IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITP4809

An Exploratory Study Exploiting Image Features for Term Assignment in Medicine

Abby A. Goodrum
Syracuse University
School of Information Studies
4-110 Center for Science & Technology
Syracuse, New York 13244-4100
aagoodru@syr.edu

ABSTRACT

In this paper, we introduce a new approach to medical image indexing. This approach utilizes image features (content) to drive the assignment of terms from a controlled vocabulary in medicine (MeSH). This paper reports preliminary results from an exploratory study to determine the effectiveness of exploiting image content features for index term assignment within the context of a library of medical images.

1. RATIONALE

1.1 Overview

The number of digital images is rapidly increasing in health care. Increasingly they are becoming a standard feature of the electronic health care record. Still images and video are also integral to education in the basic sciences of medicine as well as to clinical teaching and research, especially in domains such as radiology, pathology, ophthalmology, and dermatology.

For nearly a decade it has been possible to use image content features such as color, texture, and shape for classification and retrieval. However, human beings often prefer to locate images using words. Within the medical community, a highly structured approach to accessing text documents has been through the use of medical subject headings (MeSH). MeSH is the National Library of Medicine's thesaurus of controlled medical terms including "broader-than," narrower-than," and "related" linkages. These links represent the relationships between terms and provide a hierarchical structure that permits searching at various levels of specificity. MeSH consists of a hierarchical structure in an extensive tree structure, currently at nine levels, representing increasing levels of specificity. The manual assignment of controlled terms from this hierarchy can be costly and time consuming however, and the image indexing backlog in some medical libraries has increased dramatically. Additionally, a large number of new medical images emerges on the Web daily without benefit of controlled indexing. This leads us to question whether mechanisms for automatic image processing might be leveraged to aid in the assignment of controlled vocabulary terms.

1.2 Background

Text-based image classification has a long history and has many strengths including the ability to represent both general and specific instantiations of an object at varying levels of complexity. Unfortunately, manual assignment of textual attributes is both time consuming and costly. Furthermore, manual indexing suffers from low term agreement across indexers and between indexers and user queries. [2]

More recently, automatic assignment of textual attributes to images has been conducted utilizing the text from captions, transcripts, close captioning, or verbal description for the blind that accompany some videos [9]. While these approaches greatly reduce the labor involved in manual assignment of keywords, it must be remembered that many images are without accompanying text. Furthermore, users' image needs may occur at a primitive level that taps directly into the visual attributes of an image. These attributes may best be represented by image exemplars and retrieved by systems performing pattern matches based on color, texture, shape, and other visual features.

Problems with text-based access to images has prompted increasing interest in the development of image-based solutions. This is most often referred to as content based image retrieval (CBIR) [1]. Content based image retrieval relies on the characterization of primitive features such as color, shape, and texture that can be automatically extracted from the images themselves. By now, most are familiar with the basic features of the several widely used content based image retrieval products. The underlying techniques are far from new. However, these methods appear now to be mere technology building blocks – interesting in potential, but of value mostly to the extent that they support or assist some higher-level form of image description and retrieval [8].

There exists, for example, a considerable gap between the primitive image features such as color, texture, lines, edges, and angles, and the higher level cognition necessary to equate these features with terms that occur to human beings in the course of a search. There are several areas (e.g., landsat imagery, engineering diagrams, trademarks, etc.) where CBIR performs well, but these apply to special, limited communities, not to the usually conceived broad range of potential inquirers [2].

Even when CBIR system searches are constrained by exemplar images they frequently produce paradoxical and unexpected results. A search with an exemplar of a signpost by a highway will retrieve images of buildings. Photographs of wiring assemblies will be retrieved along with pictures of crews in space shuttle mission archives. A degree of misclassification appears to be inherent in all current approaches to CBIR. Combined use of text and images to correct misclassification has been examined [6] [11] , however this approach is dependent on the availability of some text accompanying images.

