

단위 또는 약어의 의미에 맞는 풀 네임(full name) 음성 출력 방법에 관한 연구

박안나*, 손병준*
*삼성전자주식회사

e-mail : anna1.park@samsung.co.kr

A Study on Smart Text Reader for converting Text through TTS

An na Park*, Byoung-jun Son*

* Digital Multimedia Communications R&D Center, Samsung Electronics

요 약

현재까지의 음성 출력 시스템은 텍스트를 있는 그대로 읽어 주는 것에 불과했다. 단위, 약어의 경우 알파벳을 그대로 읽어 주게 되어 그 본래의 의미를 제대로 파악하기 어려웠다. 본 연구에서는 단위나 약어의 본래의 의미를 찾아서 풀어서 음성 변환해 주는 방법을 제안함으로써 시각 장애인에게도 텍스트의 정확한 정보를 전달할 수 있다는 장점이 있다.

1. 서론

최근 음성 합성 기술은 각 분야에서 활발히 응용되고 있으며, 많은 곳에서 사용되고 있다.

하지만 기존 TTS 엔진에는 단위나 약어에 대해서 본래의 의미를 찾아서 풀어서 설명해 주는 방식이 존재하지 않았다.

이에 따라 일반인은 물론이고, 특히 시각 장애인들은 TTS 를 통해 듣는 것만으로 단위나 약어의 본래의 의미를 파악하는 데 한계가 있다.

따라서 본 논문에서는 기존의 음성 합성 기술에서 결여된 단위, 약어의 본래 의미 복원에 대한 방법을 제공하며, 사용자들에게 사용성과 편의성을 제공할 수 있도록 단위, 약어 Original Text 음성 합성 서비스를 설계하고 구현하였다.

이는 단위나 약어를 검출하여 문맥에 맞도록 본래 의미로 변환하여 전달함으로써 사용자의 편의성을 향상시키고, 다의어의 중의성을 해소시켜 정확한 의미를 전달할 수 있다는 장점이 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 기존 TTS 기술과 분석, 3 장에서는 CSR(Common-sense Reasoning)에 대해 알아보고, 4 장에서는 본 논문의 알고리즘에 대해 설명하고 실험 결과를 첨부하였다. 마지막으로 5 장에서는 결론 및 향후 연구에 대해 기술하였다.

2. 기존 TTS 기술과 분석

TTS 란 글자, 문장, 숫자, 기호 등을 사람이 일반적으로 발성하는 음성으로 변환하는 것을 말한다. TTS 기술은 시각 의존적인 작업과 병행이 가능하다는 장

점이 있다. 또한, 수시로 변할 수 있는 정보를 전달하기에 가장 쉽고 간편한 방법이며, 시각 장애인을 위한 매우 효율적인 정보 제공 수단이기도 하다.

하지만 기존 TTS 에서는 단위나 약어 등을 글자 그대로 읽어 주는 경우가 많아서 본래의 의미를 전달하는 데 어려움이 있다.

3. CSR(Common-Sense Reasoning)

Common-Sense Reasoning 이란 이미 알고 있는 상식적인 여러 지식 및 규칙을 이용하여 어떤 사실 또는 표현을 유도해 내는 것을 말한다. 사람들이 대화를 하거나 글을 쓸 때, 혹은 글을 읽을 때 글이나 대화 자체에서는 나타나지 않는 지식을 사용하는 경우가 많은데, 컴퓨터는 이러한 배경지식들을 가지고 있지 않기 때문에 여러 가지 문제가 발생할 수 있다.

따라서 ConceptNet 과 같은 전세계의 다양한 상식적 지식 기반으로 구성된 시맨틱 네트워크 툴킷을 이용하여 상식적 추론을 할 수 있다.

이를 활용하여 문맥에 적합한 뜻을 선택함으로써 단위나 약어의 중의성을 해소시켜 줄 수 있는 수단이 된다.

구체적으로 중의어가 존재할 경우, 각각의 단어에 대한 Context 정보들을 중의어 DB 에 저장하여 둔다. 이 때 ContepNet 을 활용하여 Context 정보를 추출해 낸다. 이후 중의어의 전후 문맥에 나타나는 단어들을 분석하여 수학적 연산을 통하여 관계성을 측정해 낸 후, 중의어의 본래 의미를 매칭시킴으로써 중의성을 해소시킬 수 있다.

