



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년09월17일
(11) 등록번호 10-1309954
(24) 등록일자 2013년09월11일

(51) 국제특허분류(Int. Cl.)

G06K 9/46 (2006.01)

(21) 출원번호 10-2012-0079070

(22) 출원일자 2012년07월20일

심사청구일자 2012년07월20일

(56) 선행기술조사문헌

JP11232456 A

JP2004513462 A

KR1020100081874 A

(73) 특허권자

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

이성환

서울특별시 강남구 도곡2동 대림아크로빌 B-2406

허동철

서울특별시 광진구 군자동 370-1번지 202호

크리스티안 월러벤

서울특별시 성북구 안암동5가 고려대학교 CJ하우스 302호

(74) 대리인

민영준, 최관락, 송인호

전체 청구항 수 : 총 10 항

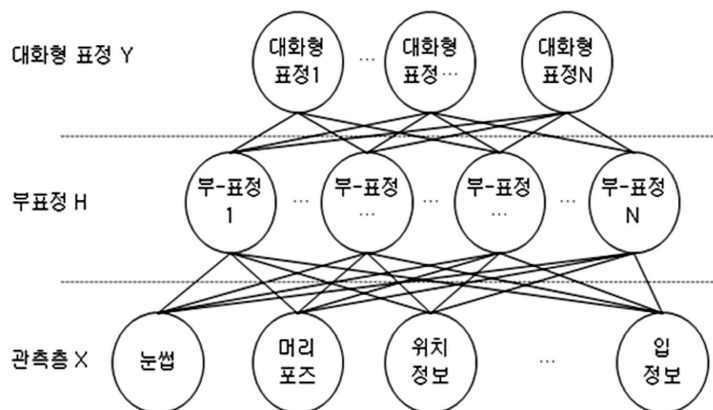
심사관 : 이별섭

(54) 발명의 명칭 대화형 표정 인식 장치 및 방법과 이에 관한 기록매체

(57) 요약

대화형 표정 인식 장치 및 방법과 이에 관한 기록매체가 개시된다. 개시된 대화형 표정 인식 장치는 입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출하는 추출부; 및 미리 설정된 특징 패턴들 중 적어도 2 이상을 소정의 배열 순서로 포함하고 있는 다수의 대화형 표정들에 관한 특징 신호와 상기 추출된 특징 신호가 상응하는지 여부로부터 상기 다수의 대화형 표정들 중 상기 입력 영상에 대응하는 대화형 표정을 결정하는 결정부;를 포함한다. 본 발명에 따르면, 대화 과정에서 나타나는 의사의 표현이나 보조를 위하여 사용되는 얼굴 표정이 인식될 수 있어 사람과 기계 사이에 보다 자유로운 인터랙션을 가능하게 하는 장점이 있다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	R312010000100080
부처명	한국과학재단
연구사업명	세계적수준의연구중심대학육성사업
연구과제명	뇌공학 융합 기술 연구
주관기관	고려대학교 산학협력단
연구기간	2010.10.01 ~ 2011.08.31

특허청구의 범위

청구항 1

입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출하는 추출부; 및
미리 설정된 특징 패턴들 중 적어도 2 이상을 소정의 배열 순서로 포함하고 있는 다수의 대화형 표정들에 관한 특징 신호와 상기 추출된 특징 신호가 상응하는지 여부로부터 상기 다수의 대화형 표정들 중 상기 입력 영상에 대응하는 대화형 표정을 결정하는 결정부;
를 포함하는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 2

제1항에 있어서,
상기 결정부는,
상기 미리 설정된 특징 패턴들이 각각 상기 추출된 특징 신호에 일치할 확률 값을 산출하는 제1 산출부;
상기 확률 값들 및 상기 확률 값들의 배열 순서가 상기 대화형 표정들에 각각 포함된 2 이상의 미리 설정된 특징 패턴들의 확률 값들 및 배열 순서와 일치할 확률 값을 산출하는 제2 산출부; 및
상기 제2 산출부에 의해 산출된 적어도 하나의 확률 값 중 가장 높은 값으로 일치하는 대화형 표정을 상기 추출된 특징 신호에 대응하는 것으로 판단하는 판단부;
를 포함하는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 3

