

자연음 TTS(Text-To-Speech) 엔진 구현

조정호 · 김태은 · 임재환
남서울대학교 멀티미디어학과

요 약

TTS(Text-To-Speech) 시스템은 텍스트 문장을 자연스러운 음성으로 출력하는 시스템이다. 자연스러운 음성을 출력하기 위해서 언어에 대한 전문적 지식을 비롯하여 많은 시간과 노력이 요구된다. 또한 영어의 음운 변환은 음소에 따라 형태소에 따라 의미에 따라 다양한 변환을 가진다. 이를 일괄적으로 처리하기란 매우 힘든 일이다. 이러한 문제들을 해결하기 위하여 모음과 자음의 변화의 규칙을 적용한 시스템을 구현한다.

이 시스템은 문장의 분석을 통해 분류하고 음소 규칙 데이터를 통해 자연스러운 음성을 출력하게 되는 이전 과정을 통해 특수문자나 숫자 등을 정규화하여 처리한다. 이렇게 처리된 문자 데이터를 운율규칙을 통해 최종 출력한다. 그 결과, 40개의 음소 규칙 데이터를 통해 보다 정확한 음성을 출력할 수 있었으며, 시스템의 효율성도 높였다. 본 논문에서 제시한 시스템은 각종 통신장비와 자동화기기에 적용하여 다양한 분야에 활용될 수 있을 것이다.

Implementation of TTS Engine for Natural Voice

Jung-Ho Cho · Tae-Eun Kim · Jae-Hwan Lim

ABSTRACT

A TTS(Text-To-Speech) System is a computer-based system that should be able to read any text aloud. To output a natural voice, we need a general knowledge of language, a lot of time, and effort. Furthermore, the sound pattern of english has a variable pattern, which consists of phonemic and morphological analysis. It is very difficult to maintain consistency of pattern. To handle these problems, we present a system based on phonemic analysis for vowel and consonant.

By analyzing phonological variations frequently found in spoken english, we have derived about phonemic contexts that would trigger the multilevel application of the corresponding phonological process, which consists of phonemic and allophonic rules. In conclusion, we have a rule data which consists of phoneme, and a engine which economize in system. The proposed system can use not only communication system, but also utilize office automation and so on.

1. 서 론

최근 인터넷의 발달로 문서형태의 정보는 홍수를 이루고 있으며, 이를 음성으로 변환하고자 하는 수요는 급격히 늘어나고 있다. 최근에는 멀티미디어·통신 등 여러 분야에서 다양한 문자정보를 음성으로 서비스하기 위한 음성합성 및 음성압축 기술을 확보하기 위한 연구가 활발히 진행되고 있으며, CTI(Computer Telephony Integration) 또는 UMS(Unified Messaging System)사업이 활기를 띠고 있다. MPEG-4(Moving Picture Export Group-4)에서 음성합성의 인터페이스에 관한 표준화 작업이 진행되고 있으며 정보수집을 보다 편리하고, 신속하게 하기 위해서 인간과 기계의 정보교환을 위한 MMI(Man-Machine Interface)가 중요하게 대두되고 있다. TTS(Text-To Speech)기술은 그 나라의 언어현상에 대한 지식이 얼마나 있는지에 따라 그 성능이 결정된다. 일반적으로 이러한 경험적 지식을 확보한 상태에서도 무제한 영역의 텍스트를 처리하는 TTS 시스템이 실생활에 서비스가 가능하기까지는 오랜 시간이 걸렸으며, 음성학, 언어학, 심리학, 전산학, 공학, 물리학, 의학 등의 여러 분야의 연구 노력의 종합적인 결과이라고 말할 수 있다.

TTS 시스템을 개발하는데 있어서 중요하게 고려해야 할 사항이 명료성(Intelligibility)과 자연성(Naturalness)이다. 그러나 명료성과 자연성은 서로 상반되는 경향을 보이고 있다. 즉, 명료성이 향상되면 자연성이 떨어지고, 자연성이 향상되면 명료성이 떨어지게 된다. 그러므로 이러한 두 성질을 최대한 절충해서 사람과 같이 자연스럽게 명료한 합성음을 갖는 대화체 음성 합성 기술의 개발을 그 목적으로 한다.

