〔**Original Paper**〕

A structured reporting system supported by speech recognition for electronic medical records

Hideto Yokoi

SUMMARY

Background: Structured reporting provides a great advantage in the secondary use of entered information. Although speech recognition is increasingly popular as a man-machine interface, the results of recognition are often provided by free-text reports. Data in prose form are not efficient for secondary usage.

Objectives: To develop a system for judging whether terms entered by speech recognition fit to the preset structured terminology. When the terms fit to the structure, they are stored in a database; otherwise, the system displays an alert to prompt correction of the term.

Methods: The system was designed for use on personal computers. The structured terminology is based on the "Minimal Standard Terminology Ver. 2 Japanese version", which was developed for gastroenterological endoscopy reporting. We compared two kinds of reporting methods, data entry through the speech recognition structured reporting system（SRSR）and through conventional handwriting.

Results: The average entering time by SRSR was 40% shorter than that by handwriting. Six of 168 words were recognized incorrectly; and were recognized correctly within second retry.

Conclusions: Data entry time, including additional free-text entries, by keyboard was shorter than that by handwriting. The SRSR system also has the potential advantage of hands-free data entry during endoscopy procedures.

I．Introduction

Electronic medical records（EMRs）are increasingly used by hospitals in Japan. Interpretations of medical images and reports, as well as ordinal clinical descriptions, are important

Department of Medical Informatics, Kagawa University Hospital, Kagawa 761-0793.
横井英人：音声認識を用いた構造化レポーティングシステム．
香川大学医学部附属病院医療情報部
Tel. 087-891-2381. Fax. 087-840-2601. E-mail: yokoi@med.kagawa-u.ac.jp
2008年10月 9 日受付，2008年10月22日受理．

aspects of EMR. The number of requirements of reports has been increasing with advancement of medical imaging modalities.

We consider that speech recognition is one of useful data entry technologies especially in radiology departments for image interpretation reports. However, the results of speech recognition are used as poorly structured free-text data and they are unfit for secondary use. Here, we describe a structured reporting system for medical images using a speech recognition engine to solve the problem with stress-free usability.

I-1 Significance of a structured reporting

Many types of image interpretation are reported by experts (i.e., radiologists, endoscopists). Thus, we call this information "reports (from experts to referring physicians)", and the systems which support these reporting procedures are called "reporting systems".

The basic process of entering information into reporting systems does not differ markedly from other health recording systems. Chiba University Hospital has developed an EMR system that is able to receive medical findings through "templates" to save appropriate data in designed fields of the database. The templates are entry forms for data capture similar to questionnaire forms on the internet. When entering data into an EMR system, a physician can choose a suitable template for an individual patient or his/her disease. The physician enters information by filling in the fields of the template system.

Both field attributes and representations to be entered are unified in the hospital system. For example, in a diabetic patient, an attribute of "thirst" provides a list of choices (−, −+, +, ++, +++).

We consider that the template concept enables structured data preservation. Moreover, if a standard terminology is used in a term list,

the template system is effective for data sharing between certain institutes.

There are two methods for information storage. One is a free-text (prose) form, and the other is a structured one. The "structured" form is preferred for secondary use (i.e. searching, analysis) of information. Many narrative text-reporting forms have fields that are broadly defined (i.e., "reasons for examination", "findings", "diagnosis", "recommendation") and are considered to be structured forms. In the broad definition of "structured reports", reporting systems using structured forms are known as "structured" systems. However, in the present study, we defined structured report in a more narrow sense, as a database with data fields for adequately granular concepts. Additionally, we did not discard narrative data, as significant amounts of information cannot be described by the preset structured terminology. However, as free-text data will inevitably include inaccurate concepts, data cleansing is necessary for secondary usage.

I-2 Conventional reporting and the proposing system

Structured reporting is important not only for secondary usage but also for standardized universal storage with a set terminology[1]. However, some reports are not suitable for structured reporting; in image interpretation, for example, reports are divided into two patterns. One consists primarily of measured values and does not depend on sentences (i.e., cardiac echo), the other is narrative (i.e., radiological interpretation).

Previously, numerical data were stored in fields for holding structured attributes, and displayed or analyzed in a time series. However, almost all radiological reporting systems in Japan are currently capturing free-text reports. Radiologists are thus required to analyze scans of various locations acquired by several different

modalities. Thus, it is difficult to standardize the terms for radiological reports. Although a radiological lexicon (RadLex) has recently been developed[2], it has not been translated into Japanese, yet.

Radiologists are entering free-text data with reference to standard models of reporting sentences. Using the standard models, radiological reports can be held to an adequate standardization of expression and can be easily read by physicians. This free-text feature of easy understanding is highly demanded by referring physicians; thus, several structured reporting systems have a free-text report-generating module for structured findings[3].

