

음성인식기반의 전자의무기록 구현을 위한 오인식 의학용어 수정 시스템 개발

A Mis-recognized Medical Vocabulary Correction System for Speech-based Electronic Medical Record

Kildong Seo, Ph.D.* , Jong Hoon Kim M.D., Ph.D.† , Taeho Choi, M.D.*

* Dept. of Medicine, Asia University School of Medicine, Seoul 123-456, kdseo@asia.ac.kr

† Dept. of Medical Informatics, East University Hospital, East University, Cheonwon-gun, Chungbuk 345-678

Abstract: Speech recognition as an input method for electronic medical record system makes data entry efficient at the point of care. The speech recognition accuracy for medical vocabulary, however, is much poorer than that for doctor-patient dialogue. We developed a mis-recognized medical vocabulary correction system based on syllable-by-syllable comparison of speech-recognized text against medical vocabulary database. Using specialty medical vocabulary, the algorithm detects and corrects mis-recognized medical vocabularies in the narrative text. Our preliminary evaluation showed 94% of accuracy in mis-recognized medical vocabulary correction.

Keyword: *electronic medical record, speech recognition, data entry, natural language processing*

I. 배경

본 연구는 음성인식 기술을 이용한 전자의무기록(EMR: Electronic Medical Record)시스템에의 효율적인 자료입력을 통하여, 단순히 행정적 목적으로 사용되던 시스템을 환자를 진료하면서 의료관계자가 직접 신속하고 정확하게 사용할 수 있는 임상정보시스템으로 전환하는 데 그 목적이 있다. 자료입력 방법으로서의 음성인식 기술은 신속하게 임상정보를 입력할 수 있다는 장점이 있다. 그러나, IBM ViaVoice™의 음성인식의 정확도를 평가해 본 결과, 의사의 질문과 환자의 대답 형식으로 구성된 일반적인 의사-환자간 대화의 음성인식 정확도는 95%로 전자의무기록 시스템에 적용 가능한 인식률을 보였지만, 환자의 진단 및 처방시 사용되는 전문 의학용어의 인식 정확도는 55%로 현저히 낮았다.

II. 방법

본 연구에서는 오인식률이 높은 전문의학용어의 인식 정확도를 높이기 위해 전문 의학용어 데이터베이스 기반 추천단어 추출 알고리즘을 적용한 오인식 의학용어의 수정 시스템을 개발하였다.

오인식 의학용어 수정 시스템은 오인식 의학용어의 추출단계와 정정단계로 구성된다. 음성인식 후에 발생하는 오인식어를 추출하는 단계는 1) 음성인식 결과인 의사-환자간의 대화 또는 의사의 처방을 음절 단위로 분해한 후, 2) 각 음절에서 의학용어 데이터베이스내의 의학용어들과 유사성을 비교한다. 유사율이 50%이상인 단어는 추천단어로 추출된다.

오인식어를 정정하는 정정단계는 1) 음절의 코드와 위치가 모두 동일한 경우와 2) 음절의 위치는 동일하지 않으나 동일한 음절의 코드가 추출단어에 존재

하는 경우로 나누어 가중치를 부여한다. 각 오인식 의학용어의 추천 단어들 중에서 가중치가 가장 큰 의학용어로 이 오인식어를 수정한다.

III. 결과

33개의 전문 의학용어가 포함된 의사-환자간 대화(33문장)를 6명의 피험자가 각각 3번씩 발성한 음성인식결과 (i.e., $33 \times 6 \times 3 = 594$ 문장)를 오인식 의학용어 수정시스템으로 시험한 결과, 오인식 의학용어 수정 정확도는 94%이었다.

IV. 결론

오인식 의학용어 수정 시스템은 인식된 문장과 단어의 어간 공백을 무시한 문장에서도 의학용어로 추정되는 단어를 추출해 낼 수 있도록 설계되었다.

또한, 오인식률이 높은 의학용어를 추출한 뒤, 가중치 부여에 의하여 제시되어지는 최적의 추천단어로 수정되므로, 본 시스템은 음성인식기반 전자의무기록 시스템 구현의 가장 큰 문제점 중의 하나인 전문 의학용어 오인식률을 개선할 수 있을 것으로 기대된다.

참고문헌

1. Tsuzaki H, Hasegawa T, Onozuka M. Research into a Radiology Reporting System Using Speech Recognition-Experience of Field Test at Hospi-net, The Institute of Telemedicine; 2000;5:4-5.
2. Rector AL, Glowinski AJ, Nowlan WA. Angelo Rossi-Morri. Medical-concepts Models and Medical Records-An Approach Based on GALEN and PEN&PAD. 1999 Yearbook of Medical Informatics 1999;:225-241.