

음운환경을 고려한 피치 기반 세기 제어 모델

Pitch-based Intensity Control Models Considering Phonetic Environments

김기중*, 김진영**, 진달복***

(Ki-Jung Kim*, Jin-Young Kim**, Dal-Bock Chin***)

요약

본 논문에서는 음성합성기에서 합성음의 자연성을 향상시키기 위한 한 방법으로 운율 제어요소중 세기 제어 모델을 통계적인 방법을 이용하여 제안하였다. 세기 제어 모델을 구성하기 위하여, 음성의 세기는 피치에 비례한다는 사실에 기반하여, 녹음된 음성 데이터에서 음절의 각 모음에 대한 평균과위와 평균피치를 구하고, 피치를 종속변수로 하는 모음의 세기 예측식을 만들었다.

예측식에 의하여 얻어진 세기값과 실제 측정된 세기값과의 상관도는 전체모음에 대하여 약 0.63으로 나타났다. 본 논문에서는 예측 세기와 실제 세기와의 상관도를 향상시키기 위하여 피치만을 고려하여 예측된 세기값을 음운환경을 고려하여 선형 오차를 보정 하는 방법을 제안하였고, 이 방법을 이용하여 얻어진 세기값은 전체모음에 대하여 실제 세기값과 0.71의 상관도로 향상시킬 수 있었다.

ABSTRACT

There require elaborate prosody control schemes to implement a natural-sounding Text-to-Speech system, one of which is the intensity control. We propose a pitch-based model for the intensity control, in this paper, statistically derived. We recorded a set of speeches, and partitioned them into syllables for corresponding vowels. Based on the well known fact that the intensity of the speech is basically proportional to the pitch frequency, we averaged the intensity powers and pitch frequencies, respectively, for each vowels, and set a predictive equation of the intensity as a function of the pitch.

The total average of the correlation between the predicted intensities and the measured, of all vowels, was about 0.63. Further to improve the correlation of the predicted values, we proposed linear compensation schemes, additionally adjusting the predicted errors depending on the combination of the syllables considering the phonetic environments. The compensated scheme is shown to afford about 0.71 of the correlation value.

I. 서론

음성합성기의 평가는 인간이 합성음을 듣고 올바르게 이해하는가를 따지는 음의 명료성(intelligibility)과 인간이 이야기하는 것처럼 자연스러운지를 따지는 자연성(naturalness)으로 고려되어진다. 지금까지의 한국어 음성합성시스템에 대한 많은 연구결과 합성음의 명료성은 어느 정도 확보되었으나 자연성은 뒤떨어져 있는 실정이다. 자연성은 운율(prosody)과 관계가 있으며 운율은 억양(pitch), 길이(duration), 세기(intensity) 정보가 포함되어 있다[1, 2]. 우리말의 경우 이들 운율정보는 정도의 차이는 있으나 운율요소 중 어느 하나라도 간과하면 자연

스러운 합성음을 얻는데는 문제가 있다. 그러나 현재 억양 및 길이에 대해서는 많은 연구가 이루어지고 있으나 [2-6], 세기에 대해서는 체계적인 연구가 부족한 실정이다.

운율정보로써 세기 정보는 어떤 요소를 문장 속의 임의의 위치에서 얼마나 큰소리로 발음할 것인가를 나타내는 것이며, 결국 이 세기는 음향학적 파워(acoustical power)로서 나타나게 된다. 세기는 피치의 상승과 같은 음향학적 조건에 따라 달라지기도 하는 것으로 알려져 있으며, 구문론 적인 요인의 영향을 받기도 한다[5, 6]. 예를 들어 평서문의 끝에서 세기는 줄어들어야 자연스러우며, 의문문의 끝에서 커져야 자연스럽다. 문장 내에서 세기는 문장의 운율성분과 관련이 있고 음절의 앞부분이 크며, 자유보다 모음이 큰 경향이 있다. 본 논문에서는 세기에 대한 현상 중 음향학적인 측면에서 고찰 하고자 한다. 음절 및 음소 세기의 구문론적인 변화는 고려하지 않고 문어체 문장에 대하여 피치에 따른 세기를 분석하고 이를 모

* 광양전문대 전산과

** 전남대학교 전자과

*** 원광대학교 전자과

접수일자: 1997년 11월 26일

델팅 하여 수학적으로 이들에 대한 규칙을 정하는 것이다. 본 연구에서 구체적으로 밝혀지겠지만, 실험에 의하면, 음절의 세기가 피치만을 사용하여서는 약 60% 정도 밖에 추정할 수 없었다. 따라서 세기조절의 방법으로서 피치와 아울러 음절의 음운학적 환경에 따른 보정값을 사용하여 세기를 적절하게 조절하는 방법을 제시한다.