1.3 Content Based Properties and Image Clustering

Common image content properties are generally encoded as a series of histograms. Each histogram reflects a scale of an image property. Commonly used properties are: greyscale, color (red, green, blue or luminance, chroma and hue), line lengths, edge intensities, angle declinations, and texture. The combination of these histograms can be used for image clustering, image searching, and image viewing

Images that have been decomposed into their respective contents may then be clustered by any of several scaling methods [4], [5]. One of the simplest methods of clustering images by similarity is to reduce the images to their histogram values and then subject them to multidimensional scaling (MDS). In MDS similarities among stimuli can be considered as distances along physical dimensions such that if stimulus A is more similar to stimulus B than it is to stimulus C, stimulus point A will be located closer in space to stimulus point B than to stimulus point C. Thus clusters which form near one another are presumed to be related. In MDS, stimuli are represented by points in an underlying multidimensional space. The points do not have a single value, for example "Darkness to Lightness," but are described by as many scale values as there are dimensions to support the fit between points. The subsequent mapping of these dimensions creates a visual display of the underlying structure of the entire stimulus set.

As any indexer would be quick to note, however, image clustering is not the same as image classification. Classification implies the imposition of intellectual categories upon material. Moreover, these categories may cross a number of dimensions that do not depend upon simple naming or identification of objects in the images themselves. In the next section, therefore, we suggest several methods by which *image clustering by content* may contribute usefully to *image classification*.

1.4 Exploiting Content Based Image Features for Image Classification

Image clustering by content can be usefully exploited. If isomorphic relationships can be expected to occur among terms and images, then it is reasonable to assume that new terms can be linked to images in the database by their proximity and that new images entering the collection should suggest existing terms by their proximity.

However, the boundaries for the effectiveness of the use of this property with medical images remains largely untested. The first published display of medical image clustering by content using multidimensional scaling appeared in 1993, [7] [10] for a series of images of white blood cells. In this work, images of blood leukocytes were clustered automatically based upon image feature similarity.

The primary application of this property is the classification of new images entering a collection. Given a set of images already described and clustered, it seems reasonable to suppose that new images clustered by content and appearing proximate to the existing images should, to some degree, inherit the terms of the older images. Use of this inheritance property would permit some term assignments to take place quickly, and would support immediate retrieval on a shallow level.

2. AIMS AND OBJECTIVES

The goal of this exploratory study is to determine the degree to which content based feature extraction may support index term assignment within the context of a library of medical images.

The specific objectives are:

- To discover the degree to which the similarity of image content measures implies the inheritability of terms. If the machine finds that two or more images are similar, does this mean that the index terms assigned to one should be assigned to another.
- To determine if there is a threshold of similarity beyond which image content similarity has little or no meaning for term assignment.
- To propose specific uses of the technique. For example, the process may be valuable for initial term assignment to new images entering the collection to provide some retrieval access by words.

3. METHOD

The specific method employed in this study is discussed in more detail in [3]

The medical images used in this study were obtained from a health sciences library multimedia cataloging project. The collection consists of images, illustrations, animations, and videos related to the health sciences. Each image in the collection has been fully cataloged by professional medical librarians. Between 3 and 5 Medical Subject Headings (MeSH) are used to describe each image, as well as local subject headings where appropriate.

We selected a set of 15 clinical images. Of these, 13 were images of the brain and 2 were images of the musculoskeleture.

The method used in this study to explore the assignment of subject headings based on image content features consists of five steps:

Step 1: Manually assign MeSH terms to images. This is done by the medical librarians at the time of digitization.
Step 2: Submit all images in the test collection to feature extraction and MDS clustering.
Step 3: Examine within-cluster term frequency and degree of term overlap for term-bearing images.
Step 4: Examine between-cluster term frequency and degree of term overlap for term-bearing images.
Step 5: Compare terms inherited from cluster to terms assigned by librarians.

4. RESULTS AND DISCUSSION

Figures 1 and 2 below display the sample of 15 medical images after clustering. In Figure 1, images can be seen to cluster together within four groups. Additionally, there appears to be an isolated image outlier (Image 8). In Figure 2 individual members of each cluster are shown to have visual similarity. Images can be seen to clearly cluster based on shape and color. Axial views of the brain whether recorded with magnetic resonance imaging (primarily black and white) or emission computed tomography (brightly colored images), are seen to cluster more closely than sagittal views of the brain recorded with magnetic resonance imaging (MRI). MRIs of the area around the knee cluster together apart from the brain images. An isolated axial view of the brain taken with emission computed tomography layers over MRI clusters midway between the black and white and the color axial views of the brain.

