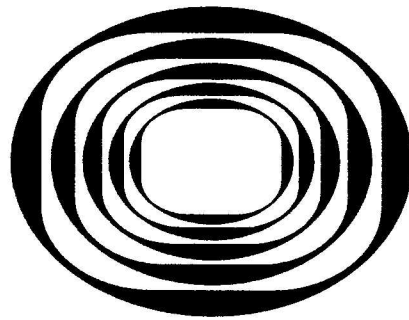


2003년도 한국방송공학회 정기총회 및 학술대회



- 일시 : 2003년 11월 15일 (토)
- 장소 : 한국정보통신대학원대학교 (화암캠퍼스)
- 주최 : 사단법인 한국방송공학회
- 후원 : 한국과학기술단체총연합회, 한국언론재단
KBS, MBC, SBS, EBS

Contents 2003. 11. 15 (Saturday)

세미나 III 실

session ① 검색기술

좌장 : 김해광 (세종대) -- 213

- ① MPEG-7 오디오 하위 서술자를 이용한 음악 검색 방법에 관한 연구 박만수, 박철의, 김희린, *강경옥 (ICU, *ETRI) -- 215
- ② 멀티모달 특징을 이용한 비디오 장르 분류 진성호, 배태면, 추진호, 노용만, *강경옥 (ICU, *ETRI) ----- 219
- ③ MPEG-7 Description을 이용한 객체 기반 MPEG-4 콘텐츠 검색 시스템 장예선, 김규현 (ETRI) ----- 223
- ④ 실시간 광고 검색 구영모, 김정림, 설상훈 (고려대) ----- 227

session ② 음향/음성 신호처리

좌장 : 홍진우 (ETRI) -- 231

- ① 차량환경에서 DMB용 음성명령어기 사용을 위한 음성개선 방법 백승권, 한민수, *남승현, **강경옥, (ICU, *배재대, **ETRI) ----- 233
- ② 가상환경에서의 4채널 오디오 출력을 이용한 실시간 입체음향 구현 양시영, 이돈응, 정제창 (한양대) ---- 237

session ③ MPEG-7/MPEG-21

좌장 : 김형중 (강원대) -- 241

- ① 방송·통신 융합 환경에서의 u-커머스 광고 아이템 연구 남재호, *문남미, 홍진우 (ETRI, *SIT) ----- 243
- ② 입체 디지털 아이템의 MPEG-21 DIA 테스트베드 손현식, 김만배, 송영주, 홍동희, 김학수, *홍진우 (강원대, *ETRI) --- 251
- ③ MPEG-21 DIA 기반 관심영역 추출을 위한 비디오 적응 엔진의 구현 박성준, 손유미, Hendry, 정현, Qonita, *김중남, *박근수, 김문철 (ICU, *KBS) ----- 257
- ④ 임베디드 시스템에서 MPEG-4 미디어 프리젠테이션 임태범, 이윤주 (전자부품연구원) ----- 261

session ④ 스트리밍/방송 소프트웨어

좌장 : 최진수 (ETRI) -- 265

- ① 동기화된 데이터 방송 애플리케이션 제작을 위한 스트림 생성기 개발 김세훈, 정문열 (서강대) ---- 267
- ② 분산환경을 위한 DirectShow의 확장 방법 강명현, 최성중 (서울시립대) ----- 271
- ③ 내용기반 검색을 위한 향상된 스케치 질의 방법 허명환, 정병희, 김희정, 임미영, 김형준, 김희율 (KBS, 한양대) ---- 275

차량환경에서 DMB용 음성명령어기 사용을 위한 음성개선방법

백승권*, 한민수*, 남승현**, 강경옥***,
* 한국정보통신대학교 공학부 멀티미디어 그룹
** 배재대학교 전자공학과
*** 한국전자통신연구원 무선방송 연구부
skbeack@icu.ac.kr

Speech Enhancement for DMB Voice commander in Car environment

Seung Kwon Beack, Minsoo Hahn, Seung Hyon Nam, Kyung Ook Kang
Multimedia Group, Information and Communications University
Electronic Engineering, PAICHAJ
Radio & Broadcasting Research Laboratory ETRI

요약

본 논문에서는 차량용 음성명령어기의 사용을 위한 전처리 과정으로 음성개선 방법을 다룬다. 특히 DMB 사용환경에서 보다 주위 소음에 자유롭고 단말 조작에 있어 안정성을 보장하기 위하여 일반적인 단일 마이크로 폰으로 처리되는 잡음뿐만 아니라 음성명령어를 제외한 오디오 신호 등 비정적 통계적 특성을 갖는 소음들도 제거 될 수 있도록 음성개선 방법을 제안한다. 우리는 2개의 마이크로폰을 가지고 BSS 알고리즘을 적용하여 비정적 신호들을 분리하고, 분리된 신호에 대하여 Kalman Filter를 이용하여 시간상 단구간 정적 잡음을 제거한다. 본 논문의 인식 실험 결과를 통하여 공간적, 시간적 음성개선 방법이 순차적으로 적용될 때, 실제 차량 환경에서 음성 개선 알고리즘으로 적용될 수 있음을 보였다.

