SurgLink: Cross-media Linking in the Operating Room

Yihan Deng

Innovation Center Computer Assisted Surgery (ICCAS)
University Leipzig
Semmelweissstr. 14
04103 Leipzig
yihan.deng@medizin.uni-leipzig.de

Abstract: The clinical daily work is accompanied by different types of patient records. A successful diagnostic and therapeutic process relies on efficient access and comprehension of a patient's status. With the proceeding of treatment, the physician can easily be overwhelmed by the rapid growth in data volume of patient records. In particular, the information seeking and navigation between different media consume tremendous time and labor. Due to the lack of a cross-media connection, the retrieval of similar cases and a causality analysis among the heterogeneous data sources can merely be conducted by hand. This paper presents the preliminary stage of the PhD study and tries to raise and explain the main research questions for finding an approach to cross-media linking in the operation room. First of all, the status quo of the multimedia data storage in clinical context is introduced. Second, the special requirement of cross-media linking in the operating room is presented. Finally, the connection possibilities between different media and the feasibility of the cross-media linking are discussed.

1 Introduction

A clinical information system (CIS) serves normally as an organizational platform for the clinical information process, which is defined as socio-technical system [Rop99]. It comprises technical systems and human activities. The patient records in the CIS platform are thus distributed over several systems, which makes it difficult to obtain an integrated view. Since the first reception of a patient in the hospital, all necessary information for the diagnosis and treatment covering laboratory test, radiological and pathological investigation are stored in different sub-systems. As can be seen in Figure 1, physicians have to face numerous heterogeneous data sources. Generally, the most commonly used types of item are: 1) Text: Finding report, nurse letter, admission note, drug prescription, discharge summary, etc. 2) Image: Radiological and pathological image, endoscope image, scanned document (legacy files, important paper-based note, referral documents). 3) Audio: Dictated report for prompt recording of diagnosis and findings. 4) Video: Operation records for the purpose of teaching and quality management.
Although methods for text retrieval have been developed and optimized for some time, little work has been done to realize a clinical cross-media linking and retrieval. Along with the rapid growth in the data volume of patient records, the information seeking activity in

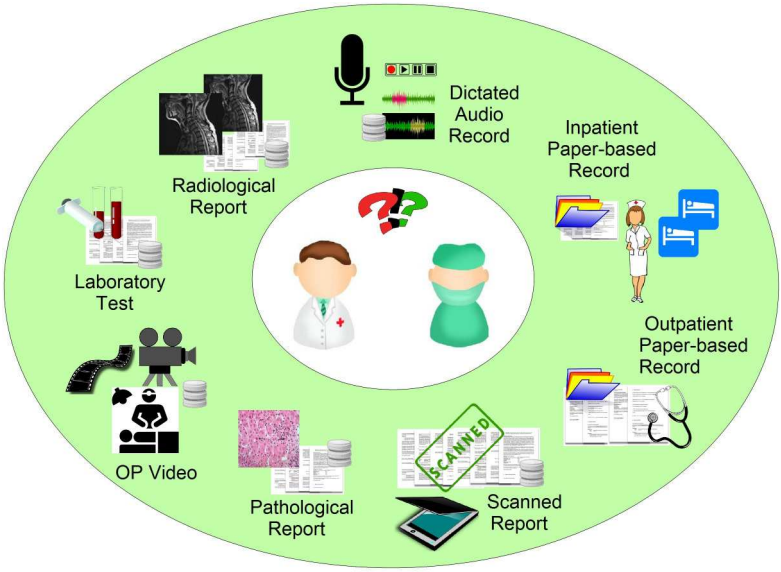

Figure 1: Status Quo of the Clinical Information System

the CIS is always a time consuming and labor-intensive task, especially for retrievals involving several media types and multiple document hierarchy. Normally, physicians need to find out the connection and diagnostic clue among different media manually. Often, they have to spend several hours on seeking a scanned document or one slide of radiological image so that the effectiveness of the treatment can be confirmed. Sometimes, it is even impossible to find out the desired document by switching between different modalities. Moreover, the searching of dictated audio snippets can only be conducted by hand, it costs a lot of time to navigate to the target excerpt in entire audio tracks, whereas the retrieval of video can be solely achieved with manually "fast forward" and "backwards". The inconvenience by cross-media retrieval leads directly to time wasting and suboptimal planning of treatment. Even treatment error can occur due to a lacking information.

Considering such information retrieval task in critical situations as those in an operating room, it becomes clear that there is a need to support physicians in retrieving and accessing relevant information. During a surgery, a physician must not only concentrate on the operation, but also make sure that each step of the procedure is completed in full accordance with the operation plan and clinical evidences. However, the information entities are scattered among different modalities. In order to obtain detailed examination results, the surgeon has to suspend the operation process and ask the nurse or the resident to search and open relevant documents stored in the CIS. The cost for information seeking is quite high.