제2항에 있어서,
상기 제1 산출부는 상기 입력 영상의 소정의 구간을 구성하고 있는 프레임들에 대해 연속적으로 상기 확률 값을 산출하는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 4

제1항에 있어서,
상기 미리 설정된 특징 패턴들은, 영상에서 얼굴을 구성하는 요소들의 움직임과 특징 신호를 회귀 분석함으로써 설정되는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 5

제1항에 있어서,
상기 추출부는 상기 입력 영상에서의 컬러 정보 및 깊이 정보를 이용하여 상기 얼굴을 구성하는 요소들을 검출하는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 6

제5항에 있어서,
상기 얼굴을 구성하는 요소들은 눈, 코, 입, 이마, 눈썹, 미간, 턱 중 적어도 하나를 포함하는 것을 특징으로 하는 대화형 표정 인식 장치.

청구항 7

입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출하는 단계; 및
미리 설정된 특징 패턴들 중 적어도 2 이상을 소정의 배열 순서로 포함하고 있는 다수의 대화형 표정들에 관한 특징 신호와 상기 추출된 특징 신호가 상응하는지 여부로부터 상기 다수의 대화형 표정들 중 상기 입력 영상에

대응하는 대화형 표정을 결정하는 단계;

를 포함하는 것을 특징으로 하는 대화형 표정 인식 방법.

청구항 8

제7항에 있어서,

상기 결정하는 단계는,

상기 미리 설정된 특징 패턴들이 각각 상기 추출된 특징 신호에 일치할 제1 확률 값을 산출하는 단계;

상기 제1 확률 값들 및 상기 제1 확률 값들의 배열 순서가 상기 대화형 표정들에 각각 포함된 2 이상의 미리 설정된 특징 패턴들의 확률 값들 및 배열 순서와 일치할 제2 확률 값을 산출하는 단계; 및

상기 적어도 하나의 제2 확률 값 중 가장 높은 값으로 일치하는 대화형 표정을 상기 추출된 특징 신호에 대응하는 것으로 판단하는 단계;

를 포함하는 것을 특징으로 하는 대화형 표정 인식 방법.

청구항 9

제8항에 있어서,

상기 제1 확률 값을 산출하는 단계는 상기 입력 영상의 소정의 구간을 구성하고 있는 프레임들에 대해 연속적으로 상기 제1 확률 값을 산출하는 것을 특징으로 하는 대화형 표정 인식 방법.

청구항 10

제7항 내지 제9항 중 어느 한 항에 기재된 대화형 표정 인식 방법을 수행하는 프로그램이 기록된 컴퓨터 판독 가능한 기록매체.

명세서

기술분야

[0001] 본 발명의 실시예들은 대화형 표정 인식 장치 및 방법과 이에 관한 기록매체에 관한 것으로서, 더욱 상세하게는 의사의 표현이나 보조를 위하여 사용되는 얼굴 표정을 인식함으로써 사람과 기계 사이에 보다 자유로운 인터랙션을 가능하게 하는 대화형 표정 인식 장치 및 방법과 이에 관한 기록매체에 관한 것이다.

배경기술

[0002] 사람과 기계 사이의 인터랙션을 위하여 영상으로부터 사람의 표정을 인식하기 위한 연구가 다양한 방식으로 수행되고 있다.

[0003] 일반적으로, 영상으로부터 표정을 인식하기 위해서는 사람들의 표정에서 공통적으로 나타나는 기본적인 표정들, 즉, 화남, 역겨움, 두려움, 행복, 슬픔, 놀람 등과 같은 기본적인 표정들을 모델링하고 이를 표정 인식에 활용하고 방식이 사용되고 있다.

[0004] 그리고, 해부학적으로 얼굴을 구성하고 있는 근육의 움직임을 Facial Action Unit으로 코드화하고 이를 표정의 인식에 활용하는 Facial Action Coding System 방식도 있다.

[0005] 이러한 표정 인식 방식들은 영상 처리 및 특징 추적의 어려움으로 인하여 정면에서 획득된 얼굴 영상으로부터 표정을 인식하고 있고, 얼굴 영역 및 구성 요소의 움직임만을 고려하여 표정을 인식하고 있다.

