

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2013-0011002 (43) 공개일자 2013년01월30일
(51) 국제특허분류(Int. Cl.) G10L 15/06 (2006.01) G10L 17/00 (2006.01) (21) 출원번호 10-2011-0071863 (22) 출원일자 2011년07월20일 심사청구일자 2011년07월20일		(71) 출원인 세종대학교산학협력단 서울특별시 광진구 능동로 209 (군자동, 세종대학교) (72) 발명자 권순일 서울특별시 강남구 압구정로 201, 82동 504호 (압구정동, 현대아파트) (74) 대리인 특허법인엠에이피에스

전체 청구항 수 : 총 12 항

(54) 발명의 명칭 음성 신호를 이용하여 사용자의 성격을 판단하는 시스템 및 방법

(57) 요약

음성 신호를 이용하여 사용자의 성격을 판단하는 시스템 및 방법이 제공된다. 음성 신호를 이용하여 사용자의 성격을 판단하는 시스템은, 사용자의 음성 신호를 수신하는 음성 신호 수신부; 상기 수신된 음성 신호의 특성을 측정하는 신호 특성 측정부; 및 상기 측정된 음성 신호의 특성에 기초하여 상기 사용자의 성격을 판단하는 사용자 성격 판단부;를 포함하며, 상기 신호 특성 측정부는, 상기 음성 신호에 포함된 묵음 구간의 길이를 측정하고, 상기 사용자 성격 판단부는 상기 측정된 묵음 구간의 길이에 기초하여 상기 사용자의 성격을 판단한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 2011-0005334

부처명 교과부

연구사업명 기초연구사업 (일반연구자지원사업)

연구과제명 사용자 적응형 음성인터페이스를 위한 음성기반 사용자 성격인식

주관기관 세종대학교산학협력단

연구기간 2010.05.01 ~ 2013.04.30

특허청구의 범위

청구항 1

음성 신호를 이용하여 사용자의 성격을 판단하는 시스템에 있어서,
 사용자의 음성 신호를 수신하는 음성 신호 수신부;
 상기 수신된 음성 신호의 특성을 측정하는 신호 특성 측정부; 및
 상기 측정된 음성 신호의 특성에 기초하여 상기 사용자의 성격을 판단하는 사용자 성격 판단부;
 를 포함하며,
 상기 신호 특성 측정부는, 상기 음성 신호에 포함된 묵음 구간의 길이를 측정하며,
 상기 사용자 성격 판단부는, 상기 측정된 묵음 구간의 길이에 기초하여 상기 사용자의 성격을 판단하는 것인,
 사용자 성격 판단 시스템.

청구항 2

제 1 항에 있어서,
 상기 사용자 성격 판단부는, 상기 음성 신호의 전체 구간 중 상기 묵음 구간이 차지하는 비율에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 시스템.

청구항 3

제 1 항에 있어서,
 상기 사용자 성격 판단부는, 상기 음성 신호에서 기 설정된 수치 이상의 길이를 갖는 묵음 구간의 개수에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 시스템.

청구항 4

제 1 항에 있어서,
 상기 신호 특성 측정부는 상기 음성 신호로부터 사용자의 말의 빠르기를 측정하며, 상기 사용자 성격 판단부는 상기 측정된 사용자의 말의 빠르기에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 시스템.

청구항 5

제 1 항에 있어서,
 상기 신호 특성 측정부는 상기 음성 신호로부터 피치의 변화 값 및 신호 세기의 변화 값을 측정하고, 상기 사용자 성격 판단부는 상기 측정된 피치 변화 값 및 신호 세기 변화 값의 변화율에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 시스템.

청구항 6

제 1 항에 있어서,
 상기 사용자 성격 판단부는, 상기 사용자의 성격을 외향적인 성격과 내향적인 성격으로 구분하여 판단하는 것인, 사용자 성격 판단 시스템.

청구항 7

사용자 성격 판단 시스템이 사용자의 음성 신호에 기초하여 사용자의 성격을 판단하는 방법에 있어서,
 사용자의 음성 신호를 수신하는 단계;
 상기 수신된 음성 신호의 신호 특성을 측정하는 단계; 및

상기 측정된 특성에 기초하여 사용자의 성격을 판단하는 단계;

를 포함하며,

상기 신호 특성을 측정하는 단계는 상기 수신된 음성 신호의 신호 세기 및 피치의 변화율을 측정하며,

상기 사용자의 성격을 판단하는 단계는 상기 측정된 신호 세기 및 피치의 변화율에 기초하여 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법.

청구항 8

제 7 항에 있어서,

상기 신호 특성을 측정하는 단계는 상기 수신된 음성 신호에 포함된 묵음 구간의 길이를 측정하고,

상기 사용자의 성격을 판단하는 단계는, 상기 측정된 묵음 구간의 길이에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법.

청구항 9

제 8 항에 있어서,

상기 사용자의 성격을 판단하는 단계는, 상기 음성 신호의 전체 구간 중 상기 묵음 구간이 차지하는 비율에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법.

