

음성인식 기반 지능형 노래 반주 시스템 개발

Developing a Speech-Recognition-Based Intelligent Karaoke Band System

오영진*, 장문수**, 정태의*, 강선미*

서경대학교 컴퓨터과학과, 서경대학교 소프트웨어학과

Youngjin Oh*, Moonsoo CHang**, Tae Eui Jeong*, Sunmee kang*

Seokyeong University Computer Science*, Seokyeong University Software**

E-mail : pennant@ihci.skuniv.ac.kr

요 약

가정용 혹은 업소용 노래 반주 시스템에서 곡명을 선택하는 데에는 목록 책자를 찾거나 키보드를 이용하여 DB를 검색하는 방법이 일반적이다. 이 방법들은 책을 찾거나 키보드로 입력하는 불편함과 함께 곡명의 첫 글자부터 알고 있지 않으면 찾을 수 없는 문제점이 있다. 본 논문에서는 음성인식기술과 정보검색기술을 결합하여 이러한 문제점들을 동시에 보완하는 기술을 제안하고, 가정용 노래방 기기와 연동하여 기존 검색 방식을 대체하는 음성인식 노래반주 시스템을 구성한다. 제안하는 시스템은 노래 곡명에 나오는 최다 빈도 단어를 이용하여 키보드 입력 없이 음성 인식으로 간단하게 원하는 곡명을 찾도록 한다. 또한, 인터넷 인기순위에 기반하여 곡명을 색인함으로써 검색 만족도를 향상시킨다.

1. 서론

국민 소득 증대에 따른 문화 욕구가 증대함으로써 소비자의 선호도가 기능의 다양화로 변화하고 있는 추세에 있다. 노래방 기기 분야에서도 이러한 요구가 자연적으로 제기되고 있다.

현재 상업 및 가정용 노래 반주기에서 노래를 선택하는 방법에는 노래곡목책자를 이용하거나 키보드를 이용한 키 입력 방법 등이 있다. 기존 검색 방법은 검색시간이 오래 걸리고, 곡명의 첫음절을 기억하지 못하면 곡목을 찾을 수 없다.

본 논문에서는 노래반주기에 음성인식과 정보검색기술을 적용하여 별도의 입력장치가 없이 음성으로 제어가 가능한 지능형 노래반주 시스템을 제안한다. 제안하는 시스템은 곡명에 나오는 고빈도 단어를 이용하여 음성인식기술

의 제약을 보완하여 곡명을 판별한다. 또한 인터넷 인기 순위를 기반으로 곡목을 색인함으로써 검색 만족도가 높은 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 시스템을 음성인식부와 곡목검색부로 나누어 설명한다. 3장에서는 곡명 인식결과 및 사용자 만족도에 대한 실험 결과, 마지막 4장에서는 본 논문의 연구 결과 및 향후 과제를 제시 한다.

2. 지능형 노래반주 시스템

기존 곡명선택 방법으로는 노래책자와 키보드를 이용하는 키 입력 방법 등이 있다. 노래책자를 이용한 검색은 첫음절의 자음을 검색 후 모음을 검색한다. 이를 반복하여 검색을 하면 시간이 오래 걸린다는 문제와 자음과 모음의 순서를 알고 있어야 되는 문제점이 있다. 또

한, 키보드를 이용한 검색은 키보드에 익숙하지 않은 고연령층에게 조작의 부담을 줄 수 있고 어두운 곳에서 키보드의 키가 잘 보이지 않는 문제점이 있다.

본 논문에서 위와 같은 문제점을 보완하기 위하여 저능형 노래 반주 시스템을 제안한다.

저능형 반주 시스템은 음성인식부와 곡명검색부로 구성된다. 음성인식부는 DSP기반 음성신호처리 시스템으로 음성인식 대상을 시스템에 탑재하고 있다. 곡명검색부는 곡명의 일부만으로 검색이 가능하도록 하기 위하여, 형태소분석[1] 및 한국어 태깅[2] 과정을 거쳐 고빈도 키워드를 추출한다. 추출한 키워드의 곡명은 인터넷 노래방 인기곡 순위에 근거하여 순위가 결정된다.