II. 분석용 DB

기존합성기의 운율제어 방법의 문제점은 기본적으로 top-down의 방식을 채택하여 운율규칙을 사용하였다는 점에 있다. 즉 몇몇의 간단한 규칙을 정하고, 정해진 규칙의 파라미터를 조절하면서 주어진 규칙 하에서 가장 자연스럽게 느껴지도록 파라미터 값을 정한 것이다. 본 연구에서는 이런 방법의 문제점을 극복하기 위하여 근본적인 접근 방법을 선택하였다. 즉, 다수의 문장을 녹음하여 음성 DB를 구축하고 구축된 음성 시료들을 분석하여 통계처리를 함으로써 규칙을 구해나가는 방법이다. 즉 통계적(bottom-up)인 방법을 이용하여 연구를 수행하도록 하였다. 따라서 음성시료를 녹음하기 위한 문장 및 녹음된 문장의 분석 등이 매우 중요하게 된다.

2.1 DB의 선택

음성 DB용 문장들은 임의의 문장을 선택하여 사용하였다. 특히 문장 선택은 문장간에 나타나는 음운현상까지 고려하기 위하여 문단 단위로 녹음하고, 문장 단위로 구분하여 DB화 하였다. 분석용 DB의 크기는 20개의 문단으로 구성되어 있으며, 이는 101개의 문장과 1855개의 어절로 이루어져 있다. 그중 하나의 문단은 표 1과 같다.

표 1. 분석용 DB의 문단 예

요즘 나라 전체의 갖가지 시끄러움은 결국 어려운 경제를 회생시키자는 여러 가지 형태의 노력의 표출일 것이다. 무역적자가 이미 2백20억 달러를 넘어섰다. 특히 반도체, 정보통신 등 첨단 정보산업분야의 무역적자가 갈수록 커지고 있다. 모두들 고효율화가 해결방안이라 한다면, 첨단 정보기술에 의한 산업 및 경제구조로의 전환이 우리 나라에 걸맞은 처방이라 할 수 있을 것이다. 따라서 효율적인 산업, 기술정책과 전략이 어느 때보다 중요하다.

2.2 문장의 녹음 및 A/D

문장의 녹음은 대학방송국의 스튜디오에서 녹음을 하여 잡음이 섞이지 못하도록 하였다. 또한 녹음은 발음훈련을 받은 적이 있는 방송국의 여자 아나운서로 하여금 녹음을 하였다. 또한 A/D는 음성 분석용으로 개발된 프로그램 "SEGMENTOR"를 사용하였다.

2.3 분석용 DB의 구조

분석용 DB는 그림 1과 같이 구분되어 있으며 음성 wave 파일은 8[Hz]로 샘플링 되었고, 피치 contour 파일은 피치의 위치정보를 포함하고 있으며 "SEGMENTOR" 상에서 수작업으로 구하였다. 레이블링 파일은 음소단위로 label 하였으며, 구문분석정보를 포함하는 구문분석 파일 등으로 구성되어 있다.

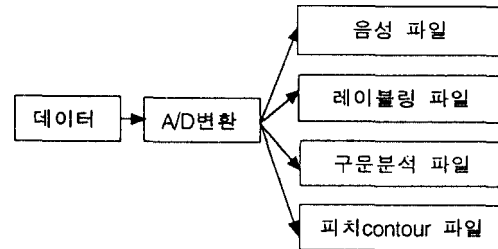


그림 1. 분석용 DB 구조

이상의 각 파일들은 하나의 문장에 대해 세트로 이루어져 발음된 문장에 나타나는 운율구조를 다양한 방법으로 연구할 수 있는 기반을 제공하며, 음성 DB는 현재 웹 상에서 공개되어지고 있다(<http://dsp.chonnam.ac.kr>).

