

분산형 시스템을 적용한 음성합성에 관한 연구

A Study on Speech Synthesizer Using Distributed System

김진우*, 민소연**, 나덕수***, 배명진*

(Jin-Woo Kim*, So-Yeon Min**, Deok-Su Na***, Myung-Jin Bae*)

*숭실대학교 정보통신공학과, **서울대학교 정보통신과, ***(주)보이스웨어

(접수일자: 2009년 12월 9일; 수정일자: 2010년 2월 11일; 채택일자: 2010년 2월 17일)

최근 광대역 무선 통신망의 보급과 소형 저장매체의 대용량화로 인하여 이동형 단말기가 주목 받고 있다. 이로 인해 이동형 단말기에 문자정보를 청취할 수 있도록 문자를 음성으로 변환해 주는 TTS (Text-to-Speech) 기능이 추가되고 있다. 사용자의 요구사항은 고품질의 음성합성이지만 고품질의 음성합성은 많은 계산량이 필요하기 때문에 낮은 성능의 이동형 단말기에는 적합하지 않다.

본 논문에서 제안하는 분산형 음성합성기 (DTTS)는 고품질 음성합성이 가능한 코퍼스 기반 음성합성 시스템을 서버와 단말기로 나누어 구성한다. 서버 음성합성 시스템은 단말기에서 전송된 텍스트를 데이터베이스 검색 후 음성파형 연결정보를 생성하여 단말기로 전송하고, 단말기 음성합성 시스템은 서버 음성합성 시스템에서 생성된 음성파형 연결정보와 단말기에 존재하는 데이터베이스를 이용하여 간단한 연산으로 고품질 합성음을 생성할 수 있는 시스템이다. 제안하는 분산형 합성기는 단말기에서의 계산량을 줄여 지가의 CPU 사용, 전력소모의 감소, 효율적인 유지보수를 할 수 있도록 하는 장점이 있다.

핵심용어: 음성합성, 분산형 음성합성기, 코퍼스 기반 TTS, 분산형 시스템

투고분야: 음성처리 분야 (2.4)

Recently portable terminal is received attention by wireless networks and mass capacity ROM. In this result, TTS(Text to Speech) system is inserted to portable terminal, Nevertheless high quality synthesis is difficult in portable terminal, users need high quality synthesis.

In this paper, we proposed Distributed TTS (DTTS) that was composed of server and terminal. The DTTS on corpus based speech synthesis can be high quality synthesis. Synthesis system in server that generate optimized speech concatenation information after database search and transmit terminal. Synthesis system in terminal make high quality speech synthesis as low computation using transmitted speech concatenation information from server. The proposed method that can be reducing complexity, smaller power consumption and efficient maintenance.

Keywords: Speech synthesis, DTTS, corpus based TTS, distributed system

ASK subject classification: Speech Signal Processing (2.4)

I. 서론

음성합성 기술은 입력된 문장을 언어학적 규칙을 통해 분석하여 음운 정보와 운율 정보를 추출하고 이것을 이용하여 합성음을 생성하는 기술로 TTS (Text-to-Speech) 시스템이 대표적이다. TTS 시스템은 사용자들의 요구로 고품질의 합성음을 얻기 위해 계속해서 연구되고 있다.

고품질의 합성음을 얻기 위한 대표적인 시스템은 대용량 음성 DB (data-base)를 사용하는 코퍼스 (corpus) 기반 음성합성이다. DB에서 합성에 필요한 최적의 합성단위를 선택하고 이것들을 적절히 연결하여 합성음을 생성하는 합성단위 선택 기반 연결형 합성 기술 (unit-selection based concatenation synthesis)이 사용되고 있다. 이 기술의 가장 큰 장점은 기존의 규칙합성 (rule-based synthesis) 시스템이 가지고 있는 제한적인 운율 변화에 의한 자연성 감소의 단점을 극복한 것이다. 대용량 음성 코퍼스의 구축을 통해 다양한 운율변화를 구현할 수 있게

됨으로써 사람의 목소리와 비슷한 음질의 합성음을 생성할 수 있게 된 것이다.