청구항 10

제 8 항에 있어서,

상기 사용자의 성격을 판단하는 단계는, 상기 음성 신호에서 기 설정된 수치 이상의 길이를 갖는 묵음 구간의 개수에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법.

청구항 11

제 7 항에 있어서,

상기 신호 특성을 측정하는 단계는, 상기 음성 신호로부터 사용자의 말의 빠르기를 측정하며, 상기 사용자 성격을 판단하는 단계는 상기 측정된 사용자의 말의 빠르기에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법.

청구항 12

제 7 항에 있어서,

상기 사용자의 성격을 판단하는 단계는, 상기 사용자의 성격을 외향적인 성격과 내향적인 성격으로 구분하여 판단하는 것인, 사용자 성격 판단 방법.

명세서

기술 분야

[0001] 본 발명은 음성 신호를 이용하여 사용자의 성격을 판단하는 시스템 및 방법에 관한 것으로서, 보다 상세하게는, 사용자의 음성 신호의 신호 특성에 기초하여 사용자의 성격을 판단하는 시스템 및 방법에 관한 것이다.

배경 기술

[0002] 최근 로봇 산업, 게임 산업, 그리고 교육 산업들과 관련하여 기능성 콘텐츠 시장은 매년 급성장을 이루고 있으며, 산업적으로뿐만 아니라 연구에 있어서도 도전 할만한 매력을 가지고 있는 분야이다. 하지만 기능성 콘텐츠에 활용되는 인터페이스와 관련된 원천과 응용 기술이 모두 취약하다. 이러한 난관을 극복하고 기능성 콘텐츠 신조류를 이용한 동반 성장을 위해서는 신개념 인터페이스 및 이를 적용한 콘텐츠 기획에 대한 집중적인 연구가 필요하다.

[0003] 수십 년간 음성 신호가 포함하고 있는 말을 문자로 바꾸는 일인 음성 인식에 관한 연구와 말이 포함하고 있는

의미를 분석하는 자연어 처리 분야의 연구가 꾸준히 지속되어 왔고, 앞으로도 발전의 여지가 많다. 음성 신호에는 언어의 의미적 요소뿐만 아니라 화자, 즉 말하는 사람의 감정 같은 심리상태나 성격 등이 포함되어 있다. 특히 현실 공간과 가상 공간을 연결하기 위하여 가상 공간에서 현실의 상황을 정보화하고 이를 활용하여 사용자 중심의 지능화된 서비스를 제공하는 기술에 있어서, 사용자의 취향이나 감성을 효율적이고 정확하게 인지하기 위해서 인간의 사고 및 감정 정보를 추출 및 처리하는 기술이 요구된다.

[0004] 한편, 선행 기술로서, 미국공개특허 제2008/0177540호에는 “Method and apparatus for recognizing reacting to user personality in accordance with speech recognition system”이라는 명칭의 발명이 개시되어 있는 바, 기설정된 단계의 질의에 대한 사용자의 답변 스피치를 형태학 및 구문론에 기초하여 분석하고 사용자의 성격을 판단하는 기술에 관한 것이다.

[0005] 하지만, 상술한 선행 기술은 사용자의 음성 신호의 의미를 해석하여 사용자의 성격을 판단하고 있어, 사용자 성격의 판단에 복잡한 알고리즘이 필요하며 많은 리소스 자원이 낭비되는 문제가 있었다.

[0006] 따라서, 사용자 음성 신호의 신호적 특성을 이용하여 효율적이면서도 정확하게 사용자의 성격을 판단할 수 있는 사용자 성격 판단 기술이 요구되고 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 일부 실시예는 사용자의 음성 신호에 포함된 언어적 정보를 사용하지 않고 사용자의 음성 신호의 신호적 특성을 이용하여 사용자의 성격을 판단할 수 있는, 사용자 성격 판단 시스템 및 방법을 제공한다.

[0008] 또한, 본 발명의 일 실시예는 사용자의 음성 신호에 포함된 묵음 구간의 길이를 이용하여 사용자의 성격을 판단할 수 있는, 사용자 성격 판단 시스템 및 방법을 제공한다.

과제의 해결 수단

[0009] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 제 1 측면은, 사용자의 음성 신호를 수신하는 음성 신호 수신부; 상기 수신된 음성 신호의 특성을 측정하는 신호 특성 측정부; 및 상기 측정된 음성 신호의 특성에 기초하여 상기 사용자의 성격을 판단하는 사용자 성격 판단부;를 포함하며, 상기 신호 특성 측정부는, 상기 음성 신호에 포함된 묵음 구간의 길이를 측정하고, 상기 사용자 성격 판단부는 상기 측정된 묵음 구간의 길이에 기초하여 상기 사용자의 성격을 판단하는 것인, 사용자 성격 판단 시스템을 제공할 수 있다.