III. 피치에 따른 세기 제어 규칙

세기 분석은 음절 단위로 이루어졌다. 그 이유는 비음 또는 유음과 같은 중성이 붙은 경우, 중성의 크기는 일정 비로 결정할 수 있기 때문이다. 따라서 음절 단위로 세기 분석을 하되, 모음의 세기를 측정하였다. 모음 구간중 세기와 피치의 대표값을 결정하기 위하여 다음과 같은 4가지 조건을 만들고 이중 상관도가 높게 나타나는 조건을 선택하여 분석할 것이다.

- 조건 1: 최대파워와 최대파워가 포함된 피치
- 조건 2: 평균파워와 안정구간의 피치
- 조건 3: 평균파워와 평균 피치
- 조건 4: 안정구간의 파워와 피치

조건 1은 음절내 각 피치값 사이의 파워를 구하여 그중 최대값을 갖는 것을 최대파워라 놓고 그곳의 피치를 선택하였다. 조건 2는 각 피치구간의 파워를 평균하여 구하고 피치는 모음의 중간지점을 안정구간이라 가정하고 중간 지점의 피치를 선택하였다. 조건 3은 모음구간내의 평균파워와 평균피치 값을 선택하고, 조건 4는 모음 중간지점의 피치와 그 구간의 파워 값을 구하여 비교하였다. 본 연구에는 위의 4가지 조건에 대해서 피치와 세기의 상관도를 분석하여 상관도가 높게 나타나는 조건을 가지고 세기의 예측식을 만들기로 하고 모음 "ㅏ"에 대하여 구축된 DB에서 상관도 분석을 하였으며, 표 2와 같은 상관

표 2. 모음 "ㅏ"에 대한 각 조건의 피치와 파워의 상관도

	조건	조건1	조건2	조건3	조건4
상관도					
(피치, 세기)		0.46	0.64	0.68	0.56

도 값을 얻었다.

본 연구에서는 위의 조건 중 조건 3의 결과가 가장 상관도가 높게 나타남을 확인하고, 모음구간내의 평균 파워와 평균 피치값을 이용하여 세기 예측식을 만드는 것이 타당하리라는 결론을 얻었다. 분석이 이루어진 모음은 단모음을 대상으로 하였다. 왜냐하면 복모음의 개수는 녹음된 음성 DB에서 개수가 매우 적었으며, 또한 복모음의 세기는 단모음에 기준 하여 정할 수 있기 때문이다.

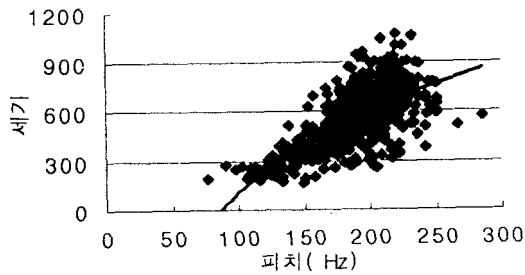


그림 2. 모음 "ㅏ"에 대한 피치에 따른 세기 분포

그림 2에 구축된 DB에서 모음 "ㅏ"를 포함하는 음절의 평균파워와 평균피치값을 나타내고 있다. 그림 2에 의하면 "ㅏ" 모음의 경우 피치와 세기는 큰 상관도를 보임을 알 수 있다. 물론, 임의의 피치에 대하여 세기의 분포가 넓게 분포되어 있기는 하지만, 피치의 함수로서 세기를 어느 정도 예측할 수 있음을 예상할 수 있다.

그림 2에서 곡선으로 나타낸 값은 로그함수를 사용하여 피치를 종속변수로 세기를 예측하는 추세곡선이다. 추세곡선은 주어진 데이터에 대한 회귀분석의 결과인데, 회귀분석은 마이크로소프트사의 엑셀에 내장되어있는 회귀분석 툴을 사용하였다. 이 곡선의 추세식은 최소자승법에 의하여 식 (1)과 같이 얻어진다[9].

$$\text{예측 세기값} = 718.41 \cdot \ln(\text{피치 주파수}) - 3220.6 \quad (1)$$

모음 "ㅏ"에 대하여 식 (1)을 이용하여 구한 예측 세기값과 실제 세기값과의 상관도 분석을 하여 0.68의 상관도를 가지고 세기를 예측할 수 있었다. 표 3에 각 모음에 대한 세기의 예측식을 나타내며, 각 모음에 대한 예측식에 의해 얻어진 예측 세기값과 실제 세기값과의 상관도 분석결과는 표 4에 나타낸다. 각 모음에 따른 예측값을 구해 상관도를 분석한 결과는 0.63의 상관도를 나타내어 피치에 의하여 어느 정도 세기를 예측할 수 있음을 알 수 있다.

