QCHARM: A Novel Computational and Scientific Visualization Framework for Facilitating Discovery and Improving Diagnostic Reliability in Medicine

Brian A. Canada[1], Keith C. Cheng[1,2], MD, PhD, and James Z. Wang[1,3], PhD
[1]Integrative Biosciences Program, [2]College of Medicine (Hershey, PA), and [3]College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA

Introduction. Histopathologists microscopically examine tissue samples, sectioned into thin slices, in order to make accurate diagnoses of tumors and other diseases; however, diagnoses based on examinations of a single tissue sample may vary considerably from one histopathologist to the next (or even for individual pathologists, for multiple examinations of the same sample)[1]. As a result, there are few guarantees about the reliability of such diagnoses, which may lead to critical errors regarding the timing and choice of cancer therapies. In response to this problem, we are developing advanced image recognition methods whose goal is to automatically recognize and quantitatively characterize abnormalities in the tissue morphology of larval and adult zebrafish. The zebrafish has been shown to be an excellent model organism for vertebrate development and human disease, largely because its transparent embryo allows the effects of mutations to be easily identified[2]. Once we have tested the algorithms to characterize zebrafish tissue morphology, we will extend them to clinical applications such as the histopathology of cancer.

These algorithms represent the central element of a novel scientific visualization framework that we are developing. This framework will not only facilitate the ability to explore zebrafish tissue morphology, but will also be compatible with biological databases representing the genomics of other model organisms (rat, mouse, primates, etc.) as well as humans. The annotated genes underlying the abnormalities in zebrafish can thus be compared against analogous genes in other species. The existence of similar traits in other species will provide evidence of the function of the analogous gene in humans. This can lead to an improved understanding of human development and also drive more targeted development of disease treatments.

Development of methods. We are developing the novel computational framework QCHARM (Quantitative Characterization of Histological Abnormalities for Research and Medicine), an extension and refinement of the Automatic Linguistic Indexing of Pictures (ALIP) method developed by Li and Wang[3]. ALIP is a content-based image retrieval (CBIR) system used to train computers to automatically recognize and annotate previously un-annotated images by extracting features from an image and comparing them to categories of images that have previously been annotated. The demonstrated high accuracy of the ALIP system's potential to index photographic images make it an ideal foundation on which to build a framework for quantitatively characterizing abnormalities in tissue morphology. QCHARM will be designed to efficiently characterize fine details of high-resolution, high-throughput biological images in real time, and we have obtained a comprehensive set of zebrafish images (spanning the degrees and causes of tissue abnormalities at various stages of development) that we will use for training and testing of the algorithms. QCHARM will also be capable of characterizing features of interactive 3D zebrafish models that have been reconstructed from images of serially sectioned slices.

In order to allow both clinicians and researchers to effectively make use of QCHARM, we are developing a user interface in the form of a series of interactive panels, including a "virtual microscope" panel for zooming in on details of a given 2D slice, a set of panels for interacting with 3D models, and an information panel for recording abnormality scores and other annotations, such as for recording the genes associated with those abnormalities. To permit compatibility with other biological databases, the interface will be integrated into the Semantic Web—an ontology-based Web architecture being embraced by the scientific community as a universal platform for the interoperability of bioinformatics applications[4]. Since the Semantic Web is based on XML (eXtensible Markup Language), our framework is being implemented using XML-based markup languages. For example, the 3D zebrafish models are being coded in X3D, the International Standards Organization (ISO) standard for communicating interactive virtual models across Web applications[5].

We will test for and refine the accuracy and usability of QCHARM through frequent feedback from collaborating histopathologists, toxicologists, members of the zebrafish community, and other interested researchers.

Acknowledgements. K.C.C. and J.Z.W. are contributing equally to this project, which is based on work supported in part by NSF, NIH, and Penn State University.

REFERENCES

1. Colquhoun P, Nogueras JJ, Dipasquale B, Petras R, Wexner SD, Woodhouse S. Interobserver and intraobserver bias exists in the interpretation of anal dysplasia. *Dis Colon Rectum*. 2003;46:1332-8.
2. Bradbury J. Small fish, big science. *PLoS Biol*. 2004;2:E148.
3. Li J, Wang JZ. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2003; 25:1075-88.
4. Salamone S. Masters of the Semantic Web. *BioIT World*. 2005;4:28-32.
5. Brutzman D, Harney J, Blais C. X3D Fundamentals. In: Geroimenko V, Chen C, editors. *Visualizing Information Using SVG and X3D: XML-based Technologies for the XML-based Web. London, England: Springer;* 2005:63-84.