[0010] 또한, 본 발명의 제 2 측면은, 사용자의 음성 신호를 수신하는 단계; 상기 수신된 음성 신호의 신호 특성을 측정하는 단계; 및 상기 측정된 특성에 기초하여 사용자의 성격을 판단하는 단계;를 포함하며, 상기 신호 특성을 측정하는 단계는 상기 수신된 음성 신호의 신호 세기 및 피치의 변화율을 측정하며, 상기 사용자의 성격을 판단하는 단계는 상기 측정된 신호 세기 및 피치의 변화율에 기초하여 사용자의 성격을 판단하는 것인, 사용자 성격 판단 방법을 제공할 수 있다.

발명의 효과

[0011] 전술한 본 발명의 과제 해결 수단에 의하면, 사용자의 음성 신호의 신호 특성을 이용하여 사용자의 음성을 해석하지 않더라도 효과적으로 사용자의 성격을 판단할 수 있다.

[0012] 또한, 전술한 본 발명의 과제 해결 수단에 의하면, 사용자의 음성 신호에 포함된 묵음 구간에 관한 정보를 이용하여 사용자의 성격을 보다 정확하게 판단할 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일 실시예에 따른, 음성 신호를 이용하여 사용자의 성격을 판단하는 시스템의 구성도이다.

도 2는 본 발명의 일 실시예에 따른 신호 특성 측정부의 세부 구성도이다.

도 3은 본 발명의 일 실시예에 따른 음성 신호를 이용한 사용자 성격 판단 방법의 세부 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 5-fold cross validation을 수행하여 예측력을 비교한 결과를 도시한 도면이다.

도 5는 본 발명의 일 실시예에 따른 실험 결과에서 가장 높은 성능을 보인 F11 feature set에서 다른 전체 음성 중 묵음 비율을 비교하여 도식한 도면이다.

도 6은 본 발명의 일 실시예에 따른 실험 결과 중 F12 feature set에서 다루어진 일정 길이 이상의 묵음 구간의 발견 횟수를 비교한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0014] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0015] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0016] 이하 첨부된 도면을 참고하여 본 발명을 상세히 설명하기로 한다.
- [0017] 도 1은 본 발명의 일 실시예에 따른, 음성 신호를 이용하여 사용자의 성격을 판단하는 시스템의 구성도이다.
- [0018] 도 1에 도시된 바와 같이, 본 발명의 일 실시예에 따른 음성 신호를 이용한 사용자 성격 판단 시스템은, 음성 신호 수신부(100), 신호 특성 측정부(200), 사용자 성격 판단부(300) 및 제어부(400)를 포함한다.
- [0019] 음성 신호 수신부(100)는 사용자의 음성 신호를 수신한다. 음성 신호 수신부(100)는 마이크 등의 장치를 통하여 사용자의 음성을 수신할 수도 있으나, 이에 제한되지 않으며, 외부 서버 또는 장치로부터 수집된 사용자의 음성 신호를 수신할 수도 있다.
- [0020] 또한, 음성 신호 수신부(100)는 사용자에게 질의를 제공하고 제공된 질의에 응답되는 사용자의 답변 음성을 수신할 수도 있다.
- [0021] 신호 특성 측정부(200)는 수신된 사용자의 음성 신호의 신호 특성을 측정한다. 신호 특성 측정부(200)는 사용자의 음성 신호의 언어적 특성(예를 들어, 음성의 의미)을 제외하고 비언어적인 특성을 측정할 수 있다. 신호 특성 측정부(200)는 사용자 음성 신호로부터 스피치 레이트(speech rate, 말의 빠르기), 피치의 변화 정도, 음성 신호의 세기 및 묵음 구간 중 적어도 하나 이상을 측정할 수 있다. 신호 특성 측정부(200)가 음성 신호의 특성을 측정하는 구체적인 방법에 관하여는 도2에서 보다 상세히 설명하기로 한다.
- [0022] 사용자 성격 판단부(300)는 측정된 사용자 음성 신호의 특성에 기초하여 사용자의 성격을 판단한다. 사용자 성격 판단부(300)는 사용자 음성 신호로부터 측정된 스피치 레이트(speech rate, 말의 빠르기), 피치의 변화 정도, 음성 신호의 세기 및 묵음 구간을 이용하여 사용자의 성격이 외향적인 성격인지 내향적인 성격인지를 판단할 수 있다.
- [0023] 사용자 성격 판단부(300)는 스피치 레이트 값에 기초하여 사용자의 말의 빠르기가 기 설정된 수치 이상인 경우에 외향적인 성격으로 판단할 수 있다.
- [0024] 또한, 사용자 성격 판단부(300)는 사용자 음성의 피치의 변화 정도에 기초하여 사용자의 성격을 판단할 수 있다. 이 경우, 사용자 성격 판단부(300)는 음성 피치의 절대 값을 고려하기 보다는 음성 피치의 평균값에 대한 상대적인 변화율을 이용하여 사용자의 성격을 판단할 수 있다. 예를 들어, 사용자 성격 판단부(300)는 음성 피치의 변화율이 평균 값에 비하여 기 설정된 수치 이상인 경우에 사용자가 외향적인 성격인 것으로 판단할 수 있다.
- [0025] 또한, 사용자 성격 판단부(300)는 사용자 음성의 신호 세기의 변화 정도에 기초하여 사용자의 성격을 판단할 수 있다. 이 경우, 사용자 성격 판단부(300)는 음성 신호 세기의 절대 값을 고려하기 보다는 음성 신호 세기의 평균값에 대한 상대적인 변화율을 이용하여 사용자의 성격을 판단할 수 있다. 예를 들어, 사용자 성격 판단부(300)는 음성 신호 세기의 변화율이 평균 값에 비하여 기 설정된 수치 이상인 경우에 사용자가 외향적인 성격인 것으로 판단할 수 있다.
- [0026] 또한, 사용자 성격 판단부(300)는 사용자 음성에 포함된 묵음 구간의 길이, 비율 및 장묵음(long silence

