

## 임베디드 TTS 시스템을 위한 아라비안 숫자의 문자 변환

정영임, 윤애선<sup>○</sup>, 권혁철

부산대학교 컴퓨터공학과, 불어불문학과<sup>○</sup>  
{acorn, asyoon<sup>○</sup>, hckwon}@pusan.ac.kr

### Grapheme-to-Phoneme Conversion of Arabic Numeral Expressions for Embedded TTS Systems

Youngim Jung, Aesun Yoon<sup>○</sup>, Hyukchul Kwon

Dept. of Computer Science, Dept. of French<sup>○</sup>, Pusan National University

#### 요 약

본 논문에서는 아라비안 숫자의 중의성을 효과적으로 제거하고 숫자 표현의 발음을 정확하게 문자화할 수 있는 임베디드 시스템용 경량화된 아라비안 숫자 읽기 시스템을 제안한다. 이를 위해 7 가지의 숫자 읽기 방식(Readings of Arabic Numerals: RAN)을 분류하였고, 문자화 규칙을 설정하기 위해, (1) 문맥 자질, (2) 패턴 자질, (3) 휴리스틱 정보를 숫자 표현의 의미에 따라 분석하였다. 그리고 숫자의 문자화 시스템을 최적화하여 임베디드 시스템에 탑재하기 위해 (1) 형태소 분석 모듈의 분리, (2) 사전 압축, (3) 인명과 지명의 제거를 하였고, 이를 통해 심각한 정확도 손실 없이 메모리 사용량과 처리 시간을 크게 줄일 수 있었다. 경량화된 mini-TAN 은 96.9~98.3%의 정확도를 보이며, 기존 상용 TTS 시스템에 비해서도 숫자 읽기의 처리에 있어 높은 정확도를 보인다.

#### 1 Introduction

최근 자연스러운 음성 합성 기술은 비약적으로 발전하였으나, 음성합성 시스템의 정확도 향상을 위한 기술의 발전은 상대적으로 미미하여, 20 여 년 전에 사용되던 '글자-소리 변환(letter-to-sound conversion)' 규칙이 숫자<sup>1</sup>, 두문자어, 약자, 인명의 처리에 여전히 사용되고 있다[2]. VIA (Voice Information Associates)에서 실시한 세계 19개 TTS 상용 제품의 정확도 테스트 결과에 따르면, 현재 TTS 시스템들은 발음의 중의성이 발생하는 영역 중 숫자의 처리 정확도가 55.6%로 가장 낮은 정확도를 보이고 있다[2].

현대 한국어에는 한국어, 한자어, 외래어의 3 개 어원을 가진 수사가 사용되며, 읽기 방식의 다양성이 존재하고, 발음의 변이형태 보여 아라비안 숫자의 정확한 발음을 생성하기 위해서는 형태소 분석과 중의성 해소 기법과 같은 자연언어처리 기술이 적용되어야 하며, 어휘 사전, 문맥 사전 및 규칙 사전을 구성해야 한다. 또한 숫자와 기호의 결합 패턴에 대한 정보 및 휴리스틱 규칙이 요구된다. 그러나 PDA, 휴대폰, 임베디드 컨트롤러와 같은 임베디드 기기들은 보조 기억 장치 없고, 가용 메모리 양이 한정되어 있으므로 가용 메모리 양을 초과하는 응용 프로그램을 돌릴 수 없다. 그러므로 응용 프로그램이 요구하는 메모리 양과 큰 어휘 사전이 차지하는 용량을 줄일 수 있는 방법이 요구된다.

따라서 본 논문에서는 경량화된 음성합성 시스템을 위해 정확하게 아라비안 숫자 표현(이하 ANEs 로 지칭함)을 한국어 음소열로 변환할 수 있는 최적화된 mini-TAN (mini-Transliteration system for Arabic-Numeral expressions) 을 제안한다. 이를 위해 본 논문에서는 (1) ANEs 의 의미와 발음을 분류하여 숫자 읽기 규칙을 설정하고, (2) 아라비안 숫자 문자화를 위한 경량화된 시스템인 mini-TAN 을 구현하며, (3) mini-TAN 과 서버용 아라비안 숫자 문자화 시스템의 성능을 비교하여 (4) 임베디드용 한국어 TTS 시스템을 개선할 수 있는 방안을 제시한다.

