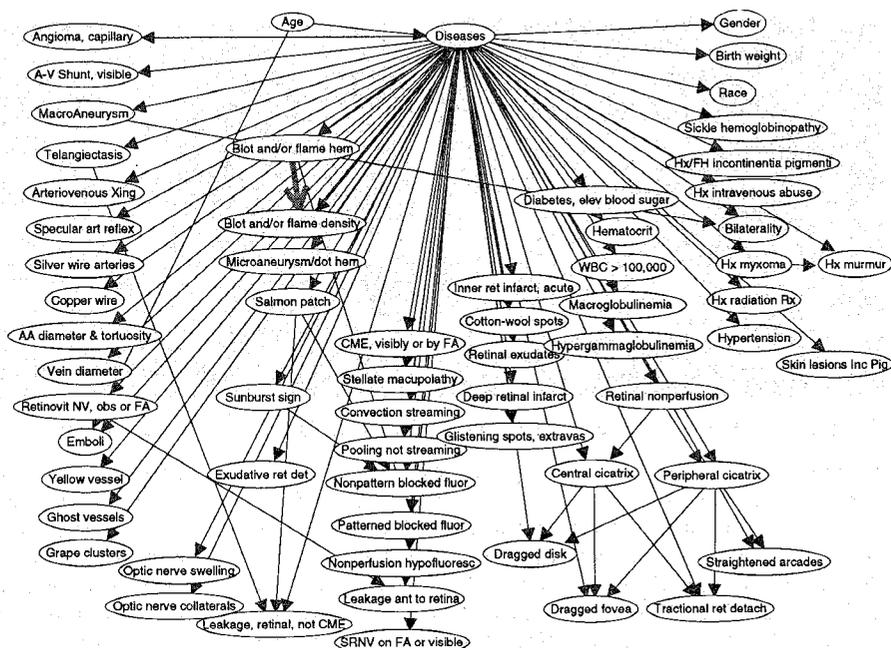


Figure 5: Influence diagram, manifestations in vascular diseases

[5] S. Chaudhuri, S. Chatterjee, N. Katz, M. Goldbaum, "Automatic Detection of the Optic Nerve in Retinal Images" Proceedings IEEE International Conference on Image Processing, Singapore, vol. 1, pp. 1-5, 1989.  
 [6] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison Wesley, 1989.  
 [7] M. Fischler, J. Tenenbaum, H. Wolf, "Detection of Roads and Linear Structures in Low Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique," Computer Graphics and Image Processing, Vol 15, 201-223, 1981.  
 [8] M.H. Goldbaum, B.L. Côté, R.F. Garcia, W.E. Hart, P. Kube, M. Nelson, "Computer Detection and Classification of Lesions Found in Diabetic Retinopathy," Invest Ophthalmol Vis Sci, vol. 33, p. 1082, 1992.  
 [9] J.S. Cramer, "The Logit Model: An Introduction for Economists," Edward Arnold, 1991.  
 [10] E. Horvitz, J. Breese, M. Henrion, "Decision Theory in Expert Systems and Artificial Intelligence," J Approximate Reasoning, Vol. 2, pp. 247-302, 1988.  
 [11] D. Heckerman, "A Tutorial on Learning with Bayesian Networks," Technical Report MSR-TR-95-06, Microsoft Research Advanced Technology Division, Microsoft Corporation, 1996.  
 [12] R. Howard, J. Matheson, "Influence Diagrams," in R. Howard, J. Matheson, editors, "Readings on the Principles and Applications of Decision Analysis, vol. II, pp. 721-726, Strategic Decisions Group, Menlo Park, CA, 1981.