Although it is important to clarify the meanings and benefits of free-text reports, the standardization of reporting terms and attributes is equally important.

Additionally, most medical information systems confine users to desktop computers due to the nature of their interface devices (mouse and keyboard), thus excluding the system from special hospital areas (i.e., operating rooms). The present system aims at releasing users from this restriction.

We, the first author and system developers, established a reporting system supported by a speech recognition system for 2 image-based examinations (abdominal ultrasonography in 2001 and upper GI endoscopy in 2002). From these experiences, we developed a conversion scheme from free-text reports to structured reports supported by a speech recognition system. We named this system the "Speech Recognizing Structure Reporting (SRSR) system".

II. Experimental application

II-1 Experimental system design

The SRSR system was developed using Microsoft Visual Basic Ver. 6.0, running on

Fig. 1 Spreadsheet of MST structure

The terminology has a hierarchical structure and each term is described with its pronunciation for speech recognition and by attributes that show entry rules for the term.
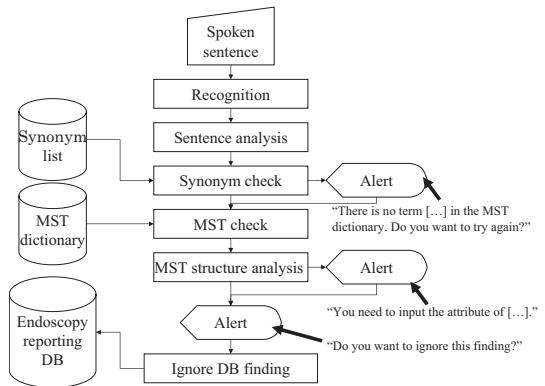

Fig. 2 Structured recognition procedure

The "Synonym list" and the "MST dictionary" are generated from a structure defining spreadsheet file. The "Synonym list" is derived from multiple pronunciation data. The list includes synonyms for the preferred term used in MST.

Microsoft Windows XP Professional SP1. The speech recognition system was AmiVoice SDK Ver. 4.0.

The pre-set terminology was prepared as a Microsoft Excel worksheet (Fig. 1). The terminology followed specific rules. We were able to express pronunciations and relation of terms on a single worksheet. The information has a feature of knowledge base for analyzing a finding's structure[4]. The system has a function for importing the structured terminology from the worksheets, analyzing the
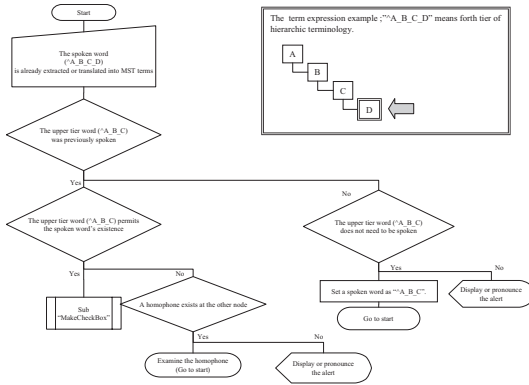
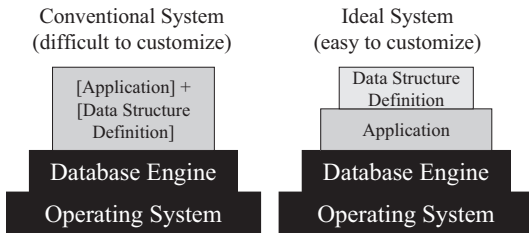Fig. 3　Flowchart of the structure check routine

Fig. 4　Two types of reporting system architecture

terminology, and creating a grammar file for AmiVoice. When a reporting sentence is spoken, the system recognizes the terms and decides whether or not they fit to the MST structure. When the terms fit to the structure, they are stored in a database; otherwise the system gives the speaker an alert to prompt correction of the term by a synthesized voice or an on-screen message. While previous studies have used a natural language processing system to output structured reports from free-text reports[5], we did not use a special module except the natural language processing function of the speech recognition system. Figures 2 and 3 show the analysis algorithm of the system.

Ⅱ-2　Methods

We installed the MST (Japanese version) in the SRSR system to investigate its usefulness. Several structural changes were needed to fit the MST to the system, although the new structure is equal to the previous system from the informatics perspective. We compared two types of reporting methods, using report contents already released 1) data entry through the SRSR and 2) conventional handwriting. Report contents were extracted from 10 randomly selected upper GI endoscopy reports conducted at Chiba University Hospital in 2002.

In this experiment, we used a screen alert instead of a voice alert because the voice alert was considered time-consuming.