본 논문의 2 장에서는 숫자표현의 문자화에 관한 선행 연구와 임베디드 음성 시스템의 요구 조건을 알아보고, 3 장에서는 숫자표현의 정확한 발음을 생성할 수 있도록 아라비안 숫자의 문자화 시스템을 제안하며, 이를 임베디드용 한국어 TTS 제품에 탑재할 수 있도록 아라비안 숫자의 문자화 시스템의 경량화 방법을 제안하고 mini-TAN 을 구현한다. 4 장에서는 mini-TAN 과 상용 한국어 TTS 제품의 성능의 비교를 통해 mini-TAN 의 성능을 평가하며, 5 장에서 결론과 향후 연구를 제시한다.

<sup>1</sup> 기계가독형 텍스트에는 한글 문자뿐만 아니라 다양한 비문자가 포함되어 있으며, 정확한 정보 전달을 목적으로 할 경우 아라비안 숫자와 비-문자 기호의 출현이 빈번하다.

#### 2 선행연구

##### 2.1 ANEs 의 문자화

[윤애선 외, 2003]에 따르면, 7 가지 요인<sup>2</sup>에 의해 한국어의 숫자가 7 가지 다른 방법으로 읽힐 수 있음을 밝혔다. 그러나 5 개 기업에서 상용화된 한국어 TTS 제품을 개발하였으나 [7, 8, 9], 이들 시스템에는 아라비안 숫자가 1~3 가지로만 읽혀 55~87.7%의 정확도<sup>3</sup>를 보인다[3]. [정영임 외., 2004]에서는 아라비안 숫자표현의 문자화를 위해 결정 트리 학습 알고리즘을 적용하였다. (1) 아라비안 숫자와 기호의 결합 패턴, (2) 문맥 자질 및 (3) 휴리스틱 정보 자질이 ANEs 의 의미와 발음을 결정지을 수 있는 학습 자질로 추출되었다.

##### 2.2 임베디드 음성 시스템의 요구조건

임베디드 시스템의 특성상 각 어플리케이션과 개별 기기에 보다 적합한 경량화 방법이 요구된다. [1]. 음성 처리를 위한 임베디드 시스템의 구현 방법론에는 메모리 사용량을 줄이고 CPU 계산 복잡도를 줄이기 위해, 첫째로, 보다 로버스트한 음성 인식을 위해 개별 기기가 아닌 원격 서버에서 음성 인식이 이루어지는 분산 음성 인식 (DSR: Distributed Speech Recognition)<sup>4</sup> 방법론이 제안되었고, 둘째로, 임베디드 기기에서 직접 음성 인식과 음성 합성이 이루어질 수 있도록 음성 처리 프로그램의 경량화에 대한 연구가 진행되어 왔다. 임베디드 기기에서의 음성 합성을 위해, 데이터베이스의 특성을 파악하여 중복된 정보를 제거함으로써 일반적인 음성 코드 축약 알고리즘보다 더 나은 성능을 보일 수 있으며 데이터베이스 축약을 위해 운운 패턴과 스펙트럼 패턴을 범주화하여 600KB 크기의 데이터베이스로 합성된 중국어 음성이 자연스럽게 허용할만한 수준의 음질을 얻었다[9]. 또한, 임베디드 TTS 시스템 구현을 위해 최소 저장 메모리, 신속한 음성 처리 및 소규모 프로그램 코드에 대한 논의가

<sup>2</sup> 수사의 어원, 수사의 품사, 숫자의 의미, 숫자의 범위, 숫자의 형태, 문자 수단위 (DSM: Decimal Scale Markers, 백 (100), 천(1,000), 만 (10,000) 등)의 결합 여부, 빈이형 여부 등의 요인이 포함된다.

<sup>3</sup> 3 개 한국어 TTS 시스템에 임의추출한 데이터를 입력하여 숫자 읽기의 정확도를 평가하였다. 더욱 신뢰할 수 있는 성능 비교를 위해 4 장에서 동일한 평가용 데이터를 이용하여, 2 개 의 상용 TTS 시스템과 mini-TAN 의 성능을 비교한다.

<sup>4</sup> DSR 응용 어플리케이션은 경량화 문제나 CPU 계산 복잡도가 더 이상 문제가 되지 않고, 소비자들에게 기기가 판매된 후라도 서버를 통해 지속적인 개선이 가능하다는 장점과, 판매된 기기를 통해 새로운 노이즈 환경이 서버에 혼합될 수 있다는 장점이 있다 [7]. 그러나 DSR 을 통한 방법은 임베디드 음성 시스템의 경량화를 위한 직접적인 방법이 아니며, 네트워크 연결 및 연결 상태 등에 그 성능이 좌우될 수 있다.