[0006] 이에 따라, 기존의 방식에 따른 인터랙션은 사람과 사람 사이에서의 일상적인 대화 과정에서 나타나는 것과 같은, 다양한 얼굴 표정의 조합으로 상대방의 의견에 대한 감정이나 의견 등을 표출하는 얼굴 표정을 인식할 수 없다는 한계를 갖게 된다.

[0007] 즉, 기존의 방식에 따르면, 영상으로부터 사람의 일시적인 감정 상태만을 인식할 수 있을 뿐, 대화 과정에서 나타나는 의사의 표현이나 보조를 위하여 사용되는 얼굴 표정은 인식될 수 없으므로, 사람과 기계 사이에 보다 자유로운 인터랙션이 불가능한 문제점이 존재한다.

발명의 내용

해결하려는 과제

[0008] 상기한 바와 같은 종래기술의 문제점을 해결하기 위해, 본 발명에서는 의사의 표현이나 보조를 위하여 사용되는 얼굴 표정을 인식함으로써 사람과 기계 사이에 보다 자유로운 인터랙션을 가능하게 하는 대화형 표정 인식 장치 및 방법과 이에 관한 기록매체를 제안하고자 한다.

[0009] 본 발명의 다른 목적들은 하기의 실시예를 통해 당업자에 의해 도출될 수 있을 것이다.

과제의 해결 수단

[0010] 상기한 목적을 달성하기 위해 본 발명의 바람직한 일 실시예에 따르면, 입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출하는 추출부; 및 미리 설정된 특징 패턴들 중 적어도 2 이상을 소정의 배열 순서로 포함하고 있는 다수의 대화형 표정들에 관한 특징 신호와 상기 추출된 특징 신호가 상응하는지 여부로부터 상기 다수의 대화형 표정들 중 상기 입력 영상에 대응하는 대화형 표정을 결정하는 결정부;를 포함하는 대화형 표정 인식 장치가 제공된다.

[0011] 상기 결정부는, 상기 미리 설정된 특징 패턴들이 각각 상기 추출된 특징 신호에 일치할 확률 값을 산출하는 제1 산출부; 상기 확률 값들 및 이들의 배열 순서가 상기 대화형 표정들에 각각 포함된 2 이상의 미리 설정된 특징 패턴들의 확률 값들 및 배열 순서와 일치할 확률 값을 산출하는 제2 산출부; 및 상기 제2 산출부에 의해 산출된 적어도 하나의 확률 값 중 가장 높은 값으로 일치하는 대화형 표정을 상기 추출된 특징 신호에 대응하는 것으로 판단하는 판단부;를 포함할 수 있다.

[0012] 상기 제1 산출부는 상기 입력 영상의 소정의 구간을 구성하고 있는 프레임들에 대해 연속적으로 상기 확률 값을 산출할 수 있다.

[0013] 상기 미리 설정된 특징 패턴들은, 영상에서 얼굴을 구성하는 요소들의 움직임과 특징 신호를 회귀 분석함으로써 설정될 수 있다.

[0014] 상기 추출부는 상기 입력 영상에서의 컬러 정보 및 깊이 정보를 이용하여 상기 얼굴을 구성하는 요소들을 검출할 수 있다.

[0015] 상기 얼굴을 구성하는 요소들은 눈, 코, 입, 이마, 눈썹, 미간, 턱 중 적어도 하나를 포함할 수 있다.

[0016] 또한, 본 발명의 다른 실시예에 따르면, 입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출하는 단계; 및 미리 설정된 특징 패턴들 중 적어도 2 이상을 소정의 배열 순서로 포함하고 있는 다수의 대화형 표정들에 관한 특징 신호와 상기 추출된 특징 신호가 상응하는지 여부로부터 상기 다수의 대화형 표정들 중 상기 입력 영상에 대응하는 대화형 표정을 결정하는 단계;를 포함하는 대화형 표정 인식 방법이 제공된다.

[0017] 또한, 본 발명의 다른 실시예에 따르면, 상기의 대화형 표정 인식 방법 중 어느 하나를 수행하는 프로그램이 기록된 컴퓨터 판독 가능한 기록매체가 제공된다.