최근 무선 통신망의 보급 및 소형 저장매체의 대용량화로 인하여 네비게이션, 전자사전, PDA, MP3, 휴대폰 등의 이동형 단말기들이 주목받고 있다. 이러한 기기들은 멀티미디어의 기본인 문자정보를 청취할 수 있도록 텍스트를 음성으로 변환해 주는 음성합성 기능이 추가되고 있다 [1-3].

이동형 단말기에서도 사용자들은 고음질의 음성합성을 요구하고 있다. 고음질의 음성합성을 수행하려면 많은 연산량이 필요하고, 멀티미디어의 발전으로 텍스트를 음성으로 변환하는 작업 이외의 여러 가지 작업을 수행하기 때문에 합성 시 소비되는 계산량을 감소시킬 필요성이 있다.

본 논문에서는 이동형 단말기에 적합한 고음질의 음성합성을 수행하기 위해서 분산형 시스템을 음성합성에 적용시켜 적은 계산량으로 고음질의 음성합성을 할 수 있도록 하는 분산형 음성합성 시스템을 제안한다. 논문의 구성은 서론에 이어 2장에서는 코퍼스 기반 음성합성 시스템, 3장에서는 분산형 음성합성 시스템을 알아보고, 4장에서는 본 논문에서 제안한 분산형 음성합성기에 대해 설명하고 5장에서는 제안한 방법에 대한 실험 및 결과를 설명하고 마지막으로 6장에서 결론을 맺는다.

II. 코퍼스 기반 음성합성 시스템

코퍼스 기반 음성합성은 합성에 사용하는 음소나 음절과 같은 합성단위에 대응하는 스펙트럼 정보를 음성 코퍼스(데이터와 그 생성, 속성 등을 포함한 총체)로부터 얻는 방법이다. 코퍼스 기반 TTS 시스템의 일반적인 구성도는 그림 1에서 보는 바와 같다.

합성 시스템에서는 입력 문장이 들어오면, 언어학적 처리부의 전처리 과정을 통해 약어나 숫자, 알파벳 표기, 그리고 특수기호 등을 텍스트 표기 형태로 변환시킨 다음, 문장 분석 과정에서 문장의 구조를 분석하게 되며, 이때 형태소 및 품사와 악센트 정보들을 사용한다. 발음 표기변환 부분에서는 음운규칙과 예외사전 등을 이용하여 사람이 문서를 읽어주는 형태의 음성학적 표현을 만들어 준다. 그 후 절, 구, 휴지기간을 결정하고, 음소의 지속시간 및 고저와 같은 억양특성 생성시킨다. 앞서 추출한 음운 정보 및 발음 정보들을 이용하여 데이터베이스의 합성 단위에 해당하는 음성파형을 선택하고 연결하여 합

성음을 만들어낸다 [4-7].

이와 같은 구조의 일반적인 코퍼스 기반 음성 합성 시스템은 동적 프로그래밍 탐색 기법인 Viterbi 알고리즘을 사용하여 최적의 합성단위를 선택한다. 대용량 DB에는 다양한 운율 및 주파수 특성을 갖는 합성 단위들이 많이 존재하므로, 한정된 합성 단위들의 특성을 직접 제어하여 합성하던 기존의 방식들 보다 훨씬 자연스러운 합성음을 얻을 수 있는 장점이 있다 [8]. 반면 방대한 양의 음성 DB가 사용됨에 따라 합성단위 선택 시 소요되는 계산량이 크게 증가됨과 동시에 메모리 요구량이 많아지며, 발생하는 성우에 지나치게 의존적이어서 다양한 음색이나 발성 스타일을 만들기 위해서는 추가적인 DB구축이 요구된다는 단점이 있다 [9,10].

III. 분산형 음성합성 시스템

분산형 음성합성기 (Distributed Text-to-Speech, DTTS)는 음성합성 시스템을 서버와 단말기로 나누어 구성하는 것을 말한다. 서버와 단말기를 사용함으로써 인해 양방향 통신이 가능하고 서버에서 전송하려는 정보를 단말기에서 수신할 수 있고 단말기에서 요구하는 정보를 서버를 통하여 제공받을 수 있다.