진행되었고, 규칙에 기반한 문자소-음소 변환 (Grapheme-to-Phoneme Conversion)이 제안되었다 [8]. 본 연구에서는, 규칙에 기반한 아라비아 숫자의 문자화 시스템을 구현하여, 경량화 기법을 통해 심각한 성능 손실 없이 임베디드 TTS 시스템에 적용될 수 있음을 보이고자 한다.

3 시스템 구현

아라비아 숫자의 읽기 규칙의 설정을 위한 분석 데이터와 구현된 mini-TAN의 평가를 위해 실험 데이터는 10개 신문의 6년치 (1998년 1월 1일-2003년 12월 31일) 기사에서 추출되었으며 <표 1>과 같이 구성되었다.

<표 1 말뭉치 구성>

말뭉치	분석	평가용 말뭉치				
말뭉치	Set1	Set2	Set3	Set4	Set5	
어절 수	170,000	1,000	1,000	1,000	1,000	10,000
출처	10개	A	A	B	B	10개
추출방법	입의	입의	균형	입의	균형	입의

3.1 ANEs의 의미 분류 및 의미 분류 규칙

말뭉치의 분석과 선행 연구의 조사를 통해 아라비아 숫자의 의미와 읽기 방식 (RAN: Readings of Arabic Numerals)을 <표 2>와 같이 분류할 수 있다.

<표 2 아라비아 숫자의 의미와 읽기 방식>

문자열	의미범주	문맥	발음열	RAN
Ex1	수량	3 그루	셋	Kca_b
Ex2		3 세	삼	Cca_b
Ex3		3 말	서	Kca_v
Ex4	순서	3 위	삼	Cor_b
Ex5		3 째	셋	Kor_b
Ex6	스포츠 용어	3 볼	스리	Eca_b
Ex7	시간	6 월	유	Cor_v

<표 3>에서는 문맥으로 사용된 분류사에 따른 RAN의 분포를 보여준다.

<표 3 분류사에 따른 RAN의 분포>

RAN	Kca_b	Kca_v	Cca_b(+D)	Cca_b(-D)
개수	188	12	479	7
종류	수량 분류사	교유도량형	도량형	지역구단위
RAN	Eca_b	Cor_v	Kor_b	Cor_b
개수	4	1	1	121
종류	스포츠 용어	월	째	순서분류사

숫자 전치어, 문자 수단위(DSM), 숫자 후치어 및 분류사 후치어와 같은 다른 문맥 정보 역시 RAN을 결정하는 단서가 되며, 이에 대한 연구는 본 논문의 선행 연구로 [윤애선 외, 2003]에서 구체적으로 분석되었으며, 본 논문에서도 이러한 문맥 정보를 RAN을 결정하는 단서로 이용한다. 아라비아 숫자와 문장 기호와 결합한 숫자 표현은 패턴화된 구조를 보이며, 의미에 따라 숫자가 산술적 특성을 가지기도 한다. 이러한 정보는 ANEs의 의미와 발음을 결정할 수 있는 휴리스틱 규칙을 설정하는 데 사용될 수 있다. [정영임, 2004]에서 RAN의 당연값을 구하였고, [윤애선 외., 2003; 2004]에서는 숫자와 결합한 기호의 읽기 방식인 RTS (Reading of Text Symbols)의 당연값을 <표 4>에서와 같이 구하였다.

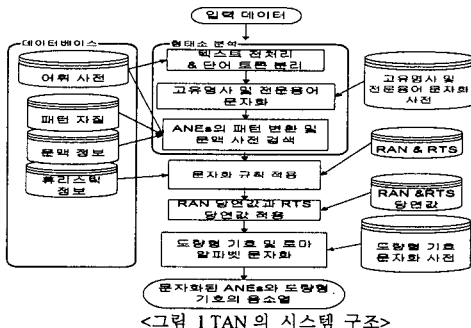
<표 4 RAN 당연값과 RTS 당연값>

RAN 당연값	RTS 당연값				
	‘	‘	‘	‘	‘
Cca_b(+D)	0	에서	점	0	대

상기에 기술한 문맥 정보, 숫자 표현의 결합 패턴 규칙 및 휴리스틱 규칙을 이용하여 숫자와 기호의 읽기 방식을 결정할 수 없는 경우 <표 4>에 나타난 RAN 당연값과 RTS 당연값을 적용하여 숫자표현의 문자화가 이루어진다.

