

텔레매틱스 환경에서 음성 콘텐츠 제공 시스템

김철수[○] 김종익 윤대섭

한국전자통신연구원

{chulsu1[○], jongik.kim, eyetracker }@etri.re.kr

Voice Contents Providing System in Telematics Environment

Chul-Su Kim[○], Jongik Kim and Daesub Yoon
Electronics and Telecommunications Research Institute

요 약

텔레매틱스는 통신장치와 컴퓨터를 이용하여 경로안내 서비스, 주변정보안내, 이메일 서비스 등과 같은 다양한 서비스를 차량 사용자에게 제공하는 통합기술이다. 운전자가 차량 운행 중에 작은 장치에 데이터를 입력하거나 디스플레이된 정보를 읽는 것은 무척 위험하다. 그래서 우리는 운전자의 손은 운전대에 그리고 눈은 도로 및 주변상황을 확인할 수 있도록 하기 위해서 개방형 텔레매틱스 환경에서 음성기반의 시스템을 개발할 수 있는 시스템 구조를 제안하였다. 그리고 제안된 시스템 구조를 이용하여 개방형 텔레매틱스 시스템을 위한 텔레매틱스 게이트웨이와 프레임워크를 기반으로 텔레매틱스용 음성기반 이메일 서비스를 개발하였다.

1 서론

텔레매틱스는 통신이 가능한 단말 장치를 이용해서 경로안내, 주변정보안내, 이메일과 같은 다양한 서비스를 제공하는 통합기술이다. 그러나 텔레매틱스 서비스를 사용하기 위해서 운전자는 텔레매틱스 단말기와 상호작용이 필요한데 이는 운전중인 차량 환경에서는 무척 위험한 행동이다. 이런 안전 문제를 해결하기 위해서는, 운전자는 텔레매틱스 단말기를 이용하는 동안에도 손은 운전대를 잡고 시선은 전방을 바라보는 등 운전 집중할 수 있어야 한다. 그래서 운전자는 음성을 이용해서 명령을 내리고 텔레매틱스 단말기는 서비스 정보를 음성을 이용하여 제공할 수 있어야 한다. 본 논문에서, 우리는 차량 사용자가 텔레매틱스 단말기를 통해 음성을 이용하는 텔레매틱스 시스템을 개발하였다.

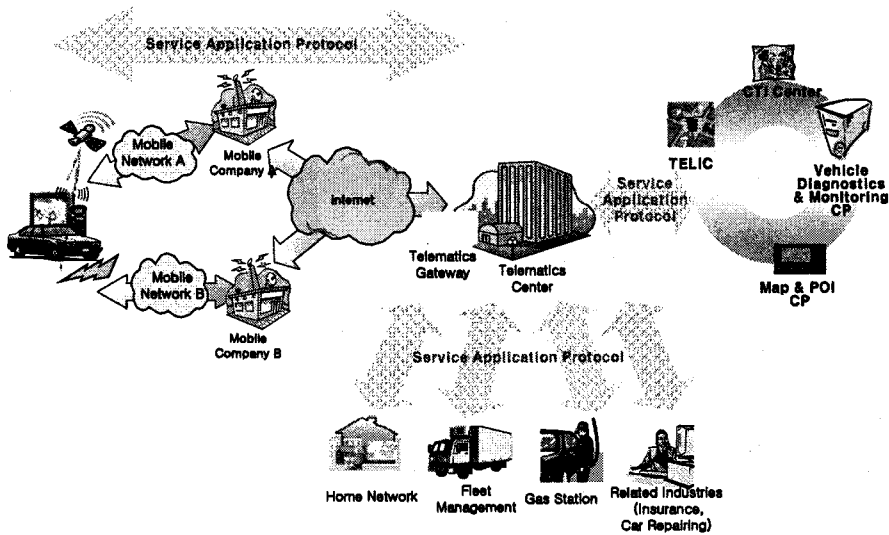
텔레매틱스 환경에서 음성서비스를 제공하는 시스템은 음성관련 기술을 제공하는 위치에 따라 크게 두 가지로 구분이 된다. 첫째로 단말기에서 음성 인식 및 합성을 처리하는 경우이다. 운전자가 음성을 이용하여 명령을 내리면 차량내의 단말 장치는 자체의 음성인식 엔진을 이용해서 음성을 명령어로 변환 한다. 운전자에게 정보를 전달할 때는 단말장치는 음성합성엔진을 이용해서 정보를 음성으로 변환하여 운전자에게 들려준다. 두 번째는 서버 측 시스템에서 음성을 인식 및 합성을 처리하는 경우이다. 서버 측 음성처리를 위해