2.1 음성인식부

현재 서버급 음성인식 시스템은 일부 적용분야에서 완성단계에 이르러 상용제품이 출시되어 있다. 예로는 음성 전화번호 등록을 할 수 있는 ‘음성인식 선불카드’[3], ‘음성인식 교통정보서비스’[4] 등이 있다. 이러한 음성인식 방법은 성능이 뛰어난 고가 장비들을 사용하는 기술들로 노래반주기에 사용되는 소규모, 소용량의 임베디드 시스템에서는 음성인식 엔진의 크기와 음성인식 엔진의 Working Memory 문제 등으로 개발 정도가 미미하다.

제안하는 시스템에서는 노래반주기와 IrDA를 이용한 인터페이스를 갖춘 임베디드 형태의 음성신호처리와 음성인식을 위한 TI사의 DSP칩과 SDRAM을 사용한다. 노래반주기와 인터페이스는 반주기에서 사용하지 않는 제어코드를 음성인식용 명령어로 표·1과 같이 할당하여 사용한다.

제어코드	명령어
0x4f	리스트의 시작
0x4e	리스트의 끝
0x4d	"단어를 말씀하세요" 표시
0x0a	"원하시는 단어가 없습니다" 표시
0x03	"잠시 기다려 주세요" 표시
0x07	노래시작

표 1 제어코드

키워드와 곡 번호는 DB로 시스템에 내장되어 인식을 한다.

소용량의 메모리를 사용하여 데이터 처리를 사용하는 본 시스템은 100개 이상의 단어인식

을 사용할 경우 인식률이 떨어져 전체 곡명을 인식대상으로 할 수 없는 한계를 가지고 있다.

2.2 곡명검색부

본 논문에서는 곡명에 출현하는 고빈도 단어를 추출하여 색인함으로써 적은수의 키워드로 최대한 많은 노래를 검색할 수 있도록 한다. 고빈도 단어를 추출하기 위하여 노래방 책자(본 논문에서는 청람디지탈사의 CLK-102모델의 노래반주기를 대상으로 실험하였다. 곡명은 노래방 책자에 실린 곡명을 추출하여 사용하였다.)를 토대로 곡명, 가수이름, 곡번호 등 총 6280개의 노래에 대한 정보를 선별했다. 노래제목은 한글과 영문을 포함한 곡명과 한글곡명, 영문곡명 3가지로 분류를 나눌 수 있다. 음성인식 대상을 한글 키워드로 제안하여 적용하였기 때문에 영문이 포함된 곡명을 가진 노래는 단어 추출범위에서 우선적으로 제외를 시켜 표 2와 같이 곡명을 선정 한다.

곡번호	제목	가수	영문명
14	겨울	정미소	말씀드 무슨 날
15	겨울도 안되는 여자	태진아	사랑이야 왜해도는
16	결혼할까	윤서함	네머지가 없잖아
17	결혼도 싫어	김성영	내 마음을 나의 맘에
18	고향의 봄	남궁다	코스모스 피어있는
19	고향무상	유기태	구름도 출고 넘는
20	김나루	남궁다	산 채비 넘는 고깃길
21	고향의 봄	김성영	울진 울진 울진 울진
22	고향의 봄	김성영	남해 나라 바다 물결
23	고향	홍인	고향의 봄
24	고향의 봄	윤서함	말씀드 무슨 날
25	고향의 봄	윤서함	말씀드 무슨 날
26	고향의 봄	윤서함	말씀드 무슨 날
27	고향의 봄	윤서함	말씀드 무슨 날
28	고향의 봄	윤서함	말씀드 무슨 날
29	고향의 봄	윤서함	말씀드 무슨 날
30	고향의 봄	윤서함	말씀드 무슨 날
31	고향의 봄	윤서함	말씀드 무슨 날
32	고향의 봄	윤서함	말씀드 무슨 날
33	고향의 봄	윤서함	말씀드 무슨 날
34	고향의 봄	윤서함	말씀드 무슨 날
35	고향의 봄	윤서함	말씀드 무슨 날

표 2 노래방 책자 정보

기존 노래방 선곡 기능에는 색인기가 없이 데이터베이스에서 단순히 키워드 매칭만을 이용하였기 때문에 곡명의 첫 음절을 알지 못하면 검색이 불가능하였다. 이에 본 논문에서는 수집한 곡명리스트에 대해서 형태소 분석을 실시하여 곡명전체로부터 필요한 키워드를 추출함으로써 곡명의 일부만으로 검색이 가능하도록 한다. 한국어 품사태깅 모델[3]을 사용한 분석 결과 중 명사(NN, NX, NP 등)와 형용사(AD), 동사(V, VI, VP 등), 조사(JX, JC 등)의 빈도수가 높다.[2] 형태소분석 결과에 대해서 품사별 빈도수를 측정하고, 고빈도 단어에 대해서 사용자의 키워드 선정선호도와 음성인식기의

제한등을 고려하여 다음의 휴리스틱규칙을 적용하여 음성인식에 사용한 키워드를 추출한다.