Ⅲ．Results

Table 1 shows the time required by the SRSR and handwriting methods. Our system can only accept terms in the installed terminology. For example, SRSR cannot understand "There is a piece of meat in the stomach" because "meat" is not included in MST, although some basic sentence patterns were loaded previously. The SRSR user must then add additional free-text expressions using the keyboard. Thus, the total speech recognition time includes both speech recognition and time for keyboard entry.

The recognition rate calculated in this experiment was from 10 reports, consisting of 49 findings. A finding means an observational

Table 1　Report entry times

Case	Speech Recognition (sec)			Handwriting (sec)
	Recognition	Additional	Total	
1	58	65	123	168
2	15	72	87	134
3	52	18	70	156
4	25	40	65	90
5	20	0	20	49
6	28	0	28	90
7	27	0	27	98
8	10	24	34	65
9	46	0	46	84
10	24	70	94	106
Ave.	30.5	28.9	59.4	104.0
S.D.	15.2	29.1	32.5	36.2

record of each lesion, such as "There was an [ulcer] at [body] of [stomach], which [shape] is [linear], and [bleeding] type is [oozing]" (words in brackets are MST terms). Thirty-five findings were recognized correctly at the first trial, and six were recognized incorrectly; eight findings consisted of non-MST terms. We calculated a sentence recognition rate of 85.4% (35/41) after removing eight non-MST sentences from the analysis. To report these eight non-MST findings, keyboard entry was required. The synonym list was used for 14 findings (34.1%), and 27 findings (65.9%) consisted of original MST terms.

In total, 168 MST terms were used in the ten reports; six terms were recognized incorrectly. Four of these six incorrectly recognized terms were recognized correctly at the second trial, and two terms at the third trial. In the experiment, no errors in morphological analysis of the sentences were observed. The word recognition rate was 96.4%.

IV. Discussion

IV-1 Customization problem of conventional reporting systems and our solution

Many structured reporting systems consist of a screen with a number of input fields. While the field layouts are not expected to fit all users' preferences, few systems have a customizing function. Thus, some of them cannot expect favorable user satisfaction.

We consider that many systems lack suitable customizing ability. This could be resolved by dividing reporting systems into two parts, a fixed system core and alterable applications that can be changed to meet user needs (Fig. 4). However, as reporting categories are not yet mature, system vendors are currently unable to divide reporting systems into these parts.

We developed a reporting system that combines EMR and computerized physician order entry systems. The reporting system has an application that provides a user interface through web-browsing software. Web contents can be used to create reporting interfaces and are easily uploaded to the system. Additionally, JavaScript and other script languages can be used in the system if they are accepted by the web browser. The system also gives system vendors' high productivity. Although we believe the advantage of the system is its flexibility and productivity, we recognize that it may not sufficiently improve upon the standard mouse and keyboard.

IV-2 Development of speech recognition interface

Our developed system has a function of recognizing not only the phonetic findings of images but also the predefined structure of the terminology. After recognizing the finding terms, it retrieves potential structures, identifies and prompts the user for missing information, and then displays the appropriate structure. This type of real-time reporting that enables the recording of findings at the time of detection is ideal for GI endoscopy[6]. In addition, the interaction of user and speech recognition system is necessary for precise and well-structured reports.

In the U.S., speech recognition in MRI reporting is considered cost-effective[7]. However, the reporting system in Japan is different, as it is not supported by medical transcribers. Instead, most radiologists use speech recognition for radiological reporting using free-text input forms. Thus, there is a substantial demand for prose radiological reports instead of speech recognition reports, and report data are not suitable for searching and analysis.

The standard terminology for GI endoscopy is the MST established by the Organisation Mondiale d'Endoscopie Digestive (World

Organization for Digestive Endoscopy; OMED) [8]. This terminology has structured relations of terms and was intended for use as a terminology for database entry. The Japan Gastroenterological Endoscopy Society (JGES) developed the Japanese version of the MST ver. 2, which is now used as the standard terminology[9]. We chose the JGES MST as the preset structure of the SRSR system, and performed a simulation of endoscopy report data entry.

Ⅳ-3 Entry time

The time required to enter the ten reports into the database by SRSR (289 seconds) was shorter than that by handwriting (1040 seconds), even including the time required for additional keyboard entry, and fatigue was quite low. Non-MST terms which we considered to be necessary for reporting were set before the experiment, consisting of certain verbs, auxiliary verbs, postpositional particles (of Japanese) and numerals, based on knowledge from our previous investigation on abdominal ultrasonography reports.