서 CTI(computer-telephony integration) [1]시스템을 이용할 수 있다. 운전자가 음성을 이용해서 명령을 내리면 음성은 전화선을 통해서 CTI 센터에 전송이 되고 CTI 센터는 음성인식엔진 또는 오퍼레이터에게 음성을 보내서 음성을 명령으로 변환한다. 운전자에게 전송될 정보는 CIT 센터의 음성합성엔진 또는 오퍼레이터를 통해 음성으로 만들어져서 사용자의 단말기로 전달된다.

차량용 텔레매틱스 단말기의 낮은 성능 때문에 첫 번째 접근법은 음성 인식에서 실패율이 높고 합성된 음성은 서비스에서 활용하기에는 질이 떨어진다. 그래서 음성인식 및 합성은 단말 측에서 처리되는 것보다 서버 측에서 처리하는 것을 선호하고 있다. 그러나 서버 측 음성처리 접근법은 복잡하고 비용이 많이 드는 CTI 기반구조가 필요하다.

본 논문에서 우리는 서버측에서 음성인식과 합성을 처리하여 음성기반의 서비스를 제공하는 새로운 텔레매틱스 시스템을 개발하였다. 개발된 시스템은 음성채널은 사용하지 않고 데이터채널만 사용하기 때문에 복잡한 CTI 기반구조가 없어도 고품질의 음성서비스 시스템을 구현할 수 있다.

2장에서는 기존의 연구들을 소개하고 3장에서는 음성컨텐츠를 제공하는 시스템을 설명하고 4장에서는 본 시스템을 이용하여 개발한 이메일 서비스를 구현한 모습을 보여주고 끝으로 5장에서 결론을 맺는다.



(그림 1) 개방형 텔레매틱스 시스템구조

2 관련연구

2.1 개방형 텔레매틱스 시스템

서비스제공을 위해 모바일 네트워크를 사용하는 대부분의 텔레매틱스 시스템은 WAP(Wireless Application Protocol) 게이트웨이를 사용하여 무선상의 단말장치와 유선상의 서버 사이에서 데이터를 교환한다. 이 경우 서비스는 게이트웨이를 제공하는 네트워크에 종속적이 된다. 이런 종속적인 문제를 해결하기 위해 개방형 텔레매틱스 시스템은 [1] WAP 게이트웨이 역할을 하는 텔레매틱스 게이트웨이를 그림 1에서 처럼 텔레매틱스 서비스센터 내부에 포함시킬 것을 제안한다.