3.2 아라비아 숫자표현의 문자화 시스템 구조

본 절에서는 ANEs 문자화의 단계적 절차와 문자화 시스템 구조를 기술한다. ANEs 문자화는 <그림 1>과 같이 아래 6 단계를 통해 이루어진다.



<그림 1 TAN의 시스템 구조>

3.3 임베디드 음성 시스템 구현을 위한 TAN 시스템의 경량화

본 절에서는 3.2 절에서 구현한 아라비아 숫자의 문자화 서버용 시스템을 임베디드 시스템용으로 최적화하기 위한 방안이 강구되었다<sup>7</sup>. 이는 (1) 문자화 시스템으로부터 형태소 분석 모듈의 분리, (2) 어휘 사전의 규모 축소, (3) 문맥 사전에서 인명 및 지명 제거의 세 가지 방안이다. 모듈 간 데이터 흐름을 최소화하고 전체 시스템 코드의 압축을 위해 형태소 분석 모듈을 제거한다. 그리고 제거된 형태소 분석 모듈을 보완하기 위해 조사 및 어미 리스트와 간단한 형태소 분석 루틴을 추가한다. 또한, 문맥 사전에서 인명과 지명으로 구성된 이름 사전을 제거하는데, 사전이 차지하는 용량이 매우 큰데 반해, 인명과 지명이 문맥으로 사용되는 빈도가 낮고 분야별로 사용되는 어휘가 한정되어 있기 때문이다.

<표 5 서버용 시스템과 MINI-TAN의 크기 비교>

크기 (MB)	서버용		mini-TAN
	학습 기반	규칙 기반	규칙 기반
프로그램	5.1	1.9	0.5
어휘 사전	20.3	11.2	2.9
인명, 지명 사전	4.5	4.5	0

<표 5>는 이상에서 제안한 3 가지 방안이 임베디드 시스템 구현을 위한 사용 공간 압축에 효율적임을 보여준다.

4 실험 및 평가

아라비아 숫자의 문자화를 위한 서버용 학습 시스템, 서버용 규칙 기반 시스템 그리고 규칙에 기반한 mini-TAN의 성능 비교를 위해, <표 1>에서 기술한 5개 실험 말뭉치 중 'Set 5'가 실험에 사용되었다. 메인 메모리 사용량, 처리 속도와

<sup>5</sup> 숫자 표현의 결합 패턴 자질과 휴리스틱 규칙의 설정에 유용한 산술적 특징에 대해서는 [정영임, 2004]을 참조하라.

<sup>6</sup> 대용량 말뭉치에서 추출된 데이터에서 아라비아 숫자와 기호 각각 최대 빈도 클래스 (Maximum Frequent Class) 결정법에 따라 구해진 당연값을 보여준다.

<sup>7</sup> 서버용 TAN 시스템을 먼저 구현하고 임베디드 시스템 탑재를 위해 최적화하였다. mini-TAN 구현에 있어 이러한 구현 절차를 따름으로써 필요에 따라 각 모듈 및 각종 사전을 포함하고 제거할 수 있어 더 높은 정확도를 유지할 수 있었다.

아라비아안 숫자표현의 문자화 정확도\*를 측정하였다. 실험 결과를 <표 6>으로 요약하여 정리하였다.

<표 6 서버용 시스템과 mini-TAN의 성능 비교>

성능 (측정 단위)	서버용 학습 기반	규칙 기반	mini-TAN 규칙 기반
메모리 사용량	104 MB	20 MB	1.5 MB
처리 속도	11.11 어절/초	30.30 어절/초	104.16 어절/초
정확도 (%)	97.29	99.1	98.2

<표 6>에서 보는 바와 같이, mini-TAN 은 서버용 학습 시스템에 비해 98.6%, 그리고 서버용 규칙 기반 시스템에 비해 92.5%의 메모리 사용량을 줄였다. 또한 아라비아안 숫자의 문자화 처리 속도를 크게 증가시켰는데, 서버용 학습 시스템에 비해 속도가 10.5 배 증가했고 규칙 기반 시스템에 비해서도 3.5 배의 속도 향상을 보였다. 이에 반해 mini-TAN 은 성능에 있어서도 심각한 성능 손실을 보이지 않아 서버용 규칙 기반 시스템에 비해 숫자의 문자화 처리 정확도가 0.9%가 낮았고, 서버용 학습 시스템에 비해서는 정확도가 오히려 0.91% 높았다.