We conclude that SRSR is a useful system with shorter entry time and gives less fatigue than handwriting to the user. Additionally, the system could be used while performing endoscopy due to its hands-free nature. Thus, the total reporting time could be significantly shorter than that of conventional reporting procedures

Ⅳ-4 Flexible application to other fields

Developing a structured and standardized reporting system is difficult because it requires knowledge of terminology, databases, and management of medical workflow. Moreover, few practitioners in Japan have experience in using a speech recognition system, and numerous minor changes are required to adapt to users' preferences. In consideration of this situation, we designed our SRSR system to be user-friendly by linking a table of terminology with a grammar file for speech recognition. The grammar file is written in "Java Speech Grammar Format", a macro language for speech recognition, and is suitable for describing a structured reporting content. Users only need to prepare the structured terminology on an Excel worksheet; the system will then generate a grammar file automatically.

Spreadsheet software has a very useful interface for editing structured content. We can describe content in a Microsoft Excel cell using set attributes such as pronunciation, required speech and so on. We defined the following three attributes that provide a relation of terms.

A. Required speech: [necessary] [optional] [not necessary]
B. A necessity of subclass terms speaking: [necessary] [optional]
C. A number of subclass terms choosing: [single] [multiple]

Using these attributes, we can describe the structure of a report and manage the structured reporting procedure on the spreadsheet software. Flexible application to various fields is an important feature of speech recognizing systems. Sistrom et al described a structured report management software[10]that enables different departments to prepare and maintain their own reporting format.

Ⅴ. Conclusion

We developed the SRSR system for a structured, standardized, and useful reporting, considering usability and obtained favorable results. Trials of establishing useful terminologies for structured reporting should be conducted by all related people such as users and vendors and, their collaborations will uncover realistic methods of applying structured reporting to EMR systems.

Acknowledgements

要　　旨

【背景】構造化レポーティングとは，自由文と違い，情報の再利用性に優れている。近年，音声認識システムが普及しているが，認識結果はフリーテキスト（自由文）であり，二次利用には適さない。

【目的】音声認識システムにより入力された文を，構造化された用語集に適合しているかを判断し，もし適合している場合，その構造に従ってデータベースに適切に保存し，適合していない場合，修正要求を発するシステムを開発する。

【方法】システムはPC上で動作するように設計し，消化器内視鏡所見用の用語集であるMinimal Standard Terminology Ver. 2日本語版を構造化用語集として採用した。実験として，千葉大学で行われた消化器内視鏡所見について，音声認識を用いた構造化レポーティングシステム（SRSR）と手書き入力の二手段で記録し，比較をした。

【結果】平均入力時間はSRSRの方が手書きより40%短かった。168語中，6語が誤認識されたが，2回目の認識で4語が，3回目の認識で2語が正しく認識された。

【結論】SRSRのデータ入力時間は，構造化用語集で入力できない付随情報をフリーテキストで入力した時間を含めても，手書き入力よりも短かった。本システムは，内視鏡施行中にハンズフリーで音声入力をすることに用いれば，更なる業務の効率化が期待される。

References

1 ）Morioka CA, Sinha U, Taira R, el-Saden S, Duckwiler G, Kangarloo H. Structured reporting in neuroradiology. Ann NY Acad Sci 2002; 980: 259-66.

2 ）Daniel LR. Creating and curating a terminology for radiology: ontology modeling and analysis. J Digit Imaging 2007 Sep 15:［Electronic publishing］

3 ）Langlotz CP. Enhancing the expressiveness of structured reporting systems. J Digit Imaging 2000; 13（2 Suppl 1）: 49-53.

4 ）Rosenthal DF, Bos JM, Sokolowski RA, Mayo JB, Quigley KA, Powell RA, Teel MM. A voice-enabled, structured medical reporting system. J Am Med Inform Assoc 1997; 4: 436-41.

5 ）Sinha U, Dai B, Johnson DB, Taira R, Dionisio J, Tashima G, Golamco M, Kangarloo H. Interactive software for generation and visualization of structured findings in radiology reports. AJR Am J Roentgenol 2000; 175: 609-12.

6 ）Molnar B, Gergely J, Toth G, Pronai L, Zagoni T, Papik K, Tulassay Z. Development of a speech-based dialogue system for report dictation and machine control in the endoscopic laboratory. Endoscopy 2000; 32: 58-61.

7 ）Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. AJR Am J Roentgenol 2000; 174: 617-22.

8 ）World Organization for Digestive Endoscopy, Minimal standard terminology for a computerized endoscopic database. Working Party Report by the Committee for Minimal Standards of Terminology and Documentation in Digestive Endoscopy of the European Society of Gastrointestinal Endoscopy（ESGE）. http://www.omed.org/index.php/resources/re_mst/.1999

9 ）The Japan Gastroenterological Endoscopy Society（JGES）. http://www.jges.net/index.html.2001

10）Sistrom CL, Honeyman JC, Mancuso A, Quisling RG. Managing predefined templates and macros for a departmental speech recognition system using common software. J Digit Imaging 2001; 14: 131-41.