분산형 음성합성 방식은 기본적으로 2가지 방식이 있으며, 그림 2과 그림 3은 분산형 음성합성 방식을 나타낸 것이다. 이 중 그림 2는 서버에서 제공하고자 하는 텍스트를 단말기로 전송하여 단말기에서는 수신된 텍스트 음성합성을 수행한다. 이 방식은 채널 정보량은 적지만 음성합성의 모든 과정을 단말기에서 수행해야 하므로 음성

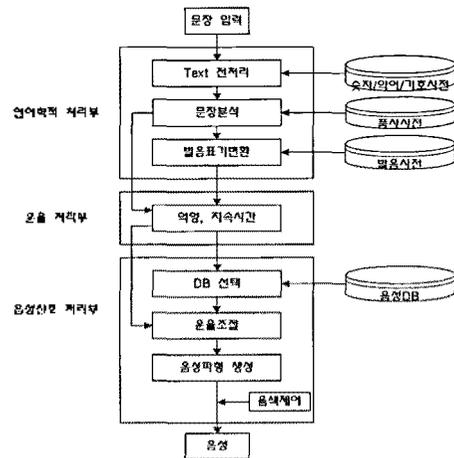


그림 1. TTS 시스템의 블록도
Fig. 1. Block diagram of TTS system.

합성 시 많은 연산량을 필요로 하기 때문에 단말기 사양이 높아야 한다. 예를 들어 차량환경에서 새로운 뉴스나 교통정보 등을 음성으로 서비스 받고자 하는 경우 서버에서 뉴스나 교통정보를 텍스트로 전송하고 단말기에서 음성합성기를 이용하여 합성하는 방법이다. 그러나 이것은 단말기의 성능에 따라 합성속도의 문제와 합성음질 열하 등을 야기할 수 있다.

그림 3은 서버에서 모든 텍스트의 음성합성을 수행하여 합성된 음성데이터를 전송하고 단말기에서 합성음을 단순 재생하는 방식이다. 이 방식은 음성합성을 서버에서 수행하기 때문에 단말기에서의 연산량이 크기 않기 때문에 사양은 높지 않아도 되지만 음성파형을 직접 전송하기 때문에 채널 정보량이 커진다는 단점이 있다. 예를 들어 차량환경에서 서버에서 뉴스나 교통정보를 전송할 때 서버에서 직접 합성음을 생성하여 전송하고 단말기에서는 재생만 하는 방법이다. 이것은 합성음의 많은 데이터양으로 인해 통신의 문제를 발생시킬 수 있다 [11].

이에 본 논문에서는 적은 데이터량의 전송과 낮은 단말기 사양을 유지하며 고음질의 음성합성을 수행하기 위하여 새로운 분산형 음성합성 시스템을 제안한다.

IV. 제안하는 방법

본 논문에서 제안하는 방식은 기존의 코퍼스 기반 음성합성기를 이동형 단말기에 적합하도록 단말기에서 모든 연산이 처리되던 시스템을 서버와 단말기 음성합성 시스템으로 나누어 단말기에서의 연산을 분산시켜 연산량을

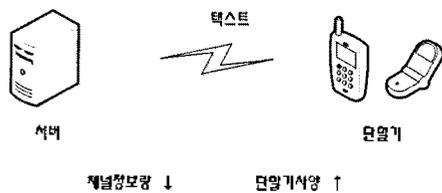


그림 2. 텍스트 제공 방식 분산형 합성기
Fig. 2. Distributed TTS using text trasmission.

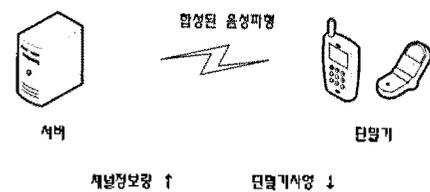


그림 3. 합성된 음성파형 제공 방식 분산형 합성기
Fig. 3. Distributed TTS using waveform trasmission.

감소시킨 분산형 음성합성기를 제안한다.

이동형 단말기와 서버간의 통신은 무선 통신 기술 WiBro 혹은 Wi-Fi를 이용한다. 현재 MIMO 기술을 이용한 WiBro는 전송 속도를 최대 하향 34.6 Mbps, 상향 8 Mbps까지 구현이 가능하다 [12].