2. 영어의 음운 변환 규칙

영어는 21개의 모음과 13개의 이중모음, 자음에는 8개의 폐쇄음, 10개의 마찰음, 3개의 비음, 2개의 유음, 3개의 반모음으로 구성되어있다. 각각의 철자는 특정 문자열을 나타낸다. 그러나 실제 발음 시 철자에 나타난 문자열은 그대로 발음되지 않고, 음운 변화 과정을 통해서 여러 가지 다양한 발음열을 생성한다. 영어의 토박이 화자라면 고도의 신뢰성을 가지고 어떤 단어가 몇 음절로 이루어졌는지(cat는 1음절 /kæt/, catfish는 2음절 /kæt-fɪʃ/, catalogue는 3음절 /kætə-lag/, catatonic은 4음절 /kæt-tə-tá-nik/로 이루어져 있다)를 결정할 수 있지만, 정확히 음절이 무엇인지에 대해서는 의견의 일치가 별로 없다. 이는 음절의 구조적 특징 때문이다. 일반적인 구조로 CV(C)라는 구조를 갖는데, 즉 이 구조는 단 하나의 자음 C뒤에 단 하나의 모음 V가 뒤따르고, 그 다음으로는 그 뒤에 단 하나의 자음(C)가 수의적으로 뒤따른다. 그리고 모음은 대체로 음절의 중심 또는 핵심을 형성하고, 그래서 종종 음절의 핵이라고 불린다[1].

이처럼 실제로 발음 되는 각 음절에 따른 구조적인 차이로 인해 음운의 변이가 달리 이루어진다. 각 모음과 자음에 따른 음절구조를 분석하고 꼭 수행하여야만 올바른 발음을 얻을 수 있다고 말할 수 있다. 하나의 음소는 음성 환경에 따라서 다른 음가의 변이음들로 실현된다. 둘 이상의 음소가 이어 나올 때 한 음소가 다른 음소로 바뀌기도 하고, 두 음소가 하나로 줄어들기도 하며, 새로운 음소가 첨가되기도 한다. 하나의 음소가 여러 변이음으로 실현되는 양상을 규칙화한 것을 변이음 규칙이라 하고, 한의 음소가 다른 음소

	p	b	m	t	d	n	k	g	ŋ	f	v	s	z	θ	ð	ʃ	ʒ	č	j	l	r	w	y	h
음절성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
자음성	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
공명성	-	-	+	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
유성성	-	+	+	-	+	+	-	+	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	-
지속성	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+
비음성	-	-	+	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
조찰성	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	+	+	+	+	-	-	-	-	-
설측성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
분포성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-
피찰성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-
순음성	+	+	+	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	+	-
원순성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-
설정성	-	-	-	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-
전방설	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	-	-	-	-	-	+	+	-	-
고설성	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	+	+	+	+	-	-	+	+	-
후설성	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
저설성	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+

(표 1) 영어자음의 변별자질 구성[2]

	i	ɪ	e	ɛ	æ	u	ʊ	ʌ	o	ɔ	ɑ	ə*	ɹ
음절성	(iy)		(ey)			(uw)			(ow)				
음절성	+	+	+	+	+	+	+	+	+	+	+	+	+
고설성	+	+	-	-	-	+	+	-	-	-	-	-	+
후설성	-	-	-	-	-	+	+	+	+	+	+	+	+
저설성	-	-	-	-	+	-	-	-	-	+	+	-	-
원순성	-	-	-	-	-	+	+	-	+	+	-	-	-
긴장성 (장음성)	+	-	+	-	-	+	-	-	+	-	-	-	-

(표 2) 영어 모음 변질자질 구성[2]

로 바뀌거나, 탈락하거나, 첨가되는 양상을 규칙화한 것을 음소 변동 규칙이라 한다. 그리고 변이음 규칙과 음소 변동 규칙을 통 털어 음운 규칙이라 한다[2]. 예를 들어, t와 d는 보통 /č/와 /j/와 같은 파찰음을 만들면서 구개음화 된다. 이와 같은 구개음화과정은 대체로 /i/와 /y/ 같은 고 전설음의 환경에서 발생한다. 예를 들면 dontcha /downčə/ 처럼 don't+you의 영어의 일상적 발화 축약에서는 don't의 마지막 /t/가 you의 활음 /y/와 결합될 때는 /č/가 된다. 이 동화 규칙은 하나의 음소가

다른 음소로 바뀌므로 음소 변동 규칙이다[2].