발명의 효과

[0018] 본 발명에 따르면, 대화 과정에서 나타나는 의사의 표현이나 보조를 위하여 사용되는 얼굴 표정이 인식될 수 있어 사람과 기계 사이에 보다 자유로운 인터랙션을 가능하게 하는 장점이 있다.

도면의 간단한 설명

[0019] 도 1은 본 발명의 일 실시예에 따른 대화형 표정 인식 장치가 입력 영상으로부터 대화형 표정을 인식하는 과정에 관한 계층도를 도시하는 도면이다.

도 2는 본 발명의 일 실시예에 따른 대화형 표정 인식 장치가 입력 영상으로부터 대화형 표정을 인식하는 과정에 관한 실시예를 도시하는 도면이다.

도 3은 본 발명의 일 실시예에 따른 대화형 표정 인식 장치가 입력 영상으로부터 대화형 표정을 인식하는 동작에 관한 블록도를 도시한다.

도 4는 본 발명의 일 실시예에 따라, 자동적 부표정 모델링을 위하여, LDCRF(Latent Dynamic Conditional

Random Fields)에 따른 모델링 과정을 도시하는 도면이다.

도 5는 본 발명의 일 실시예에 따른 대화형 표정 인식 장치의 상세한 구성을 도시하는 도면이다.

도 6은 본 발명의 일 실시예에 따른 대화형 표정 인식 방법을 시간의 흐름에 따라 상세히 도시한 순서도이다.

도 7은 본 발명의 일 실시예에 따른 결정부의 상세한 구성을 도시하는 도면이다.

도 8은 본 발명의 일 실시예에 따른 대화형 표정을 결정하는 단계를 시간의 흐름에 따라 상세히 도시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0020] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0021] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.

[0022] 도 1은 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)가 입력 영상으로부터 대화형 표정을 인식하는 과정에 관한 계층도를 도시하는 도면이다.

[0023] 그리고, 도 2는 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)가 입력 영상으로부터 대화형 표정을 인식하는 과정에 관한 실시예를 도시하는 도면이다.

[0024] 대화형 표정은 앞서 설명한 바와 같이, 일상적인 대화 과정에서 상대방에게 자신의 의사를 표현하거나 보조하기 위하여 사용하는 얼굴 표정이다.

[0025] 이를 인식하기 위하여, 대화형 표정 인식 장치(100)는 도 1 및 도 2에 도시된 바와 같이, 눈썹, 머리 포즈 등과 같은 얼굴의 구성 요소들의 움직임 조합으로부터 부표정을 유추하고, 유추된 부표정들의 조합으로부터 최종적으로 대화형 표정을 인식한다.

[0026] 이때, 부표정들은 도 2에 도시된 바와 같이, 일례로, '무표정', '고심하는', '동의하는', '거절하는' 등의 표정일 수 있으며, 대화형 표정은 이러한 부표정들의 조합으로서, 일례로, '심사숙고 후 동의', '심사숙고 후 거절' 등의 표정일 수 있다.

[0027] 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)가 이러한 계층적 구조로 대화형 표정을 인식하는 것은 다음과 같은 대화형 표정의 특징들에 기인한다.

[0028] 하나의 대화형 표정은 여러 개의 부표정으로 구성되어 있고, 다수의 대화형 표정들은 여러 개의 부표정을 공유하고 있으며, 대화형 표정들간에 부표정과 부표정 순서의 차이가 존재한다.

[0029] 즉, 다수의 대화형 표정들에서 공통적으로 나타나는 부표정들을 모델링하고, 모델링된 부표정들이 나열되는 순서로부터 대화형 표정을 모델링하는 경우, 대화형 표정의 인식이 가능하게 되는 것이다.

[0030] 도 2를 참조하면, 대화형 표정들인 '심사숙고 후 동의'와 '심사숙고 후 거절'은 각각 '무표정', '고심하는', '동의하는', '거절하는'의 부표정으로 구성되어 있고, 이들 가운데 '무표정', '고심하는' 부표정들을 공유하고 있으며, 전자의 부표정 구성 순서는 '무표정, 고심하는, 동의하는' 순으로, 후자의 부표정 구성 순서는 '무표정, 고심하는, 거절하는' 순으로 나타나고 있다.