interval) 발견 횟수에 기초하여 사용자의 성격을 판단할 수 있다. 사용자 성격 판단부(300)는 음성 신호 중에서 묵음 구간의 길이가 기 설정된 시간 이상으로 오랫동안 지속되거나 그 편차가 기 설정된 수치 이상인 경우에 내향적인 성격으로 판단할 수 있다. 또한, 사용자 성격 판단부(300)는 전체 음성 중에서 묵음 구간이 차지하는 비율이 기 설정된 수치 이상인 경우에 내향적인 성격으로 판단할 수 있다. 또한, 사용자 성격 판단부(300)는 일정 시간 이상의 묵음 구간이 발생한 횟수가 기 설정된 수치 이상인 경우에 사용자의 성격을 판단할 수 있다. 또한, 사용자 성격 판단부(300)가 사용자의 성격을 판단하는 기준은 다양한 실험에 의해 결정되어 설정될 수 있다.

[0027] 제어부(400)는 사용자 성격 판단 시스템이 사용자의 음성 신호로부터 측정된 신호 특성에 기초하여 사용자의 성격을 판단할 수 있도록 음성 신호 수신부(100), 신호 특성 측정부(200), 및 사용자 성격 판단부(300)를 제어한다.

[0028] 또한, 본 발명의 일 실시예에 따른 사용자 성격 판단 시스템은 별도의 장치로 구현되어 사용자로부터의 음성을 직접 수신하고 녹음할 수 있으나, 이에 제한되지 않으며, 별도의 서버로 구현되어 다양한 사용자 단말 또는 외부 서버로부터 사용자의 음성 신호를 수신할 수도 있다.

[0029] 이하, 도 2를 참조하여 본 발명의 일 실시예에 따른 신호 특성 측정부(200)에 대하여 설명하기로 한다.

[0030] 도 2는 본 발명의 일 실시예에 따른 신호 특성 측정부(200)의 세부 구성도이다.

[0031] 도 2에 도시된 바와 같이, 본 발명의 일 실시예에 따른 신호 특성 측정부(200)는 스피치 레이트 측정부(210), 피치 변화 측정부(220), 음성 크기 측정부(230), 및 묵음 구간 측정부(240)를 포함한다.

[0032] 스피치 레이트 측정부(210)는 사용자의 음성 신호로부터 사용자의 스피치 레이트(speech rate, 말의 빠르기)를 측정한다. 스피치 레이트 측정부(210)는 단위시간 당 사용자가 얼마나 많은 단어 또는 음절을 말하였는지를 측정할 수 있다. 또한, 스피치 레이트 측정부(210)는 사용자가 어떤 의미의 말을 하였는지에 대한 언어적인 정보를 사용하지 않고, 비언어적 단서로써 발성된 음성 신호만을 가지고 말의 빠르기를 측정할 수 있다. 스피치 레이트 측정부(210)는, 예를 들어, i)입력신호를 반파 정류한 후 저역 필터를 통과시킨 후 100Hz로 Downsampling을 실시하고, ii)1초에서 2초 사이 길이의 윈도우를 75% 이상 겹친 움직임으로 신호에 적용하고, iii)이산 푸리에 변환(DFT)을 통해 단기 스펙트럼을 구하고, iv)스펙트럴 모멘트(Spectral Moment)를 계산하는 과정을 통하여, 사용자의 음성 신호로부터 말의 빠르기를 측정할 수 있다.

[0033] 피치 변화 측정부(220)는 사용자의 음성 신호로부터 음성 피치의 변화를 측정한다. 피치 변화 측정부(220)는 음성 신호에서 W 길이의 데이터를 가지고 d 길이의 지연 간격마다 자기상관(autocorrelation) 계산을 수행할 수 있으며, 계산 결과 값들 중 반복적인 피크의 주기를 피치로 정할 수 있으며, 음성 신호에서의 피치의 변화 정도를 측정할 수 있다.

[0034] 음성 크기 측정부(230)는 사용자의 음성 신호로부터 음성 신호의 세기를 측정한다. 음성 크기 측정부(230)는 음성 신호에 대한 에너지 스펙트럼을 분석함으로써 음성 신호의 세기를 측정할 수 있다.