코퍼스 기반 음성 합성기를 서버와 단말기에 분산시키기 위하여 그림 4에 음성 합성기 각 모듈의 부하를 나타내었다. 사용된 합성기는 (주)보이스웨어의 코퍼스 기반 음성 합성기로 각 모듈의 부하 분포를 알아보기 위해 해당 함수에서 처리되는 시간을 측정하였다. 가장 많은 부하가 걸리는 모듈은 데이터베이스를 검색하여 최적의 음성파형을 선택하는 DB검색 모듈로써 시스템 전체 부하의 63%, 약 2/3를 차지하고 있고, 실제 음성 파형을 DB에서 가져와 합성음을 만드는 모듈인 음성파형 생성 모듈은 22.5%, 약 1/5 정도만 차지하고 있다. 그리고 문장분석 및 발음변환은 11.2%, Text 전처리 2.5% 억양지속시간이 0.6%의 부하율을 차지하고 있다.

그림 5는 제안하는 분산형 음성합성기의 구성도를 나타낸 것이고, 그림 6은 서버-단말기의 블록도를 나타낸 그림이다.

서버와 단말기에 동일한 데이터 베이스를 제공한다. 서버는 내장형 저장매체에 저장하고, 단말기는 대용량화된 외장형 저장매체에 저장한다. 음성합성기의 분산을 위해 계산량이 많은 모듈인 DB검색을 포함하여 문장분석, 발음변환, Text 전처리, 억양지속시간의 모듈은 서버

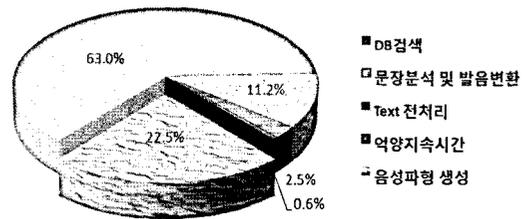


그림 4 합성기 각 모듈의 부하 분포
Fig. 4. distributed load of corpus based speech synthesizer.

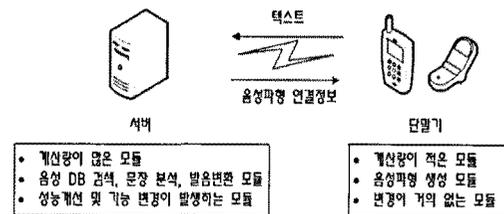


그림 5. 제안하는 분산형 음성합성기의 구성도
Fig. 5. Composition of proposed method.

에 포함시키고, 단말기에는 계산량이 적은 음성파형 생성 모듈을 포함시킨다.

서버 음성합성 시스템 부는 입력된 텍스트에 대하여 Text 전처리, 문장 분석, 발음표기 변환, DB 선택 등의 과정을 거쳐 최적의 음성파형의 정보를 선택하고 단말기로 전송한다. 선택된 최적의 음성파형의 정보를 음성파형 연결정보라 하겠다.

단말기 음성합성 시스템 부는 서버에서 생성된 음성파형 연결정보를 수신 받아 이의 인덱스 정보를 이용하여 데이터 베이스로부터 음성파형을 선택하고 합성음을 생성한다. 음성파형 생성 모듈은 전체 음성합성 시스템 중 약 22 %만 차지하는 음성파형 생성모듈만을 포함하여 기존의 단말기보다 계산량을 감소시킨다. 그림 7은 분산형 음성합성기의 서버 음성합성 시스템을 그림 8은 단말기 음성합성 시스템을 나타낸 블록도이다.

서버에서 생성된 음성파형 연결정보를 단말기로 재송

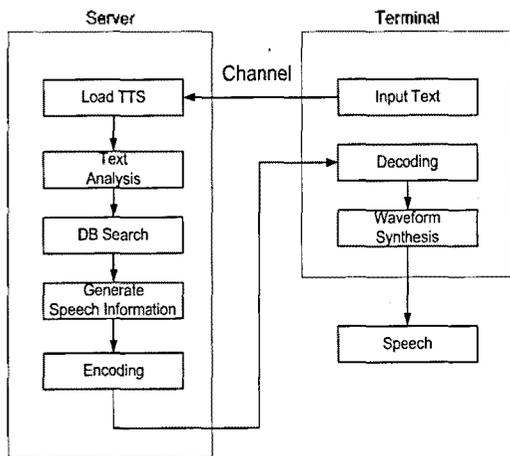


그림 6. 제안하는 분산형 음성합성기의 블록도
Fig. 6. Block diagram of proposed method.