표 3. 각 모음에 대한 기준 예측식

모음	예측식
'ㅏ'	$718.41 \cdot \ln(\text{freq}) - 3220.6$
'ㅑ'	$741.80 \cdot \ln(\text{freq}) - 3319.1$
'ㅓ'	$932.89 \cdot \ln(\text{freq}) - 4276.1$
'ㅕ'	$574.77 \cdot \ln(\text{freq}) - 2461.1$
'ㅡ'	$905.01 \cdot \ln(\text{freq}) - 4156.3$
'ㅣ'	$699.16 \cdot \ln(\text{freq}) - 3118.6$

표 4. 각 모음별 예측 세기값과 실제 세기값과의 상관도

	'ㅏ'	'ㅑ'	'ㅓ'	'ㅕ'	'ㅡ'	'ㅣ'	평균
상관도	0.68	0.67	0.66	0.78	0.51	0.46	0.63

IV. 음운환경을 고려한 세기 예측 보정

앞상에서 세기조절을 위한 기본 규칙으로서 피치와 세기와의 관계를 밝혔으며 로그함수를 사용하여 일차적인 세기 규칙을 정하였다. 그런데 표 4의 결과에 의하면, 예측치와 실제치와의 상관도가 전체모음에 대해 0.63으로 세기의 예측에 있어서 다른 성분을 참조하여 상관도를 좀 더 높일 필요성이 있다.

간단한 실험에 의하여, 중성의 세기는 초성 또는 종성 등의 음운학적 환경에 의하여 세기가 변화함을 알 수 있는데, 이는 이미 여러 연구 결과에서 지적된 바이다. 따라서, 세기의 추정을 좀 더 정확하게 하기 위해서는 전후의 음운환경을 적당히 분류하여 이에 따른 세밀한 분석을 해야 할 것이다.

표 5. 음소 환경 표

	초성조건	종성조건
제1조건	ㄱ, ㄷ, ㅂ, ㅅ	ㄱ, ㄷ, ㅂ
제2조건	ㄴ, ㄹ	ㄴ, ㄹ, ㅇ
제3조건	ㄹ	ㄹ
제4조건	ㅅ, ㅆ, ㅎ	ㄱ, ㄷ, ㅂ, ㅅ
		ㄴ, ㄹ
		ㄹ
제5조건	ㅈ, ㅊ, ㅊ, ㅈ	**
		ㅅ, ㅆ, ㅎ
제6조건	ㅋ, ㅌ, ㅍ, ㅊ	ㅈ, ㅊ, ㅊ, ㅈ
		ㅋ, ㅌ, ㅍ, ㅊ
제7조건	XX	XX

표 5는 본 논문에서 적용한 음운환경 분류를 보인 것이다. 표 5에 의하면 초성의 조건은 7가지 종성의 조건은 4가지로 총 28가지의 조건이 된다. 그러나 종성이 없는 경우 다음 음절의 초성의 조건을 고려해야 한다. 왜냐하면,

예를 들어, “아까” 같은 발음의 경우 뒤따르는 음절의 초성 “ㄱ”을 발음하기 위해서 앞 음절 “ㅏ”의 길이가 단축되는 현상을 보이기 때문이다. 따라서 종성이 없는 경우(*)는 다시 다음 음절의 초성 조건을 고려하는 7가지 조건으로 나누어진다. 따라서 종성의 조건은 모두 10개가 있다고 볼 수 있다. 따라서 이를 종합하면 모두 70개의 음운학적 조건을 만들 수 있다. 이렇게 만들어진 조건에 대해 각 모음의 전후 음운환경에 따라 표 3의 식에 의해서 얻어진 세기값과 실제 세기값과의 오차를 평균하여 대표값을 만들어 Table와 하고, 합성 시에는 입력되는 음소열의 조건에 따라 해당되는 대표값을 주는 방법으로 하고자 한다. 본 논문에서 제시한 세기제어 모델의 전반적인 개념도를 그림 3과 같이 나타낼 수 있다.