[0035] 묵음 구간 측정부(240)는 사용자의 음성 신호로부터 묵음 구간을 측정한다. 묵음 구간 측정부(240)는 배경 잡음만이 존재하는 음성 구간의 에너지 값에 기초하여 문턱치(Threshold) 값을 결정할 수 있으며, 문턱치 값 이하의 에너지를 갖는 구간을 묵음 구간으로 판단할 수 있다.

[0036] 또한, 묵음 구간 측정부(240)는 사용자의 음성 신호를 묵음 구간과 유음 구간으로 구분하고, 각각의 길이를 계산하여 아래 [수학식 1]과 같이 변환할 수 있다.

수학식 1

$$S_{user} = \{v_1, s_1, v_2, s_2, \dots, s_{n-1}, v_n\}$$

[0037]

[0038] v: 유음 구간의 길이, s: 무음 구간의 길이

[0039] 또한, 묵음 구간 측정부(240)는 위와 같이 변환된 데이터로부터, 예를 들어, 아래의 3가지 정보를 추출할 수 있

다.

[0040] 첫째, 묵음 구간 측정부(240)는 유음 구간과 묵음 구간의 길이가 얼마나 오랫동안 지속되고 그 편차가 얼마나 큰지에 관한 정보를 추출할 수 있다. 이와 관련하여, 묵음 구간 측정부(240)는, 아래 [수학식 2]와 같이, 유음/묵음 구간 길이의 평균값과 표준 편차를 산출할 수 있다.

수학식 2

$$v_{avg} = \frac{\sum_{i=1}^n v_i}{n}, \quad v_{sd} = \sqrt{\frac{\sum_{i=1}^n (v_i - v_{avg})^2}{n}}, \quad s_{avg} = \frac{\sum_{i=1}^{n-1} s_i}{n-1}, \quad s_{sd} = \sqrt{\frac{\sum_{i=1}^{n-1} (s_i - s_{avg})^2}{n-1}}$$

[0041]

[0042] 둘째, 묵음 구간 측정부(240)는 전체 음성 중에서 묵음 구간이 차지하는 비율을 아래의 [수학식 3]과 같이 산출할 수 있다.

수학식 3

$$SR = \frac{\sum_{i=1}^{n-1} s_i}{\sum_{i=1}^n v_i + \sum_{i=1}^{n-1} s_i}$$

[0043]

[0044] 셋째, 묵음 구간 측정부(240)는 일정 시간 이상의 묵음 구간과 유음 구간의 발생 횟수를 추출할 수 있다. 묵음 구간 측정부(240)는, 예를 들어, 묵음 구간은 2초 이상 지속된 경우의 발생 횟수를 합산하고, 유음 구간은 3초 이상 지속된 경우의 발생 횟수를 합산할 수 있다.

[0045] 이하, 도 3을 참조하여 본 발명의 일 실시예에 따른, 음성 신호를 이용한 사용자 성격 판단 방법에 대하여 설명하기로 한다.

[0046] 도 3은 본 발명의 일 실시예에 따른 음성 신호를 이용한 사용자 성격 판단 방법의 세부 흐름도이다.

[0047] 단계 S300은 사용자의 음성 신호를 수신하는 단계이다. 단계 S300에서는 마이크 등의 장치를 통하여 사용자의 음성을 수신할 수도 있으나, 이에 제한되지 않으며, 외부 서버 또는 장치로부터 수집된 사용자의 음성 신호를 수신할 수도 있다. 또한, 단계 S300에서는 사용자에게 질의를 제공하고 제공된 질의에 응답되는 사용자의 답변 음성을 수신할 수도 있다.

[0048] 단계 S302는 사용자 음성 신호로부터 스피치 레이트를 측정하는 단계이다. 단위시간 당 사용자가 얼마나 많은 단어 또는 음절을 말하였는지를 측정할 수 있다. 또한, 단계 S302에서는 사용자가 어떤 의미의 말을 하였는지에 대한 언어적인 정보를 사용하지 않고, 비언어적 단서로써 발생된 음성 신호만을 가지고 말의 빠르기를 측정할 수 있다.

[0049] 단계 S304는 사용자 음성 신호로부터 피치의 변화를 측정할 수 있다. 단계 S304에서는 음성 신호에서 W 길이의 데이터를 가지고 d 길이의 지연 간격마다 자기상관(autocorrelation) 계산을 수행할 수 있으며, 계산 결과 값들 중 반복적인 피크의 주기를 피치로 정할 수 있으며, 음성 신호에서의 피치의 변화 정도를 측정할 수 있다.

[0050] 단계 S306은 사용자의 음성 신호로부터 음성의 크기를 측정하는 단계이다. 단계 S306에서는 음성 신호에 대한

에너지 스펙트럼을 분석함으로써 음성 신호의 세기를 측정할 수 있다.