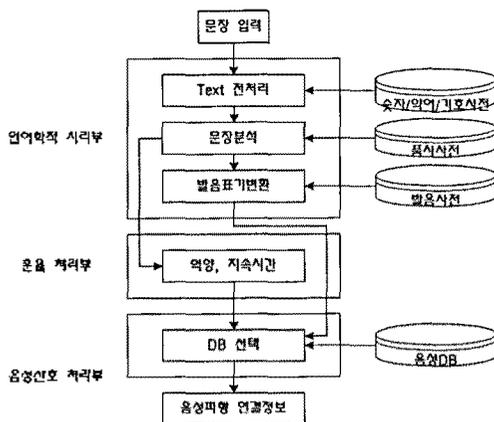


그림 7. 서버 음성합성 시스템 블록도
Fig. 7. Block diagram of server synthesizer.

신하여 단말기 음성합성 시스템에서는 수신된 음성파형 연결정보를 이용하여 합성하면 단말기에서 직접 합성하는 경우에 비해 적은 연산량으로 합성이 가능하고 음성파형 연결정보도 많은 데이터량을 필요로 하지 않으므로 텍스트나 합성음을 직접 전송하는 경우보다 매우 효율적이다.

서버 음성합성 시스템 부에서 데이터베이스 검색을 하여 음성파형을 선택 후 출력되는 음성파형 연결정보는 다음과 같다.

- 합성될 음성파형의 위치정보 (Position)
- 합성될 음성파형의 길이 (Length)
- 묵음의 종류 (Type of Pause)

입력텍스트가 '안녕하세요.' 라고 한다면 합성될 음성파형은 각각 [sil + 아 + ㄴ + ㄴ + ㅣ + ㅇ + ㅎ + ㅏ + ㅓ + ㅕ + ㅛ, + sil]에 해당하는 합성단위가 될 것이고, 그림 7과 같이 DB에서 검색 과정을 통해 최적의 음성파형을 선택하고, 그 음성파형은 각각 시작 위치 정보와 길이 정보로 나타낼 수 있다. 여기서 sil은 문장의 시작과 끝을 나타낸다. 입력된 텍스트가 문장인 경우 합성음을 생성하기 위해서는 끊어 읽기 정보가 필요하게 되는데, 그 정보를 묵음의 종류로 출력한다. 묵음의 종류는 그림 9에서 나타낸 것과 같이 3가지 종류로 구분하는데 0은 묵음이 존재하지 않는 것, 즉 뒤의 음과 바로 연결하는 것을 나타내고, 1과 2는 쉼표, 마침표 등과 같은 끊어 읽기 기호에 대응되는 것으로, 각 합성단위의 뒤에 쉼표, 마침표 등과 같은 끊어 읽기가 온다면 그것에 해당하는 묵음의 종류를 출력한다. 위와 같은 3가지의 정보를 음성파형 연결정보로 하여 단말기 음성합성 시스템 부에 전송하거나 저장하게 된다.

위의 3가지 필수 정보 이외에도 연결정보에 피치, 지속 시간, 하이라이트, 속도 등을 포함시켜 추후 음성파형 연

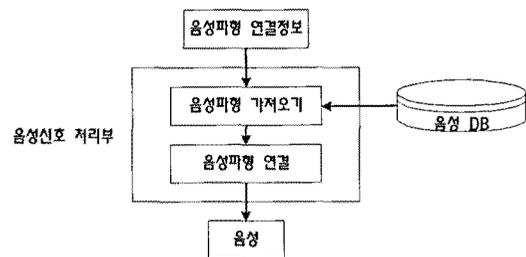


그림 8. 단말기 음성합성 시스템 블록도
Fig. 8. Block diagram of terminal synthesizer.

결정보만 가지고도 음성의 여러 가지 정보를 알 수 있다.