[0031] 일반적으로, 이러한 대화형 표정은 복잡한 움직임으로 구성되고 이에 따라 긴 시퀀스 길이를 갖는다는 점에서, 상기의 특징을 이용하는 본 발명에 따르면, 보다 신속하고 정확하게 대화형 표정의 인식이 가능한 장점이 있다.

[0032] 계속하여, 도 3은 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)가 입력 영상으로부터 대화형 표정을 인식하는 동작에 관한 블록도를 도시한다.

[0033] 도 3에 도시된 바와 같이, 대화형 표정 인식 장치(100)는 입력 영상으로부터 대화형 표정을 인식하기 위하여, 모델링된 부표정으로부터 영상에서 부표정을 인식하고, 모델링된 대화형 표정으로부터 영상에서 대화형 표정을

인식하게 된다.

- [0034] 부표정의 모델링 및 대화형 표정의 모델링을 위하여, 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)의 인식 과정에는 대화형 표정 학습 과정이 수반될 수 있다.
- [0035] 경우에 따라, 학습 과정은 대화형 표정의 인식 동작과 동시에 수행될 수 있으나, 이하에서는 설명의 편의를 위해, 학습 과정이 인식 동작과는 별개로 미리 수행되고 저장부(미도시)와 같은 저장 공간에 데이터베이스화되어 저장된 것으로 가정하여 설명하기로 한다.
- [0036] 학습 과정은 대화형 표정을 포함하는 영상 데이터와 대화형 표정에 대한 레이블 데이터의 관계를 회귀 분석함으로써 수행될 수 있다.
- [0037] 보다 상세하게, 대화형 표정 영상으로부터 얼굴 구성 요소들, 즉, 얼굴 포즈, 얼굴 위치, 눈, 코, 입, 눈썹, 이마, 미간, 턱의 움직임은 인식하여 특징 신호를 검출하고, 특징 신호로부터 빈번하게 발생하는 특징 신호 패턴을 찾으며, 찾아진 특징 신호 패턴에 대하여 부표정 인덱스를 부여하고 이들 사이의 관계를 회귀 분석을 통하여 결정함으로써, 부표정에 관한 모델링이 수행될 수 있다.
- [0038] 이러한 과정은 앞서 설명한 대화형 표정의 특징 중 하나로서, 대화형 표정들 각각은 여러 개의 부표정으로 구성되어 있고, 적어도 하나 이상의 부표정을 공유하고 있다는 점에 대응한다.
- [0039] 마찬가지로, 대화형 표정 영상으로부터 빈번하게 발생하는 특징 신호 패턴에 대응하는 부표정들의 발생 확률을 시간에 따라 연속적으로 구성하고, 연속적으로 구성된 부표정들의 발생 확률에 대화형 표정 레이블을 부여하고 이들 사이의 관계를 회귀 분석을 통하여 결정함으로써, 대화형 표정에 관한 모델링이 수행될 수 있다.
- [0040] 이러한 과정은 앞서 설명한 대화형 표정의 특징 중 하나로서, 대화형 표정들간에 부표정과 부표정 순서의 차이가 존재한다는 점에 대응한다.
- [0041] 상기의 모델링 과정을 통하여 대화형 표정의 학습이 이루어지면, 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)는 학습된 데이터를 기초로 하여 입력 영상의 특징 신호를 검출하고, 부표정을 인식하며, 최종적으로 대화형 표정을 인식한다.
- [0042] 한편, 상기의 모델링 과정들은 계층적 시퀀스 모델링 기법으로서, Sparse Hidden Dynamic Conditional Random Fields, Latent Dynamic Conditional Random Fields, Deep Conditional Random Fields, Neuro Conditional Random Fields, Conditional Neural Fields, Dynamic Bayesian Network, Hidden Markov Model과 같은 기법에 의해 수행될 수 있으며, 본 발명에서는 도 4에 도시된 바와 같이, 가장 높은 확률이 결과로 출력됨으로써 자동적으로 부표정이 모델링되는 데 적합한 LDCRF(Latent Dynamic Conditional Random Fields)를 적용할 수 있다.
- [0043] 도 5는 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)의 상세한 구성을 도시하는 도면이다.
- [0044] 그리고, 도 6은 본 발명의 일 실시예에 따른 대화형 표정 인식 방법을 시간의 흐름에 따라 상세히 도시한 순서도이다.
- [0045] 도 5 및 도 6에 도시된 바와 같이, 대화형 표정 인식 장치(100)는 추출부(110) 및 결정부(120)를 포함하며, 대화형 표정 인식 방법은 움직임에 관한 특징 신호를 추출하는 단계(S610) 및 추출된 특징 신호에 대응하는 대화형 표정을 결정하는 단계(S620)를 포함한다.
- [0046] 이하, 도 5 및 도 6을 참조하여, 본 발명의 일 실시예에 따른 대화형 표정 인식 장치(100)가 입력 영상으로부터 대화형 표정을 인식하는 동작(즉, 대화형 표정 인식 방법)을 보다 상세히 살펴본다.
- [0047] 먼저, 단계(S610)에서 추출부(110)는 입력 영상으로부터 얼굴을 구성하는 요소들의 움직임에 관한 특징 신호를 추출한다.
- [0048] 입력 영상에서 얼굴을 구성하고 있는 요소들은 영상에 나타나는 얼굴의 눈, 코, 입, 이마, 눈썹, 미간, 턱과 이들의 2차원 내지 3차원 위치로부터 유추되는 얼굴의 3차원 포즈 등일 수 있으며, 이러한 요소들의 움직임으로부터 움직임 패턴에 관한 정보가 추출될 수 있다.
- [0049] 이를 위해, 본 발명의 일 실시예에 따른 입력 영상은 가시광 카메라에 의해 획득되는 컬러 영상뿐만 아니라 적외선 카메라에 의해 획득되는 깊이 영상을 포함할 수 있다.