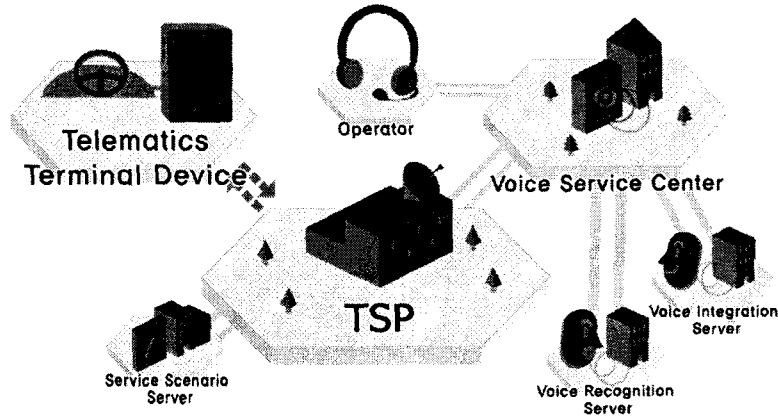
대부분의 텔레매틱스 서비스 제공자들은 서비스를 위한 콘텐츠를 직접 만들고 관리하고 있다. 그러나 이런 작업들은 작은 회사들에게 큰 부담이 되고 텔레매틱스 분야에서 새로운 사업을 시작하는데 장애물이 되고 있다. 그래서 개방형 텔레매틱스 시스템은 인터넷 상에 흩어져 있는 다양한 콘텐츠를 이용해서 텔레매틱스 서비스를 구성할 수 있는 기술을 제안한다. 본 논문에서는 개발된 시스템은 개방형 텔레매틱스 시스템을 기반으로하여 서비스 제공자들이 음성 서비스를 위한 데이터 및 처리된 데이터를 하나의 콘텐츠로 보고 외부의 콘텐츠를 이용하듯이 음성 콘텐츠를 이용하여 음성기반의 서비스를 개발할 수 있도록 한다. 그래서 서비스 제공자는 자신의 회사에 음성처리 기술 및 시스템이 없어도 음성기반의 텔레매틱스 서비스들을 제공할 수 있게 된다.

2.2 텔레매틱스&음성관련 연구

차량 운전 중에 안전성은 가장 중요한 요소이므로 텔레매틱스 서비스는 운전자의 주위를 산만하게 하지 말아야 한다. Wickens의 multiple resource theory[4]에 따르면 손을 사용하지 않는 텔레매틱스 서비스 사용이 손을 이용하는 서비스의 이용보다 안전하다는 것을 알 수 있다. 그래서 텔레매틱스 환경에서 운전자의 주위를 뺏지 않고 안전성을 높이기 위해 음성기반의 사용자 인터페이스를 사용하는 것이 필요하다.

차량에서 음성기반의 서비스를 제공하기 위해서는 음성인식 성능과 인식률이 안정적이고 높아야 하는데 McMillan 과 Eggleston의 연구에 따르면[5] 기술적인 제한과 사용자의 높은 기대감 때문에 음성인식의 성능이 신뢰받지 못하고 있다. 이 문제를 해결하기 위해서 우리는 계산 능력이 떨어지는 단말측이 아닌 고성능의 서버에서 음성처리를 하도록 시스템을 개발하였다.

Wagnier의 연구[6]에 따르면 차량이 70 mph로 운행 중일 때 음성인식률은 95%에서 82%로 떨어진다고 한다. 우리는 음성인식률 100%를 보장하기 위해 음성 오퍼레이터를 음성 서비스 센터에 두고 음성인식 엔진에서 음성인식이 실패할 경우 오퍼레이터를 통해 음성을 명령으로 변환하고 추가적인 정보를 음성으로 제공할 수 있도록 시스템을 개발하였다. 운전자의 입장에서 어떤 환경에서도 인식률이 100%인 음성 서비스를 받을 수 있게 되는 것이다.