mini-TAN 의 성능에 대해 좀 더 신뢰할 수 있는 평가를 하기 위해, 5 개의 실험 말뭉치를 추출하여, 국내 상용 제품으로 개발된 "Voiceware" TTS 시스템과 "CoreVoice" TTS 시스템의 아라비아안 숫자표현의 처리 정확도를 비교하였다.

<표 7>에서는 5 개 실험 말뭉치에 대한 mini-TAN 과 두 개 상용 TTS 시스템의 정확도 결과를 보여준다.

<표 7 mini-TAN 과 상용 TTS 제품의 정확도 비교>

실험 말뭉치 시스템	Set1	Set2	Set3	Set4	Set5
mini-TAN	97.7	96.9	98.3	97.6	98.2
VoiceWare	86.1	79.4	91.7	83.8	87.8
CoreVoice	88.7	78.8	87.8	82.9	87.1

실험 결과는 mini-TAN 의 성능이 "VoiceWare" 제품과 "CoreVoice" 제품의 성능"을 정확도 면에서 5.2~19.5% 증가함을 보여준다. 또한, 두 개 상용 제품과는 달리 mini-TAN 이 5 개 실험 말뭉치의 소스나 추출 방법에 상관없이 일관적인 정확도를 보여줌을 알 수 있다. mini-TAN 의 구체적인 성능을 분석하기 위해, 사용된 규칙의 적용 범위와 정확도를 측정하였으며, 이를 위해 실험말뭉치 'Set 5'가 사용되었다. 각 처리 단계의 규칙의 적용에 의해 문자화된 ANEs 와 적용된 규칙을 통해 정확하게 문자화된 ANEs 의 수를 어절 단위로 계산하였으며 결과는 <표 8>에 제시되었다.

<표 8 규칙의 적용 범위와 정확도>

적용 규칙	문자화된 어절 수 (A)	바르게 문자화된 어절 수 (B)	정확도 (%) C = B/A
고유명사 및 전문용어문자화	18	17	94.4
패턴 규칙	2,451	2,438	99.4
휴리스틱 규칙	160	160	100
문맥 규칙	6,654	6,639	99.7
RAN 당연값과 RTS 당연값	717	567	79.1
합계	10,000	9,821	98.2

<표 8>과 <표 9>에 제시된 결과에서 알 수 있듯이, mini-TAN 의 성능이 아주 뛰어나며, mini-TAN 의 구현을 위해 적용된 순차적 규칙이 매우 높은 정확도를 보인다.

\* 아라비아안 숫자표현을 문자화하여 주석을 달아놓은 말뭉치가 없으므로 저자에 의해 각 시스템이 생성하는 발음결과 mini-TAN 의 음소열은 저자에 의해 수작업으로 평가되었다.  
"Voiceware"와 "CoreVoice" 제품은 한국어 TTS 시스템 중 가장 성능이 뛰어난 것으로 평가된다.

그럼에도 불구하고 mini-TAN 에는 몇 가지 해결해야 할 문제가 남아있다. RAN 당연값과 RTS 당연값 적용이 다른 규칙의 적용만큼 정확하지 않다. 고유명사나 전문용어에 포함된 ANEs 는 인식이 쉽지 않으며, 이러한 경우의 ANEs 가 선택하는 RAN 역시 일관적이지 않다. 대부분 고유명사에 포함되는 이름, 두문자어, 약어 처리의 정확도가 70.7~74.1%에 불과하여[2] 음성 처리 및 자연언어처리에서 고유명사의 인식과 처리는 아직 해결하기 매우 힘든 문제이다.