- ① 두 음절 이상의 명사나 동사 형용사를 사용한다. 한 음절의 경우 인식과정 중 오인식이나 인식거절을 할 경우가 많이 발생하여 두 음절 이상으로 한다.
- ② 형용사, 동사보다 높은 호출빈도를 갖는 명사는 키워드 선정시에 가중치를 높게 준다.
- ③ 고유명사는 일반명사에 비해 낮은 호출빈도를 가지므로 제외한다.
- ④ 한 음절로 된 대명사는 인식하기 어렵기 때문에 제외한다.

가 . 가/UU 어/EM
 가거라 삼팔선 . 가/UU 여라/EM 삼팔/NN 선/NN
 가거라 들을이여 . 가/UU 여라/EM 들프/AD 으/EM 이여/JO
 가까이 . 가까이/AD
 가까이 . 가까이/AD
 가까이 허기면 너무 먼 당신 . 가까이/AD 하/UU 기/EM 예/JO 는/JO 너무/AD
 가나다라 . 가나다라/NN
 가나다라마바사 . 가나다라/NN 마/NN 바사/NN
 가난함 그대 가슴에 . 가난/NN 하/NN 나/EM 그대/NN 가슴/NN 예/JO
 가난함 마음 . 가난/NN 하/NN 나/EM 마음/NN
 가난한 연인들의 기도 . 가난/NN 하/NN 나/EM 연인/NN 들/SF 의/JO 기도/NN
 가난함 가는정 . 가/UU 는/EM 남/NN 가/UU 는/EM 정/NN
 가난함 오는봄 . 가/UU 는/EM 봄/NN 오/UU 는/EM 봄/NN
 가난함 세월 . 가/UU 는/EM 세월/NN
 가난함 오는정 . 가/UU 는/EM 정/NN 오/UU 는/EM 정/NN
 가나 . 가/UU 나/EM
 가라 가라 . 가/NN 이/CP 라/EM 가/NN 이/CP 라/EM
 가라지 . 가/NN 이/CP 라/EM 가/NN 지/EM
 가라지 . 가/NN 이/CP 라/EM 가/NN 지/EM
 가라진 시간 사이로 . 가/UU 려/EM 지/UX 나/EM 시간/NN 사이/NN 로/JO
 가로수 그늘 아래 서면 . 가로수/NN 그늘/NN 아래/NN 서/UU 면/EM
 가리워 진 길 . 가리우/UU 어/EM 지/UU 나/EM 길/NN

그림 1 한글형태소 분석 결과

위의 규칙을 적용하여 빈도수가 높은 품사중 명사와 대명사 95개와 형용사 5개의 단어를 음성인식 노래방에 사용할 키워드로 표 3과 같이 선정한다.

위의 결과를 지능형 노래반주 시스템에 그대로 적용하면 '사랑' 키워드로 찾는 노래 곡명이 600번째에 있을 경우 열번 이상의 페이지 다운을 거쳐야하기 때문에 검색시간이 오래 걸리게 되는 문제가 발생한다.

품사	빈도수	키워드	품사	빈도수	키워드
NN	607	사랑	NN	37	눈물
NP	130	그대	NN	37	여인
NN	108	이별	NP	37	하늘
NN	73	사랑	NN	35	고향
NN	64	여자	NP	34	그녀
NN	62	세상	NN	34	이야기
NP	62	당신	NN	34	기억
NN	58	바람	NN	31	이유
NN	56	남자	NN	31	행복
NN	52	마음	NN	31	겨울
NP	51	우리	NN	30	편지
NN	50	친구	NN	30	시간
NN	48	마지막	NN	28	약속
NN	48	노래	NN	28	인생
NN	41	추억	NN	28	고백