1. 서론

향후 수년 내에 차량에서의 DMB(Digital Multimedia Broadcasting) 서비스가 가능할 것이다. 이때 주행중인 운전자의 안전과 편의를 보장하기 위하여 DMB단말 서비스 요청수단으로 음성명령어기 사용이 요구된다. 일반적으로 차량 내에서 신뢰도 높은 음성명령어기 사용을 위하여 음성개선 방법이 전처리 과정으로 요구된다. DMB용 음성명령어기에서 잡음은 주행으로부터 야기되는 잡음뿐만 아니라, 동승자나 라디오로부터 발생하는 음성명령어가 아닌 음성 및 오디오신호도 잡음으로 고려되어야 한다. 전자의 경우 시간상 음성과 상호 상관관계가 적은 관계로 단일 마이크로폰을 이용한 제거가 가능하나 후자는 음성명령어와 시간적 상관관계가 무시될 수 없으므로 단일 마이크로폰을 이용한 제거가 용이하지 않다. 우리는 2개의 마이크로폰을 이용한 제거 방법을 고려한다. 먼저 음성과 시간적 상관성이 큰 신호는 2개의 마이크로폰으로부터 공간적 독립성을 이용하여 분리한 후, 아차적으로 음성신호로 기대되는 채널에 단일마이크로폰 잡음제거 기술을 적용한다. 전자의 공간상 신호분리 방법으로는 Lucas parra의 BSS(Blind Source Separation)방법을 적용하고[1] 후자의 시간 축 상의 잡음제거방법은 Kalman Filter를 이용하였다.[2] 실험에 사용된 데이터는 실제환경에서 녹취한 차량 잡음과 음악 및 음성 신호를 인식 대상 음성신호와 인위적으로 혼합하여 열화시켰다. 잡음의 공간상 혼합은 실제환경을 반영하기 위해 복적분 혼합회로(convolutive mixture)로 모델을 사용하였다.

음성인식 성능을 측정하여 알고리즘의 신뢰도를 측정하려 한다. 본 논문의 구성은 시간적 잡음제거 기술과 공간적 잡음 제거 알고리즘에 대하여 간략히 설명하고 각각의 인식성능 및 제안된 전처리 과정에 의한 인식성능을 측정하고 이를 비교 분석한다.

2. 시간성 잡음 제거 알고리즘

우리는 시간성 잡음을 시간 샘플상에서 음성신호 $s(t)$ 와 상관관계가 적은 잡음신호 $v(t)$ 로 정의한다. 이는 음성신호 $s(t)$ 에 대한 자기상관함수를 $R(\tau) = E\{s(t)s(t+\tau)\}$ 라 할 때, 잡음에 의해 열화된 $\hat{s}(t)$ ($s'(t) = s(t) + n(t)$)의 자기 상관함수 $R_s(\tau)$ 는 $R(\tau) \cong R_s(\tau)$ 로써 표현 가능하다. 만일에 시간상에서 $v(t)$ 가 백색화(whitening) 특성을 지닌다면, 이론적으로 $R(\tau) = R_s(\tau)$ 이 된다. 그러므로 시간성 잡음은 백색화가 가능할 때 시간 축 상에서 성공적으로 제거가 가능하다. 이러한 백색화가 가능하기 위해서 시간성 잡음은 분석구간 내에서 정적인 통계적 특성(stationary statistic characteristic)을 지녀야 한다. 이러한 조건하에서 시간성 잡음은 단일 마이크로폰으로 입력된 열화음성으로부터 제거가 가능하다.

차량에서 순수하게 발생하는 잡음의 대부분이 단구간 내에서 정적인 통계적 특성을 갖는다. 이는 차량 자체내에서 발생하는 엔진 소음이나 진동, 에어컨 소음등과 유체의 흐름에 의한 소음-바람과 창문, 지면과 타이어-등이 대표적이라 할 수 있다. 물론 이것이 완벽한 백색화를 통하여 음성으로부터의 제거가 가능한 것은 아닐지라도, 음성보다 장구간에서 정적인 통계특성을 유지한다는 것로부터 훌륭하게 제거 될 수 있다. 우리는 차량내에서 성공적인 시간성 잡음 제거를 위하여 단구간(200msec) 잡음정보를 기반으로 백색화를 통한



그림 1. 제안된 2개의 마이크로폰을 이용한 음성개선 방법 [x: 열화된 음성 신호, s: 복원된 위 음성 신호]
우리는 각각의 알고리즘이 열화된 음성에 적용 전 후