- [0050] 입력 영상을 분석하여 얼굴의 구성 요소들의 움직임 패턴에 관한 정보를 추출하는 단계(S610)를 보다 상세히 살펴보면 다음과 같다.
- [0051] 즉, 본 발명의 일 실시예에 따른 단계(S610)는 영상에서 얼굴 영역을 검출하는 단계(S611), 주요 얼굴 구성 요소의 3차원 위치를 검출하는 단계(S612), 얼굴의 3차원 포즈를 추출하는 단계(S613), 주요 얼굴 구성 요소를 기준으로 주변 얼굴 구성 요소를 검출하는 단계(S614) 및 검출된 얼굴 구성 요소의 움직임으로부터 패턴 정보를 추출하는 단계(S615)를 포함할 수 있다.
- [0052] 단계(S611)에서 추출부(110)는 얼굴 영역의 검출을 위하여, 얼굴이 포함된 전경과 나머지 배경을 식별하기 위한 배경 모델링에서 컬러 정보와 깊이 정보를 결합하여 픽셀 단위 가우시안(Gaussian) 모델링을 통하여 확률로서 전경을 식별한다.
- [0053] 이때, 식별된 전경에 대한 컬러 및 깊이 정보로부터 픽셀 단위 모션을 분석하여 손과 같은 인체 구성 요소보다 모션이 적게 발생하는 얼굴 영역을 식별하는 기법이 추가적으로 수행될 수 있다.
- [0054] 다음으로, 단계(S612)에서 추출부(110)는 컬러 및 깊이 정보로부터 얼굴의 주요 구성 요소, 즉, 눈, 코, 입의 2차원 위치를 추출하고, 2차원 위치로부터 3차원 위치로 변환하는 방식으로 주요 얼굴 구성 요소들의 3차원 위치를 검출한다.
- [0055] 그리고, 단계(S613)에서 추출부(110)는 주요 얼굴 구성 요소의 3차원 위치를 특징으로 사용하여 코를 중심으로 Yaw, Tilt, Roll과 같은 3차원 방향 정보를 획득함으로써 얼굴의 3차원 포즈를 추출할 수 있다.
- [0056] 그리고, 단계(S614)에서 추출부(110)는 주요 얼굴 구성 요소의 2차원 위치와 나머지 얼굴 영역의 관계를 고려하여 주변 얼굴 구성 요소, 즉, 이마, 눈썹, 미간, 턱을 검출할 수 있다.
- [0057] 마지막으로, 단계(S615)에서 추출부(110)는 상기의 단계(S611 내지 S615)에서 검출된 얼굴 구성 요소들의 움직임을 인식하여 움직임 패턴에 관한 정보를 추출할 수 있다.
- [0058] 이때, 추출부(110)는 Gabor 필터링을 통하여 색상 및 형태에 대해 분석하며 분류기를 통하여 움직임을 인식할 수 있으며, 움직임 패턴에 관한 정보는 확률로 정규화된 신호의 형태로 획득될 수 있다.
- [0059] 다시, 도 5 및 도 6을 참조하면, 본 발명의 일 실시예에 따른 단계(S620)에서 결정부(120)는 단계(S610)에서 추출된 특징 신호와 대화형 표정에 관한 특징 신호의 상응 여부로부터 단계(S610)에서 추출된 특징 신호에 대응하는 대화형 표정을 결정한다.