For this study we accepted the indexing supplied by the medical librarians as correct. Of primary interest were the occurrence of terms within each cluster and the occurrence of terms overall. We consider frequently occurring terms within a cluster to be better candidates for inheritance than terms used infrequently within a cluster. We counted as unique those terms shared by less than 50% of the members of a cluster. Table 1 presents the occurrence of terms by cluster.

The subject headings: "Central Nervous System" and "Brain" are used in all images except for those in cluster D (the knee). New images clustering with A, B, C or the outlier image #8 could safely inherit these frequently occurring terms with little effort.

The inheritance of terms that describe the recording mechanism is somewhat more complex. While 10 of the 15 images are indexed : "Magnetic resonance imaging – diagnostic use", new images entering the collection and clustering with group B would have to inherit the designation: "Tomography, Emission-computed." The real problem lies in the variability in recording and presentation. Although the selection of images used in this study exhibit clearly defined clusters based on shape and color, what can we make of outlier #8. This image exhibits an overall tendency toward black and white (MRI) but also presents a small amount of color in one area. This might for instance indicate an image created by the overlay of MRI and PET or SPECT imaging.

Figure 1. Display of 15 medical images after clustering by MDS

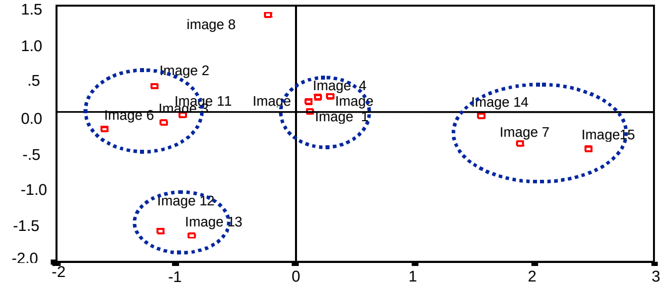
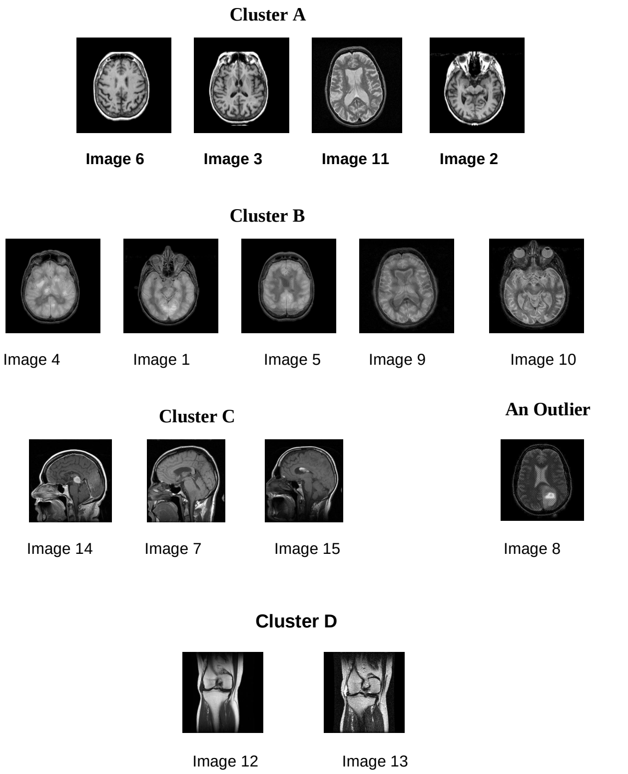


*Figure **2.** Individual images within each cluster*



Another related issue is that of the parent-child relationship within the hierarchy of MeSH. In this example, the term: "Central Nervous System" occurs one level up the terminological hierarchy than the term: "Brain." It would seem reasonable therefore that new images coming into the system should be capable of inheriting frequently occurring terms as well as terms one level up in the hierarchy, but this is not the case. The hierarchical relationships generally thought of as parent-child relationships are better understood as representing broader and narrower concepts in MeSH. The term: "Breast Neoplasms" falls under "Breast Diseases" as well as "Breast Diseases by Site."