- [0051] 단계 S308은 사용자의 음성 신호로부터 묵음 구간에 관한 정보를 측정하는 단계이다. 단계 S308에서는, 배경 잡음만이 존재하는 음성 구간의 에너지 값에 기초하여 문턱치(Threshold) 값을 결정할 수 있으며, 문턱치 값 이하의 에너지를 갖는 구간을 묵음 구간으로 판단할 수 있다. 또한, 단계 S308에서는 유음 구간과 묵음 구간이 얼마나 오랫동안 지속되고 그 편차가 얼마인지에 관한 정보, 전체 음성 중에서 묵음 구간이 차지하는 비율, 일정 시간 이상의 묵음 구간과 유음 구간의 발생 횟수를 산출할 수 있다.
- [0052] 단계 S310은 사용자의 음성 신호의 특성에 기초하여 사용자의 성격을 판단하는 단계이다. 단계 S310에서는 사용자 음성 신호로부터 측정된 스피치 레이트(speech rate, 말의 빠르기), 피치의 변화 정도, 음성 신호의 세기 및 묵음 구간에 관한 정보를 이용하여 사용자의 성격이 외향적인 성격인지 내향적인 성격인지를 판단할 수 있다.
- [0053] 또한, 단계 S310에서는, 스피치 레이트 값에 기초하여 사용자의 말의 빠르기가 기 설정된 수치 이상인 경우에 외향적인 성격으로 판단할 수 있다.
- [0054] 또한, 단계 S310에서는, 사용자 음성의 피치의 변화 정도에 기초하여 사용자의 성격을 판단할 수 있다. 이 경우, 음성 피치의 절대 값을 고려하기 보다는 음성 피치의 평균값에 대한 상대적인 변화율을 이용하여 사용자의 성격을 판단할 수 있다. 예를 들어, 사용자 성격 판단부(300)는 음성 피치의 변화율이 기 설정된 수치 이상인 경우에 외향적인 성격으로 판단할 수 있다.
- [0055] 또한, 단계 S310에서는, 사용자 음성의 신호 세기의 변화 정도에 기초하여 사용자의 성격을 판단할 수 있다. 이 경우, 음성 신호 세기의 절대 값을 고려하기 보다는 음성 신호 세기의 평균값에 대한 상대적인 변화율을 이용하여 사용자의 성격을 판단할 수 있다. 예를 들어, 사용자 성격 판단부(300)는 음성 신호 세기의 변화율이 기 설정된 수치 이상인 경우에 외향적인 성격으로 판단할 수 있다.
- [0056] 또한, 단계 S310에서는, 사용자 음성에 포함된 묵음 구간의 길이, 비율 및 장묵음(long silence interval) 발견 횟수에 기초하여 사용자의 성격을 판단할 수 있다. 예를 들어, 음성 신호 중에서 묵음 구간의 길이가 기 설정된 시간 이상으로 오랫동안 지속되거나 그 편차가 기 설정된 수치 이상인 경우에 내향적인 성격으로 판단할 수 있다. 또한, 전체 음성 중에서 묵음 구간이 차지하는 비율이 기 설정된 수치 이상인 경우에 외향적인 성격으로 판단할 수 있다. 또한, 일정 시간 이상의 묵음 구간이 발생한 횟수가 기 설정된 수치 이상인 경우에 사용자의 성격을 판단할 수 있다.
- [0057] 실험 결과
- [0058] 사용자의 성격과 음성 신호 특성과의 상관성을 분석해보기 위해 우선 1차 테스트로 사용자가 외향적인지 내향적인지를 판단하기 위한 기준 문항을 마련하여 성격 검사를 실시하였다. 이 검사는 10가지 문항으로 구성된 온라인 설문 조사를 통해 이루어졌다. 설문 조사의 참가 자격은 20세에서 29세 사이의 남녀 대학생이었고, 총 138명의 학생들 중 1점부터 5점에 해당되는 내향성 학생들 20명과 6점부터 10점의 외향성 학생들 20명을 선발하였고, 이후 2차 테스트를 실시하였다.
- [0059] 40명의 학생들에게 실시된 2차 테스트는 음성녹음을 위해 질의응답의 인터뷰 방식으로 진행하였다. 질문자가 실험자에게 가족과 일상생활에 관련된 5가지씩 두 번에 걸쳐 준비된 질문을 하고 실험대상 학생이 대답하게 하였고, 자연스러운 대화와 글을 읽을 때의 다른 점을 분석하기 위해 약 30초정도 길이 정도의 글을 두 가지로 만들어서 읽어보도록 하였다. 그리하여 총 160개, 약 280분 길이의 음성 데이터 set으로 실험하였다. 모든 음성 데이터는 16khz의 샘플링 주파수와 mono 채널, 16bits로 양자화 되었다. 데이터 녹음은 조용한 실험실 환경에서 이루어졌다.
- [0060] 또한, 획득된 음성 데이터로부터 모두 12개의 feature set을 추출하였다. 실험 type에 따라 Reading 과 Q&A로 나누었고, 각각의 type에서 과거의 연구에서 주로 사용되었던 음성 특성인 말의 빠르기(Speech Rate), Fundamental frequency(f0), Energy를 우선적으로 추출하였고, 이번 연구에서 중점적 분석대상인 Silence와 관련된 특성으로 묵음 구간의 길이, 비율, 장묵음(long silence interval) 발견 횟수 등 3가지 feature를 구성하였다.