According to the situation in the operating room, some new information organization and linking mechanism should be implemented to help the physician in overcoming the multimedia information retrieval problems. The record of the patient should be reconstructed and presented to the physician so that a comprehensive and fore-thoughtful decision can

be made.
The remainder of this paper is arranged as follows: Section 2 raises the three research questions, which will be centered in this work. Section 3 presents related work of cross-media linking in medical domain. As next, the linking possibilities, discussion over research questions and future works are described in Section 4, 5, 6 respectively.

2 Research Questions

Due to aforementioned facts in CIS and the complex environment in the operating room, the following three research questions will be addressed in this PhD thesis:

1. **From a technical point of view, what kind of approach should be developed to realize a cross-media linking in a clinical context?**
2. **From a clinical point of view, which kind of requirements should be fulfilled for cross-media retrieval in the operating room?**
3. **Can cross-media linking techniques provide required accuracy for the usage in the operating room?**

3 Related Work

The cross-media linking is not a novel technology, as it has been already studied since the 90s of the last century. Most of methods tended to create appearance models to match the media entities with the same semantic meaning automatically. However, the application of cross-media in clinical context has still not been well explored. The reasons are twofold: First, clinical data is stored in a distributed manner, which limits its accessibility. Second reason is that the sophisticated clinical work flow involves more media in the entire process, which makes the clinical data linking and cross-media retrieval more complex than in other context. The medical context has also advantages in comparison to a general domain: Various biomedical ontologies such as UMLS[1], RadLex[2], FMA[3], etc. have been developed to summarize the terminologies and their dependencies, which provides resources for semantic annotation and inference. Moreover, the concept of digital patient modeling [Den13] has extended the semantic description between patient records from different media based on probabilities model. These knowledge bases provide the semantic foundation of data linking between clinical records from terminology level to document (record) level.
In a CIS, the image data takes the largest part of the data volumes. The current research of clinical cross-media data management has therefore focused on the linking between text

[1] http://www.nlm.nih.gov/research/umls/, accessed 11.06.2014
[2] https://www.rsna.org/RadLex.aspx, accessed 11.06.2014
[3] http://sig.biostr.washington.edu/projects/fm/AboutFM.html, accessed 11.06.2014

and image. In order to process and retrieve the image in a comparable way as textual data retrieval, the corresponding model should be developed to interpret the image or store the interpretation of image. Zrimec et al. [ZLKS09] have shown a system named *Medical Image Assistant (MIA)*. It uses domain knowledge about human anatomy and pathology as well as the medical image protocols to organize the data from the PACS system [Hua04]. The image data is transformed into a set of numerical descriptors, color, texture and structure components, while the lung anatomy and disease patterns are extracted based on high resolution CT images automatically. Attributes such as color histogram, average color intensity and texture first order are calculated for every raw image by the MIA system. During the retrieval, the features from medical protocol and CT images are stored in the global description index. The MIA has employed a relatively less complex semantic representation using knowledge based image analysis to link the multimedia data and provided a swift data access. Kozuka et al. [KTK+13] have presented a new similarity comparison method for the image retrieval. It combines image features and description extracted from finding report to facilitate the cross-media retrieval. First, the diagnostic knowledge is extracted from previous cases. Then, the correlation between diagnostic knowledge and image features is determined and the most effective features for the retrieval are derived. For the retrieval of similar cases, a physician needs to choose firstly a region of interest and input short description from finding reports referring to the chosen region. The system extracts image features from the specified region of interest. Meanwhile, the diagnostic knowledge is used to weight the image features related to keywords. In this way, the extracted knowledge from past cases can adapt the weights of image features dynamically. According to the result presented in [KTK+13], the concordance rate of texture classification is improved from 57.7% to 75.9% and the precision of subjective evaluation is improved from 61.2% to 71.3%.

4 Simulated Scenario and Requirement in the Operating Room

With the aim of understanding the real information needs of a surgeon, a series of clinical work shadowing have been conducted. The typical information needs from surgeons are, e.g., "I want to see all information related to that particular operation or the medical procedures made during the staying of a patient" or some follow-up question like "The system says there were complications occurring during the treatment of patient X. Are there some video recordings of the operation or some nurse feedbacks or differential diagnoses?".
In order to answer these questions, the system should provide an overview considering detailed dependencies among different types of media. As can be seen in Figure 2, the clinical data stream from different media for one patient can be implicitly or explicitly connected based on dependencies. If a surgeon needs to make some prompt retrieval of patient data and examination results in an operating room, the system should be able to provide not only single relevant document but also a series of relevant documents regarding dependencies rapidly. The fine-grained relations between entities within one document or between documents should be also presented to the surgeon. Moreover, a user-friendly access interface should be offered so that the surgeon can raise the query conveniently.

