Digital Content for Radiotherapy Training

Shih-Tsang Tang[1], Chia-Hung Hsiao[2*]

[1]*Department of Biomedical Engineering, Ming Chuan University, Taiwan, ROC*
[2]*Department of Medical Informatics, Tzu Chi University, Taiwan, ROC*

Usually, oncological clinicians take years to deal with a great amount of patients, and then turn into a domain expert. As a result, clinical training is time-consuming and expensive. And then it's valuable to develop an eLearning structure for oncological training. This research is to develop an oncological knowledge repository, and further investigated the processes and methods to classify medical images and relative reports generated in PACS. In the repository, both the image pixel data and object description data were created, stored, and with the function of content-based retrieval. The all arrangements are for representing the knowledge concerning the features of tumors and treatment plans.

Keywords: eLearning, Radiation oncology, Therapy planning.

INTRODUCTION

Usually, oncological clinicians would take years to deal with a great amount of patients before turning into a domain expert. As a result, clinical training is time-consuming and expensive, and then it's valuable to develop an eLearning structure for oncological training.

Recently, IMS (Instructional Management Systems) Global Learning Consortium has defined the specifications for digital repositories interoperability in eLearning. Some core functions (submit, search, expose, etc.) for interoperability between systems (content repository, LMS; Learning Manage System, LCMS, Learning Content Manage System) have been defined. The requirements and functionalities of IMS definitions are similar to those protocols for DICOM (Digital Imaging and Communications in Medicine) standards. Both the content repository in eLearning and image repository in PACS (Picture Archiving and Communication Systems) have similar core functionalities (storage, search, and retrieve). DICOM standard defines the levels (patient, study, series, and image) for searching image objects. Each level has defined some attributes (such as patient's Name, ID, gender, study date, referring physician). Each attribute may support some types of matching (single value matching, wild card matching, range matching, etc.).

However, the DICOM information models and specifications are defined for using in clinical healthcare process. It is not for learning or knowledge management. Some query attributes (such as study description, procedure code, diagnoses description, etc.) that might be important for research, but are optional attributes defined in DICOM standard. DICOM server may not support those attributes for query. As a clinical content repository, medical images in PACS or in the repository should be further categorized by diseases, symptoms, image locations, etc.

This research is to develop an oncological knowledge repository, which investigated the processes and methods to classify medical image and relative reports generated in PACS. Both the image pixel data and object description data were created, stored, and with the function of content-based retrieval. The all arrangements are for representing the knowledge concerning the features of tumors and treatment plans.

METHODS

In a healthcare environment, many different types of multimedia data would be generated. The multimedia data could be stored in a single system (centralized system) or separately in many distributed systems. However, the different types of multimedia data (2D images, video, waveforms, etc.) should be processed and presented separately.

The target data for accommodation are generated in the standard workflow of radiotherapy. Although the developed repository likes a PACS server for storing the medical images, but for a knowledge-based repository the query keys and indexes would be quite different. PACS server, as defined in DICOM, supports patient- and workflow-oriented indexes (such as patient ID, access number, study date, modality types, etc.) for searching the stored images. The knowledge-based repository supports disease-oriented indexes (such as syndrome, disease types, disease stages, location and area of lesions, etc.) for sharing the stored contents. The repository scheme would be developed along with DICOM IOD (Information Object Definition) to define the tags of the collected data.

Content-based retrieval is helpful to clinician training in tumor diagnosis and treatment, which provides similar reference cases in preparing a new case. Content-based retrieval was implemented by the comparison of reference model, and the similarity is a conceptual distance measurement. The criteria for retrieval are defined by both shape and texture features of the images. A reference model can be useful in selection from the repository by means of supervised classification that uses a maxi-

mum likelihood approach as the distance measurement. Therefore the reference models would be stored in the repository with all their features.

RESULTS

A WSP (Web Service Provider) structure was developed for each type of multimedia data (Figure 1). The structure was carefully defined to support the general clinical content accessing protocol, which supports DICOM standard and collects data generated in the full therapeutic workflow, including diagnosis, planning, treatment, etc., and then transfers the proprietary data to DICOM format. Consequently, browsers can use the protocol to search, view, and edit all the clinical contents.

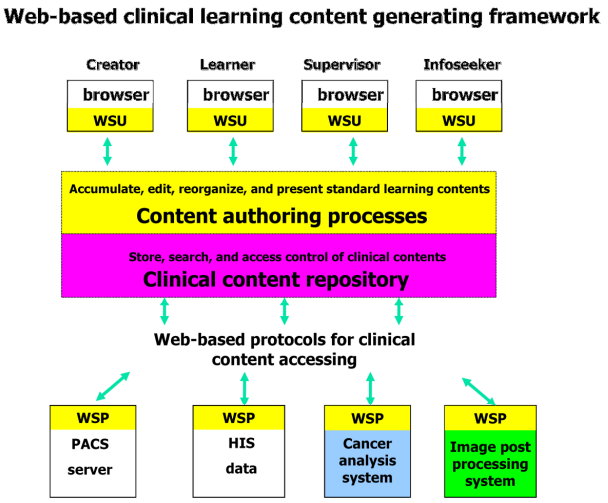

Figure 1. Dataflow of repository.

For our research goal, the disease case is far more important than the patient or workflow. In radiation oncology, the location, size, area, and shape of tumor would be the most important indexes. As a result, the "Study IOD" becomes the core index in our research. An information model was developed as show in Figure 2.