본 시스템은 영어의 대표적인 음소 변동 규칙 [표 1]과 [표2]에 나타난 자질의 존재를 나타내는 +값과 부재를 나타내는 -값을 통해 40개의 규칙[표3]을 추출하여 이를 채택하여 음소 문맥에 적용한다.

3. TTS 시스템의 구조

TTS(Text-to-Speech) 시스템은 [그림1]처럼

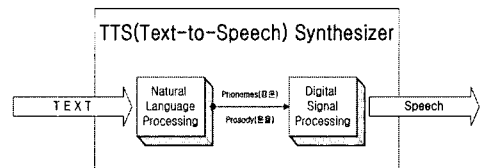
sym	예제	ID
aa	father	10
ae	cat	11
ah	cut	12
ao	dog	13
aw	foul	14
ax	ago	15
ay	bite	16
b	big	17
ch	chin	18
d	dig	19
dh	then	20
eh	pet	21
er	fur	22
ey	ate	23
f	fork	24
g	gut	25
h	help	26
ih	fil	27
iy	feel	28
jh	joy	29
k	cut	30
l	lid	31
m	mat	32

sym	예제	ID
n	no	33
ng	sing	34
ow	go	35
oy	toy	36
p	put	37
r	red	38
s	sit	39
sh	she	40
t	talk	41
th	thin	42
uh	book	43
uw	too	44
v	yat	45
w	with	46
y	yard	47
z	zap	48
zh	pleasure	49

(표 3) 음소 변동 규칙(sym : Symbol, ID : Identification).

문자를 음운 변화적으로 묘사하고 사람이 말할 때 요구되는 억양과 운율에 대한 NLP(Natural Language Processing)과 이러한 특징들을 디지털 신호로 변환해 해주는 DSP(Digital Signal Processing)로 구성되어 있다[3].

NLP(Natural Language Processing)은 높은 특질의 TTS 시스템을 위한 합성 기술인 Morpho 합성(단어의 내부구조를 살피고, 여러 텍스트 유형들에 사용된 단어들의 규칙적인 형성, 곧 단어형성의 규칙성과 가능성을 통한 합성)을 기반으로 한다.



(그림 1) TTS 시스템 구조[3].

3.1 전처리

TTS 시스템에 입력 가능한 텍스트는 간단한 문자에서부터 논술논문까지 낭독체에서 대화체까지 다양하다. 텍스트에는 영어뿐만 아니라 숫자, 약어, 기호, 전문용어 등이 있다. 이러한 정보에 대한 처리가 전처리에서 이루어진다. 예를 들어 시간 표시로 "2:20"를 "Two hour twentieth minute"으로 변경하며, 기호와 단위 "kg"을 "kilogram"등으로 변환작업을 수행한다.

3.2 Morpho 분석기

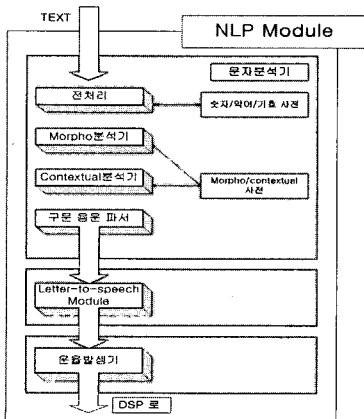
문자가 가지는 기본적인 철자를 통해 각 단어가 가지는 고유한 특성을 가능한 모든 발음 범주를 적용하는 처리가 이루어진다. 변형, 파생, 복합된 단어들은 어간과 접사의 어휘목록을 이용하는 최소화 시킨 문법 규칙을 통해 그 요소들이 분석된다. 이에 대한 처리가 이루어진다.