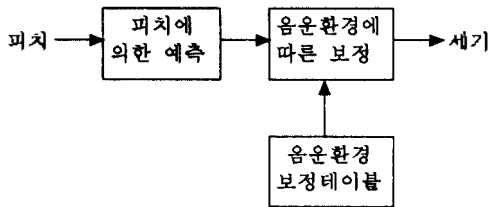


그림 3 세기 제어 흐름도

4.1 세기 제어 규칙의 보정

그림 2의 모음 “ㅏ”에 대한 예에서 피치와 세기의 값이 추세곡선에 대해 일차함수 형태의 오차값을 갖고 있음을 알 수 있다. 본 절에서는 표 3의 식에 의해 얻어진 세기 예측의 경우에 있어서보다 세기예측의 상관성을 향상시키기 위한 방법으로 다음과 같은 세 가지 오차 보정 모델 방법을 제시하고 분석하고자 한다.

- i) 영차 가산 보정 모델
- ii) 영차 승산 보정 모델
- iii) 일차 선형 보정 모델.

본 논문에서 제시하는 세 가지 오차 보정 모델은 피치만을 고려하여 예측된 세기값에 보정치를 식(2~4)와 같이 연산함으로써 보정을 취한다.

$$\text{영차 가산 보정 모델: } \widehat{I}_{P_i} = I_{P_i} + Ea(P_i) \quad (2)$$

$$\text{영차 승산 보정 모델: } \widehat{I}_{P_i} = I_{P_i} \cdot Em(P_i) \quad (3)$$

$$\text{일차 선형 보정 모델: } \widehat{I}_{P_i} = I_{P_i} \cdot E_1(P_i) + E_2(P_i) \quad (4)$$

여기서, \widehat{I}_{P_i} = 보정후 세기값

I_{P_i} = 표 3의 식에 의한 피치만을 고려한 세기 예측값

$$Ea(P_i) = \frac{1}{N_i} \sum_i (P_i - I_{P_i})$$

$$Em(P_i) = \frac{\sum_i I_{P_i} \cdot P_i}{\sum_i I_{P_i}^2}$$

$$E_1(P_i) = \frac{\sum(I_{P_i} P_i) - \sum I_{P_i} \sum P_i}{N_i \sum I_{P_i}^2 - (\sum I_{P_i})^2}$$

$$E_2(P_i) = \frac{\sum I_{P_i}^2 \sum P_i - \sum I_{P_i} \sum (I_{P_i} P_i)}{N_i \sum I_{P_i}^2 - (\sum I_{P_i})^2}$$

위에서 논의된 다양한 보정 방법의 특성을 비교하기 위해 본 논문에서는 보정후 예측된 세기모델이 갖는 자승오차 및 실제 측정된 세기와 보정 예측된 세기를 성능 기준으로 설정하고 식 (5), 식 (6)과 같이 정의한다. 보정후 자승평균오차는 피치에 의한 예측 및 보정이 완료된 모델에서 발생되는 세기와 실제 세기간의 오차의 자승값으로 보정이 양호할수록 그 값이 작아지며, 보정후 상관도는 그 모델에서 발생되는 세기와 실제세기의 관련을 나타내주는 지수로서 보정이 양호할수록 1에 근접하는 값을 갖는다.

$$\text{보정후 자승평균오차: } Er_2^2 = \frac{1}{N_i} \sum (P_i - \widehat{I}_{P_i})^2 \quad (5)$$

$$\text{보정후 상관도: } \rho(P_i, \widehat{I}_{P_i}) =$$

$$\frac{\sum [P_i - \frac{1}{N_i} \sum P_i] \cdot [\widehat{I}_{P_i} - \frac{1}{N_i} \sum \widehat{I}_{P_i}]}{\sqrt{\sum [P_i - \frac{1}{N_i} \sum P_i]^2 \cdot \sum [\widehat{I}_{P_i} - \frac{1}{N_i} \sum \widehat{I}_{P_i}]^2}} \quad (6)$$