5 결론 및 향후 연구

본 논문에서 아라비아안 숫자표현 (ANEs)의 의미와 발음의 중의성을 효과적으로 해소할 수 있고, 임베디드 시스템을 위해 최적화된 아라비아안 숫자의 문자화 시스템 mini-TAN 을 제안하였다. 숫자의 문자화 규칙을 설정하기 위해 문맥 정보, 패턴 자질 및 휴리스틱 정보가 사용되었다. 그리고 (1) 문자화 시스템으로부터 형태소 분석 모듈의 분리, (2) 어휘 사전 규모의 축소, (3) 문맥 사전으로부터 인명과 지명 사전의 제거를 통해 mini-TAN 은 시스템의 정확도의 심각한 손실 없이 메모리 사용량과 숫자표현의 문자화 처리 시간을 매우 크게 줄일 수 있어, 임베디드 시스템에 최적화되었다. 실험을 통해 96.9~98.3%의 숫자표현의 문자화 정확도를 얻었으며, 기존의 두 개 상용 TTS 제품보다 5.2~19.5% 높은 정확도를 보임을 알 수 있었다. 그러나 mini-TAN 은 고유명사나 전문용어에 포함된 ANEs 의 처리가 정확하지 않고, 단순화된 형태소 분석의 오류에 의한 오류를 해결할 수 없는 문제를 가졌으며, 문맥 정보가 부족하여 발생하는 중의성을 해결할 수 없었다. 고유명사의 인식과 처리 문제의 해결을 위해, 본 논문에서는 고유명사 인식 및 고유명사에 포함된 숫자표현의 문자화를 위한 추론 규칙을 제안하였으나 더 정교화된 추론 규칙에 대한 연구가 이후 진행되어야 할 것으로 본다. 또한, 모든 자연언어처리 및 음성 처리의 기반이 되는 형태소 분석기의 경량화 역시 계속 진행되어야 할 연구 분야이며, 한국어 음성 합성을 위한 텍스트 전처리 단계에서 아라비아안 숫자 및 문장 기호 외에 본 시스템을 로마 알파벳, 한자 등 모든 비-한글 문자 및 기호를 문자화할 수 있는 통합 시스템으로 확장할 수 있는 방안이 연구되어야 한다.

Acknowledgement

본 연구는 산업자원부 신기술 실용화 기술 개발 사업 (프로젝트 번호: 10009812)에 의해 지원을 받았음을 밝힌다.

참고 문헌

- [1] Frostad, Kathy "The State of Embedded Speech", *Speech Technology*, April 15, 2003, <http://www.speechtechmag.com/pub/industry/1926-1.html> (referred to on July 23, 2004).
- [2] Tetschner, Walt "Text-to-Speech - Naturalness and Accuracy", *ASR News*, July, 2003 <http://www.asrnews.com/tsapp/tsapp11.htm> (referred to on June 7, 2004).
- [3] 정영임(2004), "한국어 음성합성을 위한 숫자 읽기 시스템 구현", 석사학위논문, 부산대학교, 2004.
- [4] Yoon, Aesun, Kwon, Hyuk-Chul, Lee, Man-Hyeong (2003), "An Automatic Transcription System for Arabic Numerals in Korean", *Proceedings of 2003 International Conference on Natural Language Processing and Knowledge Engineering*, pp.221-226.
- [5] 윤애선, 권혁철 (2003), "한국어 텍스트에 사용된 이음표의 문자 전사" *언어와 정보*, Vol.7, pp.23-40.
- [6] 윤애선, 정영임, 권혁철 (2004) "아라비아안 숫자를 동반한 중의적 기호의 자동전사: 은점, 쌍점, 빗금을 중심으로", *언어와 정보*, Vol.8, pp.117-136.
- [7] Deng, Li, Wang, Kuansan, Acero, Alex, Hon, Hsiao-Wuen, Droppo, Jasha, Boulis, Constantinos, Wang, Ye-Yi, Jacoby, Derek, Mahajan, Milind, Chelba, Ciprian, Huang, Xuedong (2002), "Distributed Speech Processing in Mipad's Multimodal User Interface", *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING*, Vol.10, No.8, pp.605-619.
- [8] Hoffmann, R., Jokisch, O., Hirschfeld, D., Strecha, G., Kruschke, H., Kordon, U., Koloska, U.(2003), "A MULTILINGUAL TTS SYSTEM WITH LESS THAN 1 MBYTE FOOTPRINT FOR EMBEDDED APPLICATIONS", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.1, pp.1-532-535.
- [9] Ling, Zhen-Hua, Hu, Yu, Shuang, Zhi-Wei, Wang, Ren-Hua (2004), "Compression of Speech Database by Feature Separation and Pattern Clustering Using STRAIGHT", *Proceedings of INTERSPEECH2004-ICSLP Vol.II*, pp.1201-1204.