3.3 Contextual 분석기

Contextual 분석기는 문장 성분의 분석을 처리한다. 자연스러운 운율 정보를 생성하기 위해 문장의 구문 정보를 추출해 내는 역할을 해야 하는데 이에 대한 처리가 이루어지는 부분이다.

3.4 구문 음운 파서

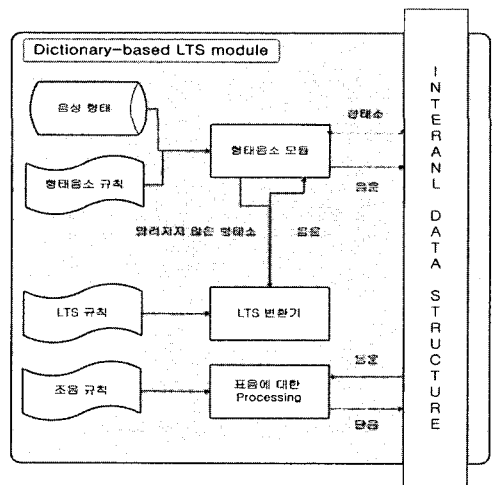
임의의 문장을 빠른 시간에 고유한 결과를 낼 수 있도록 구현되며 음성출력에 영향을 미치는 운율 정보를 생성하기 위한 표층구조를 주목적으로 한다. 모든 단어들에 대한 음성 정보를 사전(Data base)에서 찾을 수는 없다. 이에 대해 사전정보를 기반으로 새로운 단어에 대한 정보를 추출해 내는 처리를 한다.

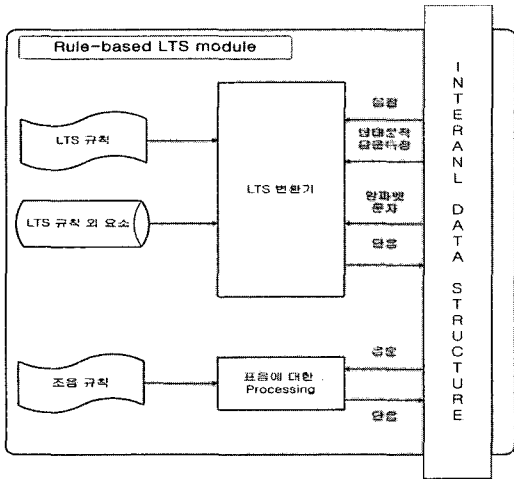


(그림 2) NLP Module 구조.

3.5 Letter to Speech 모듈

문자는 동일한 철자에 다른 발음을 가지는 동형어의어들이 있다. 예를 들어 "record"는 /rikɔɑrd/는 "기록하다"라는 뜻을 가지고 있다. 그러나 /rekɔɑrd/는 "기록"이라는 명사이다. 동사형과 명사형에 따른 발음에 차이를 가지고 있다. 이에 대해 Letter to Speech 모듈에서는 이에 대한 분석을 이행한다. 각 단어의 품사에 따른 분류를 실시하여 발음에 차이를 두어 정보를 처리한다. 처리방법에는 사전목록을 기초로 하는 방법 있고, 변환 규칙을 통한 방법이 있다. 사전목록에 의한 방법은 데이터베이스에 저장되어 있는 문자에 대한 정보를 통해 동형어의어등에 대한 처리를 한다. 이 방법은 12,000개가 넘는 형태소들을 데이터베이스에 저장해 두어야 하며, 현재는 43,000개 가량으로 늘어나 있다. 반면에 변환 규칙을 통한 방법은 2,000개 정도의 규칙에 대한 처리를 통해 70%이상의 문자들에 대한 처리가 가능하다.



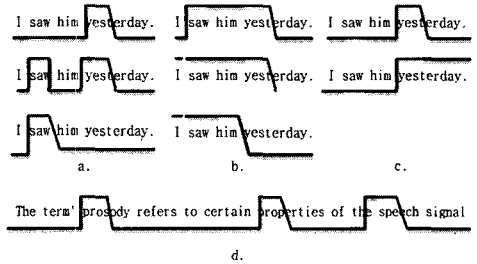


(그림 3) Dictionary-based(위)와 Rule-based (아래) 비교.