(그림 2) 음성기반 텔레매틱스 서비스 시스템

3 음성기반 텔레매틱스 서비스 시스템 개발

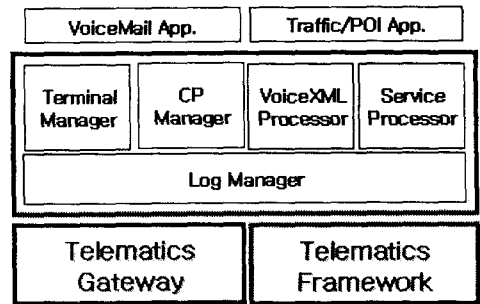
본 연구를 통해 개발한 텔레매틱스 서비스 시스템은 세가지 특징을 가진다. 첫째, 대부분의 음성 기반의 서비스 시스템은 음성처리 엔진을 내부에 장착하여 사용하지만 제안된 시스템은 그림 2에서처럼 음성처리 시스템을 외부에 독립적으로 둔다. 그래서 서비스 제공자가 음성처리 관련 모든 기반구조를 구축하지 않고도 음성서비스를 제공할 수 있도록 한다. 둘째, 음성 서비스 센터는 오퍼레이터를 두어서 음성인식엔진을 통해 인식이 실패한 경우 오퍼레이터가 듣고 명령을 전달하여 100%의 인식률을 보장할 수 있게 된다. 셋째, 제안된 시스템은 음성기반의 서비스 시스템을 개발하기 위해 VoiceXML 기술을 사용한다. 서비스 시나리오와 서비스 처리기가 분리되어 서비스 내용이 변경되면 간단히 XML 형태의 서비스 시나리오만 수정하여 음성 서비스를 제공할 수 있다.

본 장에서는 텔레매틱스 환경에서 음성 콘텐츠를 제공하는 시스템의 각 구성요소에 대해서 설명한다.

3.1 텔레매틱스 서비스 프로바이더

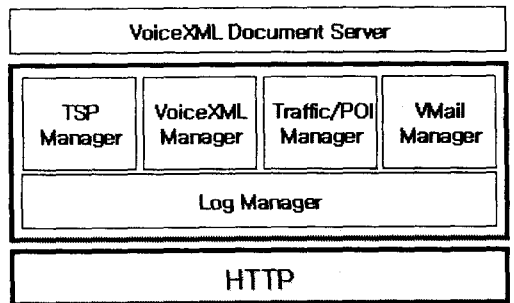
텔레매틱스 서비스 제공 서버는 텔레매틱스 게이트웨이를 통해 무선망에 있는 단말기와 연결하여 사용자에게 서비스를 제공한다. 또한 텔레매틱스 프레임워크 기술을 이용해 외부의 콘텐츠 제공자와 연결하여 서비스를 구성하기 위한 데이터를 받아온다.

그림 3에서 VoiceMail App.와 Traffic/POI App.는 각자 분리되어 운영되지만 단말기 관리, 콘텐츠 제공자 관리 VoiceXML 처리, 서비스 처리, 로그관리 등을 위해서 공통의 모듈을 사용한다.



(그림 3) 서비스 제공 서버

3.2 서비스 시나리오 서버



(그림 4) 서비스 시나리오 서버

서비스 시나리오 서버는 서비스 제공 서버에게 받은 요청을 처리해서 서비스 시나리오인 VoiceXML을 생성해서 응답으로 돌려준다. 서비스 시나리오 서버는 VoiceXML 표준에서 document server의 역할을 수행한다.

시나리오 서버는 이메일 데이터와 주변정보 데이터와 같이 서비스와 관련된 데이터를 TSP 서버에게 받아

서 서비스를 제공하기 위한 다음 시나리오인 VoiceXML 문서를 생성해서 다시 TSP 서버에게 돌려준다.

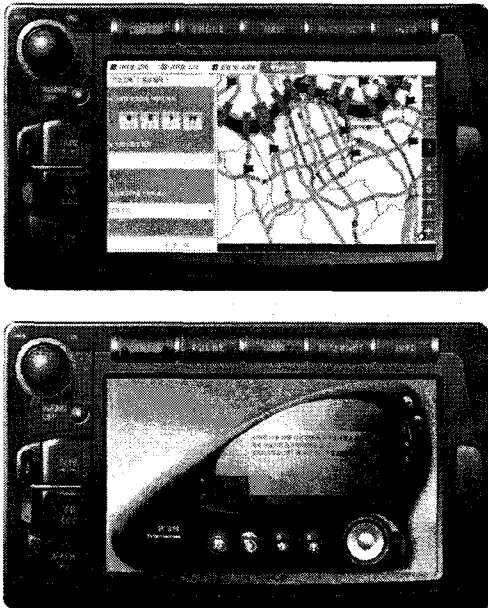
3.3 음성 서비스 센터

음성 서비스 센터는 음성인식 엔진과 음성합성 엔진을 이용하여 음성을 텍스트로 또는 텍스트를 음성으로 변환하는 서비스를 제공한다. 센터는 TSP 서버가 사용자에게 서비스로서 제공할 데이터를 음성으로 변환하고, 사용자가 TSP 서버에게 주는 음성명령을 텍스트 형태로 변환하여 준다.