4.2 보정된 세기의 성능평가

실제 세기값과 앞 절에서 제시된 3가지 보정 모델에 의하여 얻어진 예측값과의 상관도 분석 및 보정후 자승오차를 비교하기 위하여 모음 “ㅏ”에 대하여 분석을 하여 표 6과 같은 결과를 얻었다. 표 6에 의하면 상관도는 식 (1)에 의해 구해진 세기예측과 가산형 보정 모델(식 2)에 의하여 얻어진 세기값과의 상관도는 0.04의 상관도 차이를 보이나, 승산 보정 모델(식 3)이나 일차 선형 보정 모델(식 4)에 의한 방법에 의해서 얻어진 상관도는 차이가 거의 나지 않았다. 또한 보정후 자승평균오차에서도 기준예측에 비하여 가산형이 오차를 다른 방법에 비하여

표 6. 보정 방식별 상관도 및 보정후 자승오차 비교

'ㅏ'	상관도	보정후 자승오차	자승오차차이
기준예측	0.68	16514	-
가산형	0.72	13172	3342
승산형	0.72	12811	361
일차선형	0.73	11784	1027
비교	$\rho(\widehat{I}_{P_i}, P_i)$	Er_2^2	

현격하게 줄임을 알 수 있다.

표 6의 결과만을 고려한다면 일차선형 보정 모델을 이용하여 세기의 보정 모델을 선택하여야 타당하겠지만 메모리의 크기 및 연산속도 등을 고려하여 본 연구에서는 가장 손쉽고 경제적으로 구성할 수 있는 가산형을 선택하여 전체 모음에 대하여 상관도 분석을 하였다. 이를 위하여 각 모음의 전후음소의 음운환경에 따른 세기의 오차값을 평균하여 table화 하였고 그중 모음 "ㅏ"의 초성 1 조건에 대한 오차평균값을 표 7에 나타내었다.

표 7. 모음 "ㅏ"의 초성조건 1에 대한 가산모델 예측오차값

중성	초성	중성	예측오차	
ㅏ	ㄱ, ㄷ, ㅂ, ㅅ	ㄱ, ㄷ, ㅂ	-28.30	
		ㄴ, ㄹ, ㅇ	47.80	
		ㄹ	6.27	
		**	XX	-45.58
			ㄱ, ㄷ, ㅂ, ㅅ	2.10
			ㄴ, ㄹ	18.09
			ㄹ	22.17
			ㅅ, ㅆ, ㅈ	12.43
			ㅈ, ㅊ, ㅍ, ㅍ	-51.55
			ㅋ, ㆁ, ㅍ, ㅍ	-31.58

모음 'ㅏ'에 대한 세기를 음운환경 조건에 따라 가산형 보정 방법을 적용하여 구한 최종 추정값과 실측치의 상관도를 계산한 결과 상관도는 0.72로서 피치만을 파라미터로 하여 예측한 값의 상관도 0.68에 비하여 세기에측을 향상시킴을 알 수 있다. 표 8은 각 모음별로 음운환경에 따른 가산형 오차 보정 상관도 결과를 보인 것이다. 표 8에 보인 값은 본 연구에서 제안한 세기 규칙을 적용했을 때의 상관도와 피치 정보만을 이용한 경우의 상관도를 보인다. 표 8을 살펴보면 제안한 세기 제어 규칙이 매우 우수한 성능을 보임을 확인할 수 있다. 특히 모음 'ㅏ'의 경우에는 피치만을 고려한 경우 0.66의 상관도를 보였으나, 음운환경을 고려한 결과에서는 상관도가 0.77로 향상됨을 알 수 있다.