3.6 운율 발생기

운율 발생기는 음조, 음량, 음절의 길이의 변화에 의한 음성 신호의 특성들을 참조한다. 예를 들어 문맥상 어디를 강조하느냐에 따라, 단어와 단어 사이의 관계에서, 문맥상 끝과 문맥의 연결에 따라, 연결음절 내 문맥의 의미론적인 차이에 의해서 차이를 나타낸다.

운율 발생기에서는 이런 차이에 대한 처리를 한다. 그러나 문맥상 강조에 의한 차이와 단어와 단어 사이의 관계에 의한 차이는 매우 적다. 반면에 문맥의 끝과 연결에 의한 차이와 문맥 내 의미에 의한 차이는 매우 크며 중요하다. 운율 발생기에서는 앞에 두 차이를 최소화 시키고 뒤에 두 경우를 강화하였다.



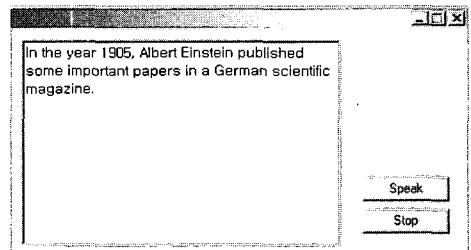
(그림 4) 억양에 의한 정보의 차이 비교.

- a. 문맥상 강조에 의한 차이.
- b. 단어와 단어사이의 관계에 의한 차이.
- c. 문맥의 끝과 연결에 의한 차이.
- d. 문맥 내 의미에 의한 차이.

4. 실험 및 평가

4.1 실험

자연 발음 생성기는 윈도우즈 환경에서 C++로 구현하였다. 실험에 대한 자연 발음 생성에 대한 평가를 위해 "Spectrogram version 7"와 "Goldwave version 4.26"를 사용하여 실제 원어 대화체 음성과 구현된 엔진에 의한 음성데이터를 분석하였다. 실험 데이터는 아인슈타인의 상대성 이론에 관한 강연의 내용으로 44개의 문장과 1,259어절을 실제 한 문장씩 입력하여 수동으로 실험 분석하였다.



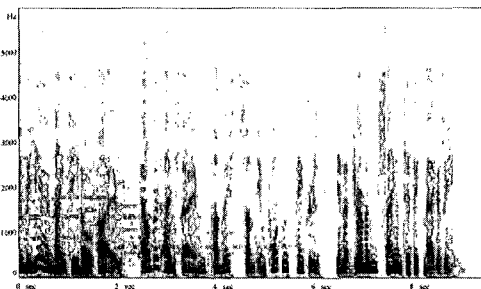
(그림 5) 실행 데모 화면.

sym	TTS	실제음성
aa	47	46
ae	51	52
ah	58	59
ao	48	50
aw	67	65
ax	69	67
ay	68	68
b	59	58
ch	70	73
d	75	75
dh	65	64
eh	61	61
er	66	69
ey	70	69
f	69	69
g	66	66
h	63	63
ih	67	66
iy	64	68
jh	67	67
k	55	55
l	64	63
m	65	64

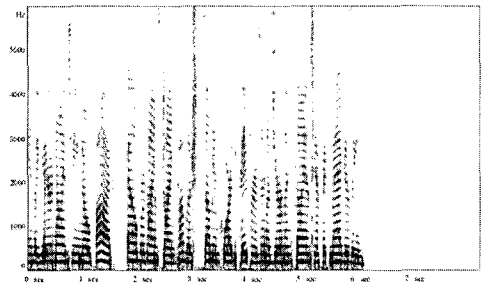
sym	TTS	실제음성
n	70	69
ng	73	73
ow	71	71
oy	67	66
p	37	41
r	65	65
s	64	64
sh	61	63
t	62	62
th	58	57
uh	68	70
uw	60	56
v	65	65
w	67	67
y	65	65
z	67	67
zh	62	62

(표 4) Spectrogram을 통한 dB 분석표(sym : Symbol).