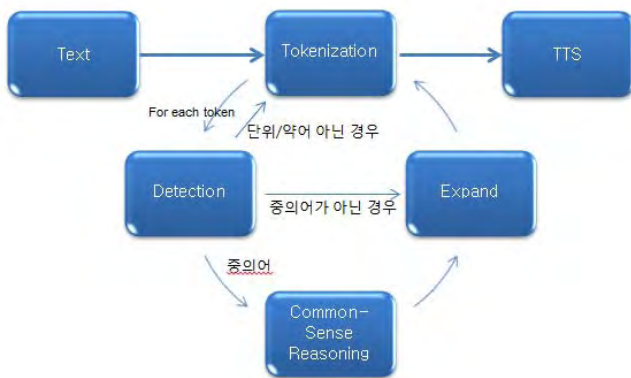
4. 본 연구의 동작 원리

본 연구는 그림 1 에 표현된 흐름도와 같이 텍스트를 음성으로 출력하는 시스템에서 단위와 약어를 본래의 의미에 맞도록 변환하여 음성 출력시켜 주는 방법에 관한 연구이다.

먼저, 텍스트에서 단위나 약어를 검출하는 모듈이 필요하다. 이를 위해 텍스트를 토큰 단위로 분리시키고, 각 토큰이 단위나 약어를 포함하고 있는지 여부를 특정한 규칙에 의해 확인해낸다.

검출된 토큰이 단위나 약어를 포함하고 있는 경우에 사전 DB 에 접근하여 검색한다. 이 때 중의어가 포함되어 있을 경우 앞서 설명한 Common-Sense Reasoning 기법을 활용하여 가장 적합한 의미를 검출해 준다. 검출 결과를 활용하여 본래의 의미로 변환시켜 음성으로 출력해 준다.

사전 DB 의 구축 시 단어 리스트 선정에는 인터넷 기반 사용자 통계를 활용하였고, 문맥적 연관 Graph 인 ConceptNet 를 활용하여 DB 의 Context 정보를 구성하였다.



(그림 1) 동작 Flow Chart

본 연구의 결과와 기존의 TTS 들과의 단위, 통화, 약어에 대한 비교 분석 결과는 다음과 같다.

단위/약어예	Samsung TTS (S5)	Google TTS (S5)	A#TTS (S5)	본 논문
1KRW	1-K-R-W	1 Korean Won	1-K-R-W	1 Korean Won
1AUD	1 Australian dollar	1 Australian dollar	1 aud	1 Australian Dollar
1VND	1-V-N-D	1 Vietmanise Dong	1-V-N-D	1 Viet Nam Dong
1USD	1 US Dollar	1 United States Dollar	1 US Dollar	1 United States Dollar
\$2	2 dollars	2 dollars	2 dollars	2 dollars
\$1M	1 million dollar	1 million dollars	1 million dollars	1 million dollars
\$1B	1 billion dollar	1 billion dollars		1 billion dollars
\$1T	1 trillion dollar	1 trillion dollars		1 trillion dollars
€ 7	7 Euros	7 Euros	7 Euros	7 Euros
€1M	1 million Euro	1 million Euros	1 million Euros	1 million Euros
€1B	1 billion Euro	1 billion Euros	1 billion Euros	1 billion Euros
¥2	2 yen	2 yen	2 yen	2 yens
¥1M	1 million yen	1 million yen	1 million yen	1 million yen

<표 1> 기존 TTS 와 비교 실험 : 단위

단위/약어예	Samsung TTS (S5)	Google TTS (S5)	A#TTS (S5)	본 논문
1m	1 meter	1 meter	1 m	1 meter
2m	2 meters	2 meters	2 m	2 meters
1m/s	1 meter per second	1 meter per second	1 m slash s	1 meter per second
2m/s	2 meters per second	2 meters per second	2 m slash s	2 meters per second
2rad	2 rad	2 radians	2 rad	2 radians
2cd	2 candela	2 candelas	2 c-d	2 candelas
10mV	10 m-V	10 millivolts	10 m-V	10 millivolts
7000km	7000 kilometers	7000 kilometers	7000 kilometers	7000 kilometers
1π	1	1 pi	1 pi	1 pi
2°C	2	2	2 Degrees Celsius	2 Degrees Celsius
3°F	3	3	3 Degrees Fahrenheit	3 Degrees Fahrenheit
2A	2 A	2 A	2 A	2 Ampare
1W	1 W	1 W	1 W	1 Watt
2Pa	2 Pascal	2 pa	2 pa	2 Pascals
2ft	2 feet	2 feet	2 feet	2 feet
1J	1 J	1 J	1 Joule	1 Joule