표 3 고빈도 상위 30개 키워드

본 논문에서는 이 문제를 해결하는 방법으로 인터넷 노래 인기순위를 곡명의 가중치로 적용

하여 순위를 조정함으로써 검색만족도를 높인다. 그림 2 는 포털사이트 네이버에서 노래방 검색의 예를 보여 주고 있다. 노래방 인기순위는 시간에 따라 빠르게 변화한다. 제안하는 시스템은 검색어별로 곡명이 색인되어 음성인식 시스템의 DB에 축적되므로, 곡명 색인에 인터넷 메타검색[5] 기술을 이용하여 색인할 때 실시간으로 인터넷 검색순위가 적용 되도록 한다.

▶ 금주 노래방 인기순위

인기순위	노래제목	가수	작사/작곡
1위	용궁성(대결총합OST)	비	신동우/신동우
2위	영디발대마	SPAPA	최희건/김경우
3위	머더나	정윤경	윤경선/김영선
4위	사랑은...해리디	담미(하)	조은희/황세준
5위	올더서	최연	김정관/김형찬
6위	귀로	나눔	이민/이민
7위	행날(행날OST)	이이	이현우/박정일
8위	눈의꽃(대안하다사랑한다OST)	박호신	KENZIE/최국복
9위	선녀가 지나도(솔문연가OST)	송승현	류재현/류재현
10위	담백하지않대요	MC THE MAX	양재선/최국복

그림 2 인터넷 노래방 순위

3. 실험 및 평가

지능형 반주 시스템의 평가를 위하여 음성인식실험(인식률)과 곡명검색실험(검색만족도)에 대하여 실험을 하였다. 실험자는 서울대생 남성 4명과 여성 4명으로 선정하였고, 실험 환경은 조용한 환경에서의 지향성 마이크 사용, 시끄러운 환경에서의 지향성 마이크 사용, 조용한 환경에서의 무지향성 마이크, 시끄러운 환경에서의 무지향성 마이크를 사용하여 인식실험을 하였다. 곡명검색실험은 위의 조건에서 실험자가 50번의 똑같은 곡명의 검색과 각기 다른 곡명의 검색결과에 대한 검색만족도를 평가하였다.

표 4와 그림 3을 보아 주변 환경과 마이크 종류에 따라 인식률에 많은 영향을 주는 결과를 알 수 있다. 또한 남성 실험자가 여성 실험자에 비해 높은 인식률을 보였다. 성별에 따른 인식률의 차이는 남성과 여성의 음성의 크기의 차에 의한 것임을 음성파형 분석으로 알 수 있었다.

	조용한 환경, 지향성 마이크	시끄러운 환경, 지향성 마이크	조용한 환경, 무지향성 마이크	시끄러운 환경, 무지향성 마이크
남 자	88.4%	83.57%	85.2%	78.2%
여 자	83.2%	76.8%	78.2%	75.6%

표4 환경과 마이크 종류에 따른 성능

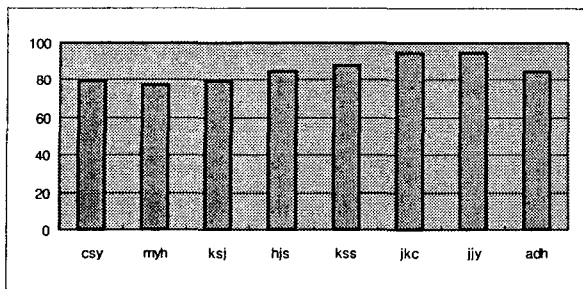


그림 3 인식률

검색만족도측정을 위하여 실험자에게 곡명이 같은 50개의 곡명에 대한 검색과 실험자 자유에 의한 50개의 곡명을 검색하였다. 100개의 키워드를 사용하였으며 노래곡명의 검색만족도는 그림 4와 같이 높은 만족도를 보였다.