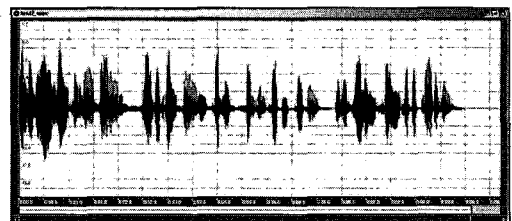
“아인슈타인의 상대성 이론”을 실험 데이터로 선택한 것은 전문용어가 많이 사용되었으며, 특수기호(숫자, 약어, 기호)가 많이 사용되었기 때문이다. TTS엔진을 통하지 않은 분석 데이터는 실제 원어민의 음성을 녹음하여 사용하였다.



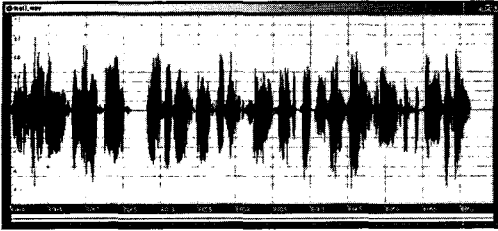
(그림 6) 실제 음성에 따른 Spectrogram 결과.



(그림 7) TTS 엔진을 통한 Spectrogram 결과.



(그림 8) 실제 음성에 따른 wave 결과.



(그림 9) TTS 엔진을 통한 wave 결과

실험에 대한 분석은 Spectrogram을 통해 음향을 측정하였다. 그리고 wave 파형에 대한 분석을 하였다. 실험 데이터에 음소변환 규칙의 적용을 통해 1,295 어절 중 1,042 어절에 대해 정확성을 나타내 81%의 정확성을 얻었다. 이에 대한 실험은 각 음소의 음향과 wave 파형을 상호 비교 분석하였다. 예를 들어 숫자 “1905” 경우 32~34

dB를 나타내었다. 이에 TTS 엔진을 통한 결과 같은 범위에서 나타났으며, 파형도 +0.5~-0.2로 동일하게 나타났다. 그러나 0.25초의 운율에 대한 오차를 범하고 있다. 이는 운율 발생기에서 단어와 단어 사이의 음절의 길이에 대한 처리를 최소화하였기 때문이다. 본 시스템에서는 연속된 문맥에 대한 처리를 강화하였기 때문에 음절길이에 따른 오차가 발생하였다. 또한 시스템 메모리에 따른 차이 또한 오차 발생의 원인이다.

측정 데이터에 대해 음소 변화 규칙에 적용에 따른 정확성을 분석하기 위해 Spectrogram을 통해 dB에 대한 분석표[표4]와 대조하여 실험결과를 나타내었다. 음소에 대한 평균오차는 ±1.1dB가 나왔으며, wave를 통한 Hz에 대한 분석표[표5]와 대조한 결과 평균오차는 0.02Hz가 나왔다.

sym	TTS	실제음성	sym	TTS	실제음성
aa	-0.80~080	-0.80~0.77	n	-0.46~0.49	-0.46~0.49
ae	-0.45~0.35	-0.44~0.30	ng	-0.60~0.46	-0.63~0.48
ah	-0.45~0.30	-0.45~0.27	ow	-0.50~0.34	-0.52~0.32
ao	-0.75~0.75	-0.80~0.65	oy	-0.46~0.41	-0.44~0.40
aw	-0.50~0.45	-0.40~0.43	p	-0.39~0.54	-0.39~0.53
ax	-1.00~0.75	-0.90~0.78	r	-0.55~0.54	-0.54~0.55
ay	-0.75~0.70	-0.77~0.72	s	-0.50~0.48	-0.51~0.48
b	-0.55~0.40	-0.50~0.43	sh	-0.39~0.53	-0.37~0.53
ch	-0.80~0.80	-0.83~0.77	t	-0.45~0.37	-0.45~0.37
d	-0.60~0.70	-0.60~0.70	th	-0.56~0.54	-0.54~0.54
dh	-0.60~0.50	-0.58~0.52	uh	-0.50~0.47	-0.50~0.46
eh	-0.40~0.54	-0.42~0.53	uw	-0.60~0.45	-0.60~0.43
er	-0.55~0.57	-0.52~0.55	v	-0.67~0.55	-0.66~0.55
ey	-0.69~0.48	-0.62~0.48	w	-0.50~0.70	-0.50~0.70
f	-0.65~0.64	-0.65~0.64	y	-1.00~0.47	-0.97~0.45
g	-0.64~0.46	-0.64~0.46	z	-0.63~0.45	-0.62~0.44
h	-0.41~0.25	-0.40~0.25	zh	-0.37~0.40	-0.36~0.40
ih	-0.50~0.60	-0.54~0.52			
iy	-0.58~0.60	-0.54~0.62			
ih	-0.83~0.65	-0.80~0.65			
k	-0.65~0.55	-0.60~0.55			
l	-1.00~0.71	-1.00~0.69			
m	-0.55~0.49	-0.55~0.49			