These results are limited by several factors. The depth of the MeSH indexing may be too shallow to drive term inheritance with any meaning. For example, a sudden acquisition of thousands of axial MRIs of the brain all inheriting the terms: "Magnetic Resonance Imaging," "Brain" and "Central Nervous System". Another limiting factor in this study is the small number and homogenous content of the selected images. Finally, term inheritance in this study was predicated on term frequency, but certainly other measures might have been explored. For example, it might be useful to consider the relative distances between images such that a new images would only inherit terms of its nearest neighbor.

5. FUTURE WORK

We are currently engaged in exploring these issues with a much larger and more heterogeneous collection of medical images. Our primary goal is to implement various techniques on a trial basis (contingent on discovery of their usefulness) in a medical library environment in order to understand conditions of their acceptability to working staff.

Table 1. Term Occurrence By Cluster

CLUSTER	IMAGES IN CLUSTER	TERMS IN COMMON	UNIQUE TERMS
A	6, 3, 11, 2	Magnetic resonance imaging – diagnostic use Central Nervous System Brain	Multiple Sclerosis Seizures
B	4, 1, 5, 9, 10	Central Nervous System Brain Tomography, Emission-Computed	Central Blood Flow Parkinsons Disease
C	14, 7, 15	Magnetic resonance imaging – diagnostic use Central Nervous System Brain	Encephalitis
D	12, 13	Knee Magnetic resonance imaging – diagnostic use medial meniscus	NONE
OUTLIER	8	Central Nervous System Brain Magnetic resonance imaging – diagnostic use	Astrocytoma

6. REFERENCES

1. Eakins, J.P., Graham, M.E.: Content-based image retrieval: A report to the JISC Technology Applications Programme, Institute for Image Data Research, University of Northumbria at Newcastle, January, 1999.

2. Enser, P. & McGregor, C.: Analysis of visual information retrieval queries. British Library Research and Development Report, 6104. (1993).

3. Goodrum, A., Rorvig, M., Jeong, K. & Suresh, C: An Open Source Agenda for Research Linking text and Image Content Features. Journal of the American Society for Information Science & Technology, 52(11), . (2001) 948-953.

4. Han, K., Myaeng, S: Image Organization and Retrieval with Automatically Constructed Feature Vectors, 19th International ACM SIGIR Conference on Research and Development in Information Retrieval (Frei, H., Harman, D., Schauble, P., Wilkinson, R., Eds.), Zurich, Switzerland, August 18-22, (1996). 157-165, ACM: New York.

5. Leouski. A., Allan, J.: Visual Interactions with a Multidimensionally Ranked List, in Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, August 24-28, (1998), 21:353-355.

6. Patrick, B., Seivert, M., & Popescu, M.: Text indexing of images based on graphical image content. 62nd Annual Meeting of the American Society for Information Science, (Woods, L. Ed.), Washington, DC, October 31-November 4, (1999), pp. 675-680.

7. Rorvig, M., Fitzpatrick, S., Ladoulis, C.T., Vitthal, S.: A New Machine Classification Method Applied to Human Peripheral Blood Leukocytes, Information Processing and Management, 29(6): (1993) 765-774.

8. Srihari, R.K., Zhang, Z., Manmatha, R., Ravela, S.: Multimedia Indexing and Retrieval," Proceedings of the 22st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Berkeley, California, August 25-29, (1999) ACM: New York.

9. Turner, J.: Some characteristics of audio description and the corresponding moving image. 61st Annual Meeting of the American Society for Information Science, (Preston, C. Ed.), Pittsburgh, PA, October 25-29, (1998) 108-117.

10. U.S. Patent No. 5,181,259: General Method of Pattern Classification Using the Two Domain Theory (1993).

11. Westerveld, T. Image Retrieval: Content versus Context. Proceedings of RIAO 2000.