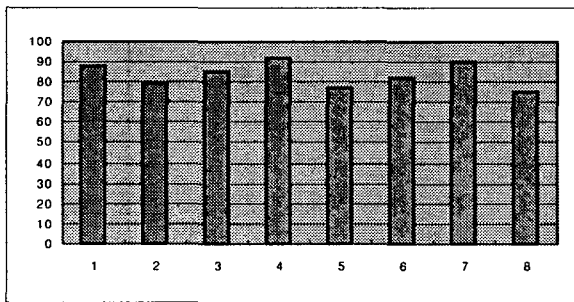


그림 4 검색만족도

그림 5는 키워드 중에서 고빈도 단어를 뽑아 객관적인 지표와의 상관관계를 알 수 있었다. 고빈도 단어로 인해 전체 만족도를 낮추는 결과를 줄 수 있다는 것을 알 수 있다.

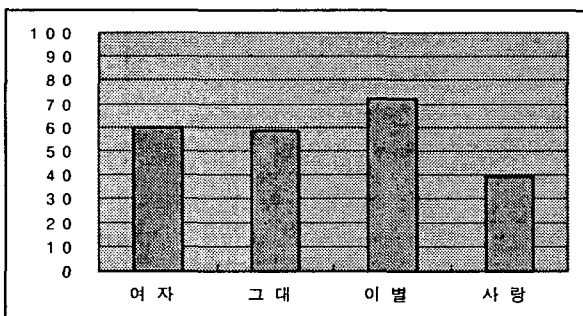


그림 5 고빈도 단어 만족도

그림 6은 곡명의 첫 음절의 키워드를 사용하지 않고 두 번째나 세 번째에 위치한 키워드를 사용한 검색만족도 결과로써, 그림 5과 그림 6을 분석한 결과 고빈도의 단어일수록 곡명의 위치에 상관없이 포함하는 곡명의 수가 많아져 검색만족도를 낮추는 결과를 얻는다는 것을 알 수 있다.

곡명	키워드	사실의 할기	인생은 나그네	하늘도 모르는 사람
키워드	사람	할기	인생	나그네
만족도(%)	0	98	90	98
키워드	마지막	약속	당신의 모습	흐린가을하늘에편지들써
만족도(%)	100	100	28	98
키워드	그녀의 연인에게	바람에 실려간 사람	고향은 내 사람	
만족도(%)	100	100	60	0
키워드	아름까지사랑하게	사나이	번 가슴	사랑 그리고 이별
만족도(%)	100	70	98	98
키워드	내일은 사랑	여자의 마음	당신의 눈물	
만족도(%)	98	0	28	28
키워드	내일	사랑	여지	마음
만족도(%)	98	0	28	28
키워드	바보	사랑	계절	사람의 이름표
만족도(%)	70	0	0	98

그림 6 키워드의 위치에 따른 만족도

4. 결론 및 향후과제

본 논문에서는 음성인식기술과 정보검색기술을 결합하여 기존 방식의 시간 및 사용상의 문제점을 개선한 시스템을 설계 하였다. 키보드 없이 음성인식으로 곡명을 찾도록 하고 음성인식기술의 한계를 보완할 수 있도록 정보검색기술을 사용하여 고빈도 키워드를 추출하여 사용할 수 있게 하였다. 또한 곡명의 인기 순위 변동을 빠르게 적용할 수 있는 지능형 노래 반주 시스템을 구현하였다.

향후 과제로써 그림 6의 객관적 만족도에서는 표 7, 표 8 높은 만족도를 보이는 것은 고빈도 키워드를 포함한 곡명의 검색이 객관적 만족도 검색에 많이 포함되지 않아 높은 만족도를 보였다. 높은 인식률을 가질 수 있도록 키워드간의 오인식을 줄이는 연구와 검색 만족도의 실험결과, 만족도가 낮은 '사랑', '이별' 등의 고빈도 키워드 등에 2차 질의검색을 사용하여 검색의 만족도를 높여야 할 것이다.

5. 참고문헌

- [1] 김영택 외 공저, "자연언어처리", pp69-98, 1994.
- [2] 임희석. 1997. "언어지식과 통계 정보를 이용한 한국어 품사 태깅 모델" 고려대학교 박사학위논문.
- [3] (주)세중 나모 인터랙티브, 음성인식 선블카드, <http://www.namo.co.kr/>
- [4] 예스테크놀로지, "음성인식 교통정보서비스", <http://www.roadi.com>
- [5] 양명석, 이석형, 강남규, 윤화목, "메타검색에서 링크정보와 요약정보를 이용한 검색결과 통합" 한국정보과학회 02 가을 학술발표논문집 (1), pp118-120, 2002