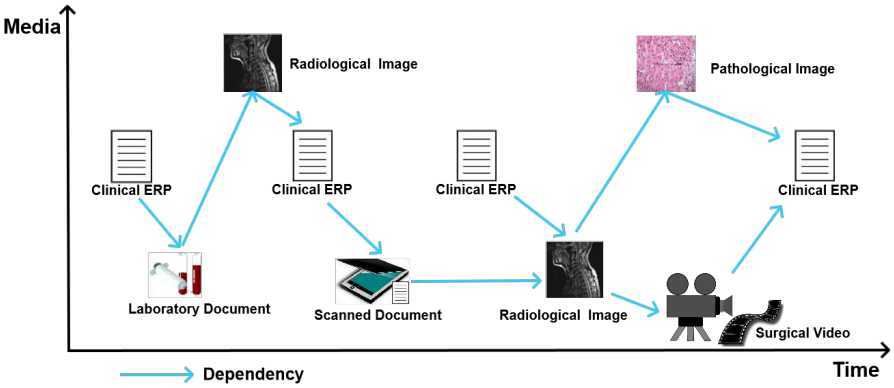

Figure 2: Dependency Chains from Different Media for One Patient

5 Linking Possibilities

Various connections among media are of interest in an operating room. The linking possibilities between text and other media types are crucial, since the text has constituted not only the representation within knowledge bases and other textual resources, but also the connection between natural language utterance. If other media can be annotated with text, the semantic roles and dependencies can be inherited from well-defined knowledge bases. Thus, the semantic description in knowledge bases can be expanded to multimedia resources. Due to the scope of the PhD study, it is assumed that each multimedia document or snippet is annotated with the corresponding text tags. Hence, the multimedia linking will mainly focus on the linking of multimedia resource through knowledge mapping. The special snippet in image, scanned document, video and audio labeled with the same semantic meaning will be clustered together. The links types between different media documents or snippets will be derived from a knowledge base such as medical ontology or patient modeling for certain diseases. Reversely, the machine learning techniques can also be applied to discover the newly emergent terminologies and dependencies so that the potential update entry for knowledge bases can be offered.

6 Conclusions and Future Work

As a preliminary stage of the PhD program, the status quo of CIS and requirement for cross-media linking in the operating room have been presented based on the clinical work shadowing and the state of the art analysis. The three research questions are answered as follows:

- **RQ1:** For cross-media linking, a unified data structure should be defined to map the documents from different media so that they can be grouped together according to

their content, while their relations can be determined based on medical ontologies and patient modeling for special disease through concept mapping. The user could get the possibility to "see both the forest and the trees". Especially for large data volume, the implicit clues hidden in the multimedia documents can be explored through the linking process, which offers additional serendipity for users.

- **RQ2:** From a clinical point view, a successful retrieval in the operating room depends on the full accordance of surgeons' need for information seeking. A swift delivery of configurable overview of patient status and an easy interaction interface will play the crucial role. The audio query interface should also be integrated to let the surgeon generate the query by their voices so that the surgeon can conduct non-stop operation with both hands.

- **RQ3:** Under the assumption that each multimedia document or snippet is annotated with the corresponding text tags, an acceptable accuracy and efficiency can be expected by using various of medical ontologies, patient modelings and machine learning methods. For optimization, the compromise between precision and recall should be decided by concrete scenarios, e.g., a result with high precision and moderate recall can help the surgeon focusing on several most relevant information instead of a result set with several pages.

At the present stage, the real data obtained from surgical data stream will be further analyzed. The features from different media will be compared. The experiment for the linking between finding report and radiological image will be conducted as a first step. In the short term, the prototype for cross-media linking will be implemented to audio and video resources. In the long term, the validated prototypical system will be integrated in the operating room to evaluate its performance.

References

[Den13] Kerstin Denecke. Model-based Decision Support: Requirements and Future for its Application in Surgery. *Biomedical Engineering*, 58(1), 2013.

[Hua04] H.K. Huang. *PACS and imaging informatics: basic principles and applications*. Wiley-Liss, 2004.

[KTK+13] K. Kozuka, K. Takata, K. Kondo, M. Kiyono, M. Tanaka, T. Sakai, and H. Kimura. Development of image-retrieval technology using radiological knowledge extracted from a clinical database of lung CT images. In *International Journal of Computer Assisted Radiology and Surgery*, pages 71–74, 2013.

[Rop99] Günter Ropohl. Philosophy Of Socio-Technical Systems. *IN SOCIETY FOR PHILOSOPHY AND TECHNOLOGY 4(3)*, 4:59–71, 1999.

[ZLKS09] Tatjana Zrimec, Iztok Lajovic, Kres, and Kreativni Sistemi. A Medical Image Assistant for Efficient Access to Medical Images and Patient Data. *Journal on Information Technology in Healthcare*, 7:261-271:pp.261–271, 2009.