(표 5) wave를 통한 진폭 분석표.

이러한 오차는 실제 음원에 대한 녹음 환경에 따른 오차로 각 대조군에 대한 결과는 매우 근소하다는 결론을 얻을 수 있다.

5. 결론

본 논문의 시스템은 입력 텍스트를 전처리 과정과 Contextual 분석기를 통해 영어와 함께 쓰이는 특수 문자와 기호들을 정규화 한다. 특히 Letter-to-Speech 모듈에서 기존의 Dictionary-based 방식이 아닌 Rule-based 방식을 채택하여 효율성 높였으며, 음소에 대한 규정된 데이터를 통한 처리하여 불규칙한 음운 변화의 발음열을 얻어낸다.

실험 분석 결과, 이 논문에서 제안한 시스템은 음운론적인 분석을 통해 모음과 자음에 규칙이 적용되어 성능을 개선하였고 효율적인 메모리 사용을 제시하고 있다.

향후 연구로는 본 논문의 시스템을 각종 장치에 적용하여 응용범위를 넓히는 시도를 하여야 하겠다. 그리고 음절 길이에 대한 오류를 보완하기 위한 방안을 모색하여 보다 정확한 TTS 엔진 구현을 위한 노력을 해야 할 것이다.

참고문헌

[1] 구회산, 고도홍, "음성학과 음운론", 한신문화사, 1998.8
 [2] 이남근, 전병만, "MIT 언어학", 한신문화사, 1999.10
 [3] Breen A, "Speech synthesis models: a review", Electronics & Communication Engineering Journal, v. 4(#1), pp.

19-31, 1992

- [4] 한진수, "음성신호처리", 오성미디어, 2000.11
 [5] 남현석, "LUT(Look Up Table)을 이용한 TTS(Text-To-Speech)프로그램 구현에 관한 연구", 세종대학교 정보통신대학원 석사학위논문, 2002
 [6] 정경석, "형태 음운론적 분석에 기반한 자동 발음열 생성", 전남대학교 대학원 석사학위논문, 2002
 [7] 장석복, "TTS를 위한 자동 억양곡선 생성방식에 관한 연구", 부산대학교 대학원 석사학위논문, 1999
 [8] Sanghun Kim, Youngjik Lee, and Keikichi Hirose, "Unit Generation Based on Phrase Break Strength and Pruning for Corpus-Based Text-to-Speech," ETRI Journal, Vol. 23, No. 4, Dec. 2001
 [9] Sanghun Kim, Jong-Jin Kim, and Youngjik Lee, "SYNTHESIS UNIT GENERATION FOR CORPUS-BASED TEXT-TO-SPEECH CONVERSION," ICSP2001, Seoul, Vol. 2, pp. 415-418, 2001
 [10] Richard Sproat and Joseph Olive, "An approach to text-to-speech synthesis", Speech coding and synthesis, Elsevier press, 1995



조 정 호

2003년 남서울대학교 멀티미디어학과 졸업
현재 인큐브테크놀로지 연구원



김 태 은

1989년 중앙대학교 전기공학과 (학사)
1992년 중앙대학교대학원 전자공학과(석사)

1997년 중앙대학교대학원 전자공학과(박사)
1995년 삼성휴먼테크 논문대상 은상수상
1994년~1996년 한국과학재단(ERC) 참여연구원
1997년~현재 남서울대학교 멀티미디어학과 부교수



임 재 한

1998년 숭실대학교전자공학과(공학박사)
1999년 ~ 현재 남서울대학교 멀티미디어학과 조교수