V. 실험 및 결과

5.1. 실험 및 데이터 환경

본 논문에서 제안한 분산형 음성합성기의 성능평가를 위해서 데스크탑 PC (Intel Pentium processor 1.5 GHz)에서 시뮬레이션을 구현하였다. 사용된 음성합성기는 보이스웨어사의 코퍼스 기반 음성합성 시스템을 사용하였다. 코퍼스의 용량은 947 MByte이고 합성음은 16 kHz로 샘플링 되고 16 bits로 양자화된 여성의 음성이다. 코퍼스 기반 음성합성 시스템을 서버와 단말기의 분산형 음성합성 시스템으로 분리 하여 음성파형 연결정보가 생성되도록 하였다.

서버 음성합성 시스템 부에서 출력되는 음성파형 연결 정보에는 음성합성을 위한 필수 요소인 음성파형의 위치 정보, 음성파형의 길이, 녹음의 종류와 선택요소인 음소열, 음성파형의 시작 피치, 음성파형의 마지막 피치, 피

치마크위치 (한 피치 주기에서 에너지가 최대인 위치), 피치마크 길이의 정보가 포함되어 있다.

실험을 위한 텍스트는 한글 50 kbyte (약 2만 5천 음절에 해당)의 중복되지 않는 문장을 선택하여서 음성합성시의 음소차이를 평준화 하였다.

5.2. 시뮬레이션 구현

시뮬레이션 구현을 위해서 TCP/IP를 통해 네트워크가 구성되어 있는 PC 2대를 사용하여 DTTS 시뮬레이션을 구현하였다. 아래 그림 10은 분산형 음성합성기를 구현하기 위한 개요도를 나타냈다.

PC는 서버와 클라이언트로 나누어서 VC++ MFC를 이용하여 프로그램 하였다. 그림 11는 클라이언트 PC에서의 화면으로 입력창에 텍스트를 입력하여 음성합성 버튼을 클릭하면 서버 PC에 텍스트를 전송하고 서버에서 생성된 음성파형 연결정보를 전송 받아서 음성을 생성한다. 그림 12은 서버 PC에서의 화면을 제시한 것이다. 서버 PC에서는 클라이언트 PC에서 제공된 텍스트를 분석

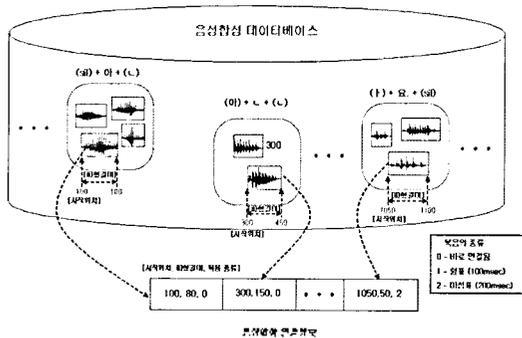


그림 9. 음성파형 연결정보
Fig. 9. concatenation information of speech waveform.

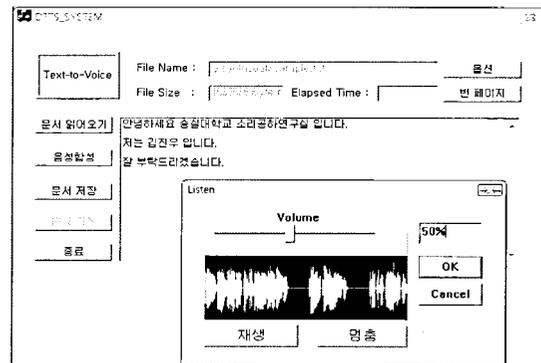


그림 11. 단말기 TTS 시스템
Fig. 11. TTS system of terminal.

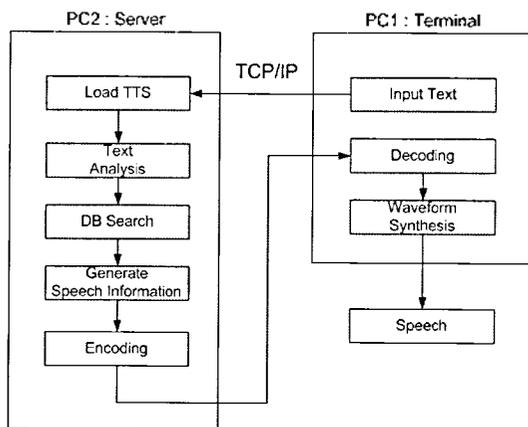


그림 10. 시스템 구현 개요도
Fig. 10. diagram of speech synthesizer system.