사용자가 차내에서 음성으로 명령을 내리면 명령은 TSP 서버에게 전달되고 음성명령은 다시 TSP가 인식할 수 있는 명령으로 변환하도록 음성 서비스 센터에 보내진다. 음성 서비스 센터는 음성인식 엔진을 이용하여 음성을 텍스트로 변환하는데 만일 실패한다면 오퍼레이터가 음성을 듣고 해당 명령을 텍스트로 입력하게 된다.

4 시스템 구현

텔레매틱스 환경에서 음성 콘텐츠 제공 시스템을 이용하여 두 가지 텔레매틱스 서비스를 개발하였다. 교통정보/주변정보 제공 서비스를 개발하여 음성을 이용하여 교통정보와 주변의 유용한 정보를 요청하고 받을 수 있도록 하였다. 그리고 음성기반 이메일 서비스를 개발해서 사용자는 메일 목록 및 내용을 음성을 이용하여 요청하고 메일 내용을 청취할 수 있게 하였다.



(그림 5) 음성기반 텔레매틱스 서비스

5 결론

자동차 산업과 통신장치 및 컴퓨터의 발전으로 움직이는 차에서도 다양한 서비스를 사용할 수 있게 되어 텔레매틱스 분야는 눈부신 발전을 보이고 있다.

본 논문에서 음성 기반의 텔레매틱스를 제공할 수 있는 시스템을 소개하였다. 이 시스템은 개방형 텔레매틱스 시스템을 기반으로 개발되었고 이는 서비스 개발자가 외부의 콘텐츠를 활용하여 서비스를 개발할 수 있게 하였다. 우리는 음성기반 서비스에 필요한 음성 및 변환된 텍스트를 하나의 콘텐츠로 보고 서비스 제공자가 음성처리를 위한 기반시설 없이도 음성 서비스를 제공할 수 있는 방안을 제시하였다. 또한 음성 서비스 센터에 오퍼레이터를 두어서 텔레매틱스 환경에서 필요한 정확한 음성 처리 성능을 제공할 수 있도록 하였고 VoiceXML 기술을 사용하여 서비스 내용의 변경에도 쉽게 적용할 수 있도록 하였다.

참고문헌

- [1] W. Y. Han, O. C. Kwon, J. H. Park & J. Kang, A Gateway and Framework for Telematics Systems Independent on Mobile Networks, ETRI Journal, Vol. 27, No. 1, Feb. 2005, p106-109
- [2] Stein, A.C., Parseghian, Z., and Allen, R.W. (1987). A simulator study of the safety implications of cellular mobile phone use (Paper No. 405). Hawthorne, CA: Systems Technology, Inc.
- [3] Zwahlen, H.T., Adams, C.C., Jr., and Schwartz, P.J. (1988). Safety aspects of cellular telephones in automobiles (Paper No. 88058). Proceedings of the ISATA Conference, Florence, Italy.
- [4] Wickens, C.D. (1984) Processing resources in attention. In R. Parasuraman and D.R. Davies (Eds.), Varieties of Attention. London: Academic Press.
- [5] McMillan, G.R. and Eggleston, R.G. (1997). Nonconventional controls. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics. New York: John Wiley & Sons, Inc.
- [6] Wargnier, J. (2002). Considerations for robust speech recognition and sound quality for automotive handsfree kits. 21st Annual Conference of the Applied Voice Input/Output Society (AVIOS) pp.339-348
- [7] Asatani, K. (1998). Standardization on Multimedia Communications: Computer-Telephony-Integration-Related Issues, IEEE Communications Magazine pp.105-109