The DICOM data and protocols are not suitable in WEB-based knowledge management or eLearning application. Therefore the collected data were further transferred as XML formatted structure. Consequently, the contents can be easily accessed and handled via Internet everywhere and anytime.

The reference models were also developed for content-based retrieval. In order to store the reference model in XML format, a specialized DTD (Document Type Definition) would be developed. [1–6] The developed DTD structure should be independently usable for any similarity criteria and features. Hence, a meta-grammar was developed to comprise the feature descriptions as well as the association with reference models regarding image content. The structure of the developed DTD is shown in Figure 3. It is compulsory to start with the <similarity> element.

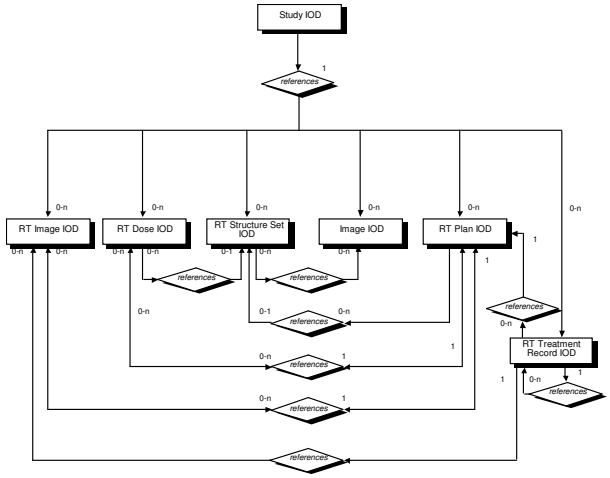

Figure 2. Information model for repository scheme.

Within the root element, one <featurespace> element describes the entire set of defined images and one <trainingset> element describes the associations between the medical images. To define a medical data set and its features, the <dataset> element is used. Thus, the DTD defines that the <featurespace> contains <dataset> elements. The <trainingset> describes associations between reference data sets and clusters. The parser reads file and creates a document object model. [7] In this model it is easy to add new elements of reference data sets and training set associations. The <dataset> element describes the path of the file in the physical disk management system and the parameters of the data set. Within the <dataset> element, features are stored with <feature> elements, which describe the shape and texture features of a data set. In that case various features are stored that represent the characteristics of a data set. Each <feature> element has a numerical value field to define the feature's value. This element contains all data sets that belong to one class in the feature space. The <cluster> element is used to fit the supervised classification process. The <cluster> element contains parameters for the file name and the path. The <example> element describes data sets whose features are used to train the classification. The values defined in <example> element are the similarity to the <cluster> element. Furthermore, the methods of storage and access are important for interactive systems to keep response time reasonable. Dimension reduction techniques or pruning methods was introduced to allow efficient and quick accesses. [8]

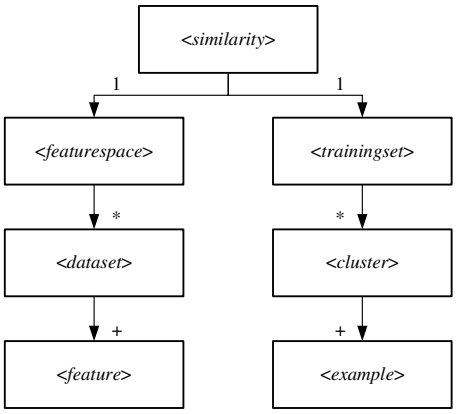

Figure 3. Document type definition.

DISCUSSION AND CONCLUSIONS

This research is to develop an oncological knowledge repository for representing the knowledge of treatment, which is basing on the current workflows of oncology diagnosis and treatment, and to accumulate, present, and share clinical content and domain knowledge to provide resident education and training in tumor treatment planning.

It is interesting and important to investigate the process and method to further classify medical image and relative reports generated in PACS, which can provide even more attributes on the search interface and explore the potential of the image content repository. The stored data could additionally capsulate into a content package, followed SCORM/CAM (Sharable Content Object Reference Model/Content Aggregation Model) standard. And then the repository is available for future study in adopting standard learning and content management system in developing the interactive learning platform.

ACKNOWLEDGMENTS

This study was supported by grants from National Science Council (NSC-97-2221-E-320-003-), and none of the authors have any conflict of interest.

REFERENCES

1. Kunz, H., Derz, C., Tolxdorff, T., and Bernarding, J. 2004. XML knowledge database of MRI-derived eye models. Comput Methods Programs Biomed. 73:203–208.
2. Dugas, M., Kuhn, K., Kaiser, N., and Uberla, K. 2001. XML-based visualization of design and completeness in medical databases. Med Inform Internet. 26:237–250.
3. Hoelzer, S., Schweiger, R. K., Boettcher, H. A., Tafazzoli, A. G., and Dudeck, J. 2001. Value of XML in the implementation of clinical practice guidelines—the issue of content retrieval and presentation. Med Inform Internet. 26:131–146.
4. 2002. World Wide Web Consortium Extensible Markup Language (XML). Available at http://www.w3c.org/xml.
5. Aviram, M. H. 1998. XML for dummies. IDG Books Worldwide, Foster City.
6. Kahn, C. E. 1999. Standard generalized markup language for self defining structured reports. Int J Med Inform. 53:203–211.
7. 2002. Document Object Model. Available at http://www.w3.org/ DOM/.
8. Muller, H., Squire, D. M., Muller, W., and Pun, T. 1999. Efficient access methods for content-based image retrieval with inverted files, p. 461–472. Proceedings of the 4th SPIE Conference on Multimedia Storage and Archiving Systems, SPIE, Boston, USA.