표 1

Type	Feature Set	Description
Reading	F1	reading 시 말의 speed
	F2	reading 시 f0
	F3	reading 시 energy
	F4	reading 시 묵음/유음구간의 길이
	F5	reading 시 전체 구간 중 묵음구간의 비율
	F6	reading 시 기준길이 이상의 묵음/유음 구간의 발견 횟수
Question & Answer	F7	Q/A 시 말의 speed
	F8	Q/A 시 f0
	F9	Q/A 시 energy
	F10	Q/A 시 묵음/유음구간의 길이
	F11	Q/A 시 전체 구간 중 묵음구간의 비율
	F12	Q/A 시 일정길이 이상의 묵음/유음 구간의 발견 횟수

[0061]

[0062]

[표 1]은 이번 실험에서 이용한 12가지의 feature set와 각각의 의미를 설명한 표이다. 각각의 feature set은 변수 별로 0부터 1사이의 값으로 정규화(normalization) 하였다. C-SVC에서 데이터를 정규화 하는 목적은 값이 큰 변수가 작은 변수들을 압도하여 더 큰 영향력을 행사하는 것을 막기 위함이다. 정규화를 통해 각각의 변수가 동등하게 target variable에 영향을 미치도록 하였다.

[0063]

C-SVC를 이용한 실험에서 사전에 필요한 것은 kernel function을 선택하는 것과, parameter값을 결정하는 일이다. Linear kernel 을 이용할 경우에는 Penalty parameter인 C만을 결정하면 되지만, 이번 실험에서는 Nonlinear kenel 인 RBF를 이용하기 때문에, 값도 결정되어야 했다. C와 의 최적 값을 찾기 위해, -10부터 10까지 1씩 증가시키는 grid search 방법을 통해 각각의 feature set 별로 최적 값을 탐색하였으며, 최적 값은 C=25, =22 으로 탐색되었다.

[0064]

도 4는 본 발명의 일 실시예에 따른 총 12개의 feature set 에 대해 5-fold cross validation을 수행하여 예측력을 비교한 결과를 도시한 도면이다.

[0065]

전체적으로 Reading과 Q&A로 나누어 보면 평균적으로 각각 65.4%와 69.0%로 나타났고, 이는 Q&A에서 발화자의 성격이 더 반영이 되는 것으로 분석할 수 있다. Feature 측면에서는 Reading과 Q&A와 관계없이 Energy의 통계치가 평균 72.5%로 높은 정확도를 보였는데, 이는 기대한 것처럼 외향적인 사람이 목소리가 크다고 분석할 수 있다. 하지만 이는 물리적으로 목소리가 클 수도 있지만 심리적으로 적극적인 발화 행태에 기인한 것으로 생각할 수도 있겠다. 이 실험에서 발화자의 성격이 가장 잘 반영된 것으로는 F11인 Q/A 시 전체 구간 중 묵음 구간의 비율로 accuracy가 76.25%로 가장 높게 나왔으며, Q&A의 silence 관련 feature들인 F10, F11, F12가 높은 정확도를 보여주었다.

[0066]

결론적으로 Q&A와 같은 자연스러운 대화 시에 묵음과 관련된 여러 가지 통계 자료가 발화자의 성격을 가장 잘 반영하였다. Reading은 모든 피실험자가 동일한 script를 읽게 되어 피실험자의 성향이 반영될 여지가 많지 않지만, Q&A 실험에서는 피실험자의 응답이 모두 다르게 나올 뿐만 아니라 사용자의 성격적인 성향이 질문에 대한 응답이라는 행태 속에 반영이 되는 것으로 분석되었다. 특히 sound/ silence 길이와 비율에 영향을 미치게 됨을 실험 결과를 통해 알 수 있었다.

[0067]

도 5는 앞선 실험 결과에서 가장 높은 성능을 보인 F11 feature set에서 다룬 전체 음성 중 묵음 비율을 비교하여 도시한 도면이다.

[0068]

Reading 실험에서는 내향/외향과 상관없이 거의 비슷한 묵음 비율을 보였지만, Q&A 실험에서는 외향적 피실험자들의 묵음 비율이 내향적인 피실험자의 묵음 비율과 많이 차이가 나는 것을 볼 수 있었다. 이는 피실험자에게 주어지는 자유도(Degree of Freedom)가 높은 Q&A 실험에서 외향적인 사람은 말의 비중이 상대적으로 높고, 내향적인 사람은 말을 하지 않는 시간이 상대적으로 길다는 것을 알 수 있었다. 이러한 특성이 내향적/외향적 사람

을 구분해내는 예측변수로 중요한 역할을 한다는 것을 이번 실험을 통해 알 수 있다.