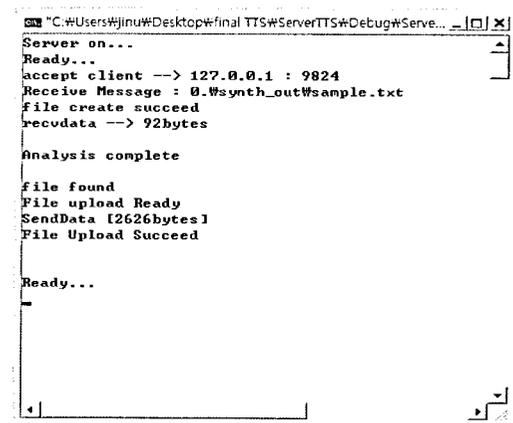


그림 12. 서버 TTS 시스템
Fig. 12. TTS system of server.

표 1. 음성파형 연결정보량과 합성기별 소요시간
Table 1. consumption time and capacity of speech information.

Text	음성파형 크기	음성파형 연결정보 크기	단말기 부 처리시간	
			일체형	분산형
50 kbyte	117 Mbyte	1.3 Mbyte	35.3 sec	6.6 sec

하여 DB검색 후 음성파형 연결정보를 생성하여 바이너리 코드로 인코딩하여 클라이언트 PC로 다시 전송한다.

5.3. 단말기에서의 합성속도

일체형 합성기가 50 kbyte의 음성을 합성하는데 소요 되는 시간과 분산형 합성기의 단말기가 텍스트 50 kbyte의 음성파형 연결정보의 크기와 합성 시 소요되는 시간을 비교하여 단말기에서의 연산량과 전송 용량의 차이를 확인하였다.

실험 결과 음성파형 대신 연결정보를 전송함으로 채널 용량이 약 10배 정도 줄어들고, 일체형 합성기에서의 소요시간 보다 분산형 합성기의 단말기 소요시간이 약 6배 정도 감소된 것을 확인 할 수 있다. 이를 통해 일체형 합성기에서의 단말기 부하가 분산형 합성기에서의 단말기 부하가 더 적다는 사실을 알 수 있다. 서버와 단말기의 성능 차이가 클수록 처리시간은 감소할 수 있다.

VI. 결론

본 논문에서는 대용량화된 저장매체와 광대역 통신망이 보급된 점을 감안하여 이동형 단말기에서도 고음질의 음성합성을 수행 할 수 있도록 분산형 음성합성 시스템을 단말기에 적용하여 단말기에서의 계산량을 감소시켜 이동형 단말기에서도 고음질의 음성합성을 구현하기 위한 방법을 제안하였다.

제안한 방법의 성능 확인을 위하여 코퍼스 기반 음성합성 시스템을 서버와 단말기로 나누어 서버에서 생성된 음성파형 연결정보의 정보량을 확인하고, 일체형 음성합성기와 분산형 음성합성기 중 단말기에서의 음성합성시간을 비교하였다. 50 kbyte의 텍스트를 음성파형 연결정보로 표현하기 위해 약 1.3 Mbyte의 정보량이 필요하였으며, 분산형 합성기에서 계산속도가 약 6배 정도 빨라짐을 알 수 있었다.

분산형 음성합성기의 제안으로 갖는 장점으로는 첫째, 고정된 텍스트 환경에서 효율적으로 이용이 가능하다. 음성파형 연결정보의 생성으로 인해 소설, 기사와 같은 입력텍스트가 바뀌지 않는 환경에서 적은 용량의 음성파

형 연결정보를 생성해 두면 재생 시 간단한 연산만으로 합성음을 생성할 수 있다. 둘째, 단말기 가격의 저렴화이다. 기존 단말기에서 모두 처리되던 데이터베이스 검색, 전처리 과정 등의 연산이 서버에서 이루어짐에 따라서 단말기는 낮은 성능의 음성합성 칩의 사용이 가능해졌다. 셋째, 시스템의 유지보수 및 업그레이드의 용이성이다. 서버에서 음성합성 연산 후 단말기로 전송을 하기 때문에 TTS 시스템의 기능을 업그레이드 하거나 유지보수 작업을 할 때 서버의 개선만으로 가능하기 때문에 기존 보다 용이하다. 마지막으로 전력소모의 감소이다. 이동형 단말기는 휴대성뿐만 아니라 배터리의 사용시간도 매우 중요한 요소로써, 음성합성 뿐만 아니라 다른 여러 가지 작업을 수행하기 때문에 단말기에서 계산량의 감소로 인해 전력소모도 감소된다.