<표 2> 기존 TTS 와 비교 실험 : 통화

단위/약어예	Samsung TTS (S5)	Google TTS (S5)	A#TTS (S5)	본 논문
USA	U-S-A	U-S-A	U-S-A	United States of America
FAQ	F-A-Q	faq	Frequently Asked Questions	Frequently Asked Questions
Q&A	Q-and-A	Q-and-A	Q-ampersand-A	question and answer
ASAP	A-S-A-P	A-S-A-P	Asap	as soon as possible
PLZ	P-L-Z	P-L-Z	P-L-Z	Please
CPA	C-P-A	C-P-A	C-P-A	certified public accountant
UA370 flight	U-A 370 flight	U-A 370 flight	U-A 370 flight	United Airlines 370 flight
CA	C-A	C-A	C-A	California
Fig 1	fig 1	fig 1	figure 1	figure 1
Fig.1	figure 1	fig 1	fig period 1	figure 1
2MPH	2 M-P-H	2 M-P-H	2 Miles per Hour	2 Miles per Hour
BTW	By the way	B-T-W	By the way	By the way
LOL	L-O-L	lol	lol	Laugh out loud
AFAIK	A-F-A-I-K	A-F-A-I-K	afaik	As far as I know

<표 3> 기존 TTS 와 비교 실험 : 약어

5. 결론 및 향후 연구

기존의 TTS 엔진에서는 단위, 약어에 대해서 별도로 처리해 주는 모듈이 없어, 텍스트를 글자 그대로 읽어 주었다. 이로 인해 단위, 약어에 대한 본래의 의미를 파악하기 어렵다는 한계가 있었다.

본 연구에서는 단위, 약어를 검출하는 알고리즘을 제안하고, 단위, 약어를 검출하여 규칙에 맞게 각각의 본래의 텍스트로 대체하여 원래의 의미대로 음성 합성하여 알려줄 수 있도록 구현하였다.

본 연구의 방법에 의하면, 일반적으로 TTS 를 이용하는 경우는 물론이고, 직접 텍스트를 보기 힘든 운전 중이거나 요리 중인 상황 뿐만 아니라, 특히 시각 장애인에게 단위, 약어 등의 전문 지식을 본래의 의미에 맞도록 읽어 주는 것은 매우 유용한 기능이 될 것이다.

음성 합성과 재생을 동시에 수행하기 위해, 자바의 쓰레드 기법을 사용하여 사용자의 TTS 요청에 대한 응답 시간을 최소화하고, DB 접근 시간을 최소화하였다.

또한 문장을 의미 있는 단위인 토큰별로 분리하여 각각의 분산 병렬 처리 과정을 통해 처리 및 음성 합성 시간을 최소화하여 긴 문장이 입력된 경우에도 사용자가 많은 시간을 기다리지 않고, 빠르게 병렬로 응답을 받을 수 있도록 구조화하여 최적화하였다. 이에 따라 기존 TTS 대비 처리 속도 저하의 문제는 없었다.

이로 인해 사용자에게 보다 정확한 단위, 약어의

의미 정보를 제공하는 데 있어서 DB 접근 및 음성 합성 처리 속도도 최적화하여 제공할 수 있다는 데 의미가 있다고 판단된다.

향후 연구 과제로는 단위, 약어의 중의어 해소에 대한 방법을 보완하여 보다 정확한 정보 제공을 할 수 있도록 보완하고자 한다.

참고문헌

- [1] 이희만 "TTS 적용을 위한 음성합성엔진"
- [2] 이희만 "홈페이지의 음성합성 시스템"
- [3] 양승정, 이종태 "자연스러운 음성 합성을 위한 음성합성기술 개발"
- [4] 이종석, 박기태, 이준우 "시각 및 언어장애인을 위한 음성합성 기술의 현황"
- [5] 이희연, 홍기형 "시각장애이용 음성합성기에 대한 사용자 요구분석"
- [6] Hugo Liu, et al., "ConceptNet: A Practical Commonsense Reasoning Toolkit," Springer BT Technology Journal, vol. 22, no. 4, pp. 211-226, 2004.
- [7] Hugo Liu, et al., "A Model of Textual Affect Sensing using Real-World Knowledge," Proc. of the 8th Int. Conf.