표 8. 각 모음별 가산형 오차보정 상관도 비교

모음	'ㅏ'	'ㅓ'	'ㅗ'	'ㅜ'	'ㅡ'	'ㅣ'	평균
$P(\hat{p}_i, P_i)$	0.68	0.67	0.66	0.78	0.51	0.46	0.63
$P(\hat{p}_i, P_i)$	0.72	0.74	0.77	0.80	0.61	0.59	0.71

V. 결 론

본 연구에서는 운율 제어 규칙중 세기 제어 규칙에 대하여 모델링 하였다. 세기의 적절한 제어요인을 구하기 위하여 각 음절에 나타나는 평균파워, 최대파워 등과 피치의 상관도를 분석한 결과 모음부의 평균파워와 평균파치가 가장 상관도가 높음을 확인하고, 각 모음에서 나타나는 평균파워와 평균 피치값을 이용하여 각 모음에 대한 예측 세기를 구하기 위하여 피치를 종속변수로 하는 예측식을 구하였다. 그러나 준비된 음성 DB를 대상으로 실제 세기값과 예측 세기값의 상관도 분석을 통한 조사에 의하면 각 모음에 대하여 0.63 정도의 상관성을 갖음을 알 수 있었다. 그런데 실험결과에 의하면 피치를 고려한 세기 예측식에 오차값을 음운환경 조건을 고려하여 보정을 하는 경우 더욱 적절한 세기를 조절할 수 있음을 확인하였다. 구체적으로 본 논문에서는 기준음소 전후의 음운환경 조건을 초성조건은 7가지로, 중성조건은 중성이 없는 경우에 다음 음절의 초성을 고려하는 것을 포함하는 10개 조건으로 분류하여 작아지거나 커지도록 가산형 방법, 승산형 방법, 일차 선형 방법을 제시하였다. 각각의 보정 방법에 대해 나타나는 음성 DB에서 실제 세기값과 예측된 세기값에 대한 상관도 분석 결과, 일차 선형 방법이 가장 높은 상관도를 나타내고, 승산형 방법, 가산형 방법, 무보정 기준예측 방법 순으로 성능이 좋음을 확인하였으며, 각각의 경우에 대한 보정후 오차의 추이도와 일치함을 확인하였다. 본 연구에서는 시스템의 간소화를 위하여 가산형 보정 방법을 이용하여 100여 개의 문장에 대하여 음운환경별 보정 Table값을 도출하고, 가산형 세기 제어 규칙에 적용하여 얻은 세기값과 실제 세기값의 상관도가 0.71 이상으로 적절하게 제어함으로써 제안된 규칙의 타당성을 확인하였다.

향후, 제안된 규칙의 일반성을 확보하기 위하여 다양한 문법 규칙이 반영된 대량의 CORPUS를 구축하여 세기 제어 규칙의 확장이 필요하며, 본 제어규칙을 실제 음성합성기에 적용하여 제안된 모델의 타당성을 확인할 필요가 있다.

참 고 문 헌

1. 이 윤근, 안 승권, "음성합성 기술 분야," 전자공학회지, 제 20권, 제5호, pp. 523-532, 1993.
2. 배 명진, "TTS 음성합성기술," 통신학회지, 제11권, 9호, pp. 761-772, 1994.
3. 안 승권, 국어 음성변환 시스템에서의 합성음의 자연도 향상 기법에 관한 연구, 서울대학교 박사학위 논문, 1992.
4. 김 정수, 이 혜정, "언어정보 및 통계데이터를 이용한 한국어 운율생성," 제13회 음성통신 및 신호처리 워크샵 논문집, pp. 227-231, 1996.
5. 김 진영, 한국어 음성합성을 위한 운율제어 규칙, 서울대학교 박사학위논문, 1994.

6. 지 민재, 성 철재, 전 은주, "한국어 돌돌림에 관한 실험음성학적 연구," 제9회 음성통신 및 신호처리 워크샵 논문집, pp. 21-217, 1992.
7. 이 현복, 한국어의 표준발음. 교육과학사, 1989.
8. Sagisaka Y., "Recent Advances in Japanese Speech Synthesis Research," ISSIPNNI'94, pp. 1927-1933, 1993.
9. Microsoft, User's Guide:Microsoft Excel 7.0, Microsoft Corp., 1997.

▲김 기 중(Ki-Jung Kim)

한국음향학회지 16권 7호 참조

현재:광양전문대 전산과 전임강사

▲김 진 영(Jin-Young Kim)

한국음향학회지 14권 1호 참조

현재:전남대학교 전자공학과 조교수

▲진 달 복(Dal-Bock Chin)

한국음향학회지 16권 7호 참조

현재:원광대학교 전자과 교수