향후 서버에서 생성되는 음성파형 연결정보의 최적화와 무선 통신망 (Wi-Fi, Mobile Wi-Max)을 이용한 실제 환경에서 분산형 음성합성기와 일체형 음성합성기의 처리속도 비교와 성능 향상을 위한 연구가 이루어져야 할 것이다.

참고 문헌

1. 나덕수, 이종석, 배명진, "이동형 단말기 및 음성합성 칩에 적합한 저전력 소모 및 저 계산량 고음질 음성합성 시스템," *한국음향학회 2007년도 정기총회 및 추계학술발표대회 논문집*, 26권, 2(s)호, 53-56쪽, 2007.
2. 김영길, 박창현, 양운기, "데이터베이스 분산을 통한 소용량 문자-음성 합성 단말기 구현," *대한전자공학회 2003년도 하계종합학술대회 논문집*, 26권, 1호, 2431-2434쪽, 2003.
3. T. V. Raman, *Auditory User Interfaces Toward the speaking computer*, Kluwer academic publishers, 1997.
4. Thierry dutoit, *An introduction to text-to-speech synthesis*, Kluwer academic publishers, 1997.
5. 나덕수, "코퍼스 기반 일본어 음성 합성기의 성능 향상에 관한 연구", 박사학위 논문, 송실대학교, 2009.
6. 박진수, 박종세, 김형순, "코퍼스 기반 TTS 시스템의 개발을 위한 객관적 척도의 적용," *정보통신산업진흥원 정보통신연구진흥원 학술기사*, vol. 19, no. 1, 2003.
7. A. Hunt and A. Black, "Unit selection in a concatenative speech synthesis system using a large speech database," *in Proc of ICASSP, Atlanta*, vol. 1, pp.373-375, 1996.
8. Akemi Iida, Nick Campbell, Fumito Higuchi and Michiaki Yasumura, "A corpus-based speech synthesis system with emotion," *Speech Communication*, vol. 40, Issues 1-2, pp.161-187, 2003.
9. T. Mizutani and T. Kagosima, "Concatenative Speech Synthesis Based on the Plural Unit Selection and Fusion Method," *IEICE Trans. Inf. & Syst.*, Vol.E88-D, No.11, pp.2565-2572, 2005.

10. H. Kawai and T. Toda, "An Evaluation of Automatic Phone Segmentation for Concatenative Speech Synthesis," *Proc. IEEE ICASSP2004*, 2004.
11. 김진우, 민소연, 나덕수, 배명진, "분산형 시스템을 이용한 음성 합성의 계산량 감소에 관한 연구," *한국음향학회 2009년 추계 학술발표대회 논문집*, 28권, 2(s)호, 358-359쪽, 2009.
12. 최두현, 문연태, 김도균, "Wave 2 규격을 위한 와이브로 기지국용 일체형 복합 RF unit 연구," *대한전기학회지*, 59권, 9호, 1660~1668쪽, 2008.

저자 약력

• 김진우 (Jin-Woo Kim)



2008년 2월 : 한경대학교 응용수학과 (이학사)
 2008년 3월~현재: 송실대학교 정보통신공학과 석사과정

• 민소연 (So-Yeon Min)



2003년 2월 : 송실대학교 전자공학과 (공학박사)
 현재 : 서울대학 정보통신과 교수
 한국음향학회지 제21권 제3호 참조

• 나덕수 (Deok-Su Na)

2009년 2월: 송실대학교 정보통신공학과 (공학박사)
 현재: (주)보이스웨어 연구원
 한국음향학회지 제19권 제2호 참조

• 배명진 (Myung-Jin Bae)



현재: 송실대학교 정보통신전자공학부 교수
 한국음향학회지 제 26권 제4호 참조