[0069] 도 6은 F12 feature set에서 다루어진 일정 길이 이상의 묵음 구간의 발견 횟수를 비교한 그림이다. 본 실험에서는 2초 이상 묵음이 지속된 경우의 횟수를 사용자별로 합산 하였다. 좌측은 외향적 피실험자들의 묵음 구간의 발견 횟수이고, 우측은 내향적 피실험자들의 묵음 구간 발견 횟수이다. 외향적 그룹은 대부분 0에서 3회까지가 대부분이어서 평균 값은 1.1회로 나왔고, 내향적 그룹은 3회 이상 발견된 경우가 많아 평균값은 3.2회로 나왔다. 이는 내향적인 사람들이 외향적인 사람들에 비해 장시간 말을 하지 않는 횟수가 많음을 의미하며, 머릿 거림의 또 다른 지표로써 성격을 구분하는 중요한 변수로 이용될 수 있음을 보여준다.

[0070] 위와 같은 실험 결과에서 알 수 있는 바와 같이, 본 발명의 일 실시예에 따른 사용자 성격 판단 시스템 명 방법에서는 사용자의 음성을 언어적으로 분석하지 않더라도 사용자의 음성 신호로부터 스피치 레이트, 피치 변화, 음성 크기 및 묵음 구간을 측정하고 이용함으로써 사용자의 성격을 보다 정확하게 판단할 수가 있다.

[0071] 본 발명의 일 실시예는 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독 가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈, 또는 반송파와 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.

[0072] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

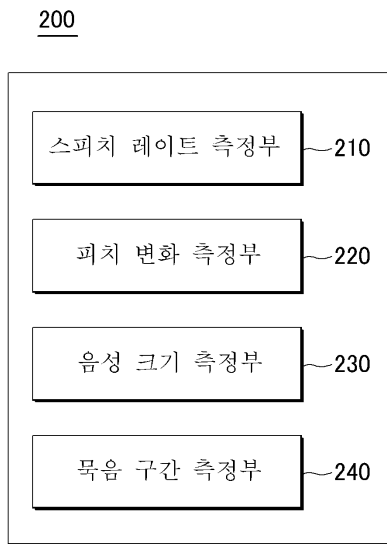
[0073] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

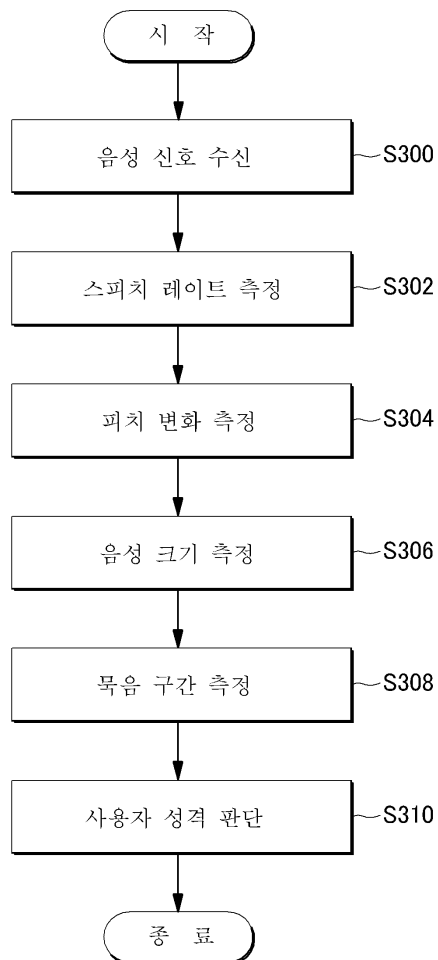
도면1



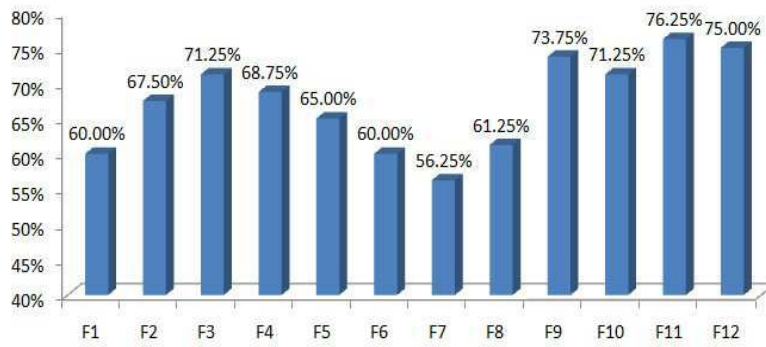
도면2



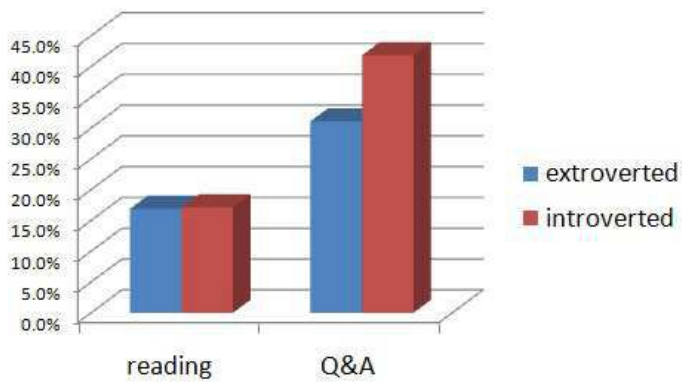
도면3



도면4



도면5



도면6

