

CervigramFinder: A Tool for Uterine Cervix Cancer Research

Zhiyun Xue¹ PhD, Sameer Antani¹ PhD, L. Rodney Long¹ MA, Mark Schiffman² MD,
George R. Thoma¹ PhD

¹ National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

² National Cancer Institute, 6120 Executive Blvd., Bethesda, MD 20892

Abstract

The need for medical image retrieval by visual content has steadily increased with greater use of images in clinical medicine and research. This paper presents CervigramFinder, a new Web-accessible Content-Based Image Retrieval (CBIR) prototype system created by the National Library of Medicine to aid uterine cervical cancer research and education for cervigram database access. It represents a significant effort toward developing what may become essential capabilities of biomedical image database systems.

Introduction

Cervigrams are color optical images of the uterine cervix taken during a gynecological exam and may provide clinically important information for cervical cancer diagnosis. The National Library of Medicine (NLM) has created a database of 100,000 cervigrams as well as other correlated (text) clinical data collected during two major studies in cervical cancer conducted by the National Cancer Institute (NCI) [1]. NLM is developing a suite of tools [2] that include text data search capability to access this information. In an effort to overcome limitations of text-only searching, research efforts exploring information retrieval by directly querying the visual content have resulted in the CervigramFinder system. It is the first Content-Based Image Retrieval (CBIR) system for accessing cervigram image databases.

System Design

CervigramFinder is implemented using a distributed client/server framework which is expected not only to be a significant benefit in evaluating the system, but also for its eventual use at multiple research sites. It employs open standards and open-source software and can be accessed via the Web. We followed principles of modular software development that enable straightforward enhancements to the user interface and easy integration of new algorithms, resulting from our ongoing research on the topic, with minimal impact on other components.

Interpreting cervigrams requires gynecologic oncologists to assess visual attributes of local image regions. CervigramFinder provides the capability for visual search of clinically important regions using local image characteristics, making it different from

many other CBIR systems which are limited to using visual attributes computed over the entire image. The system has been developed with a user-in-the-loop approach to minimize the “semantic gap”, with user-annotated query regions and automatic image processing algorithms computing their visual attributes. CervigramFinder uses color, texture, size and location attributes to describe the image regions and retrieve images with similar regions from an expert-annotated subset of the cervigram database.

System Testing

CervigramFinder incorporates fundamental CBIR functions, such as feature extraction, normalization, dimension reduction, combination, and various similarity measures. Methods selected for the system resulted from an extensive performance evaluation of key techniques for cervigram retrieval [3], e.g., image color and texture representation algorithms. In addition, we collaborated with medical experts at different stages of development and also conducted a usability study to minimize the “usability gap” which is an understudied, important aspect in biomedical image informatics. Study outcome resulted in an improved UI and enhanced system capabilities.

Conclusions

CervigramFinder represents efforts toward filling an important need in biomedical image informatics research. The tool is among a suite of tools that provide capability for accessing and annotating data resulting from NCI’s uterine cervix cancer research, and provides access to cervigram images through visual queries on localized image regions. We expect the results from this Web-accessible CBIR system to provide valuable support in cervical cancer research and training.

References

1. Herrero R, Schiffman MH, et al. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa-Rica: the Guanacaste Project. *Rev Panam Salud Publica* 1997; 1(5):362-75.
2. Jeronimo J, Massad S, Schiffman M, et al. Visual appearance of the uterine cervix: correlation with human papillomavirus detection and type. *Am J Obstet Gynecol.* 2007; 197(1):47.e1-47.e8.
3. Xue Z, Antani SK, Long LR, et al. Investigating CBIR Techniques for Cervicographic Images. *Proc AMIA Annu Fall Symp.* 2007; 826-30