- [0060] 이때, 대화형 표정에 관한 특징 신호는 미리 설정된 특징 신호 패턴들 중 적어도 하나를 소정의 배열 순서로 포함하고 있다.
- [0061] 즉, 앞서 설명한 바와 같이, 모델링된 대화형 표정은 빈번하게 발생하는 특징 신호 패턴에 대응하는 부표정들의 발생 확률이 연속적으로 구성된 것이므로, 추출된 특징 신호가 모델링된 대화형 표정의 특징 신호에 상응하는지를 판단하게 되면, 어떠한 대화형 표정인지를 결정할 수 있게 된다.
- [0062] 이하, 도 7 및 도 8을 참조하여, 본 발명의 일 실시예에 따른 결정부(120)가 추출된 특징 신호로부터 대화형 표정을 결정하는 동작을 보다 상세히 살펴보기로 한다.
- [0063] 도 7은 본 발명의 일 실시예에 따른 결정부(120)의 상세한 구성을 도시하는 도면이다.
- [0064] 그리고, 도 8은 본 발명의 일 실시예에 따른 대화형 표정을 결정하는 단계(S620)를 시간의 흐름에 따라 상세히 도시한 순서도이다.
- [0065] 도 7 및 도 8에 도시된 바와 같이, 결정부(120)는 제1 산출부(122), 제2 산출부(124) 및 판단부(126)를 포함할 수 있고, 대화형 표정을 결정하는 단계(S620)는 제1 확률 값을 산출하는 단계(S622), 제2 확률 값을 산출하는 단계(S624) 및 가장 높은 확률로 일치하는 대화형 표정을 판단하는 단계(S626)를 포함할 수 있다.
- [0066] 먼저, 단계(S622)에서 결정부(120)의 제1 산출부(122)는 미리 설정된 특징 패턴들이 각각 단계(S610)에서 추출된 특징 신호에 일치할 확률 값을 산출한다.
- [0067] 즉, 단계(S622)에서 결정부(120)는 회귀 분석을 통하여 모델링된 부표정에 관한 특징 신호에 관한 패턴이 단계(S610)에서 추출된 특징 신호에 얼마나 포함되어 있는지를 확률 값으로 판단한다.

- [0068] 일례로, 미리 설정된 특징 신호에 관한 패턴들로서, α , β , γ , δ 가 정의된 경우, 제1 산출부(122)는 추출된 특징 신호에 α 가 70%, β 가 70%, γ 80%, δ 가 10% 일치할 확률로 포함되어 있다고 판단할 수 있다.
- [0069] 이와 같이 산출된 확률 값들은 단계(S624)로 제공되게 되며, 이하 설명하는 바와 같이, 결정부(620)가 확률 값들의 분포로부터 대응하는 대화형 표정을 결정할 수 있도록 한다.
- [0070] 이때, 산출된 확률 값들의 배열 순서가 단계(S624)에서 활용될 수 있도록, 단계(S622)는 입력 영상의 소정의 구간을 구성하고 있는 프레임들(일례로, 64프레임)에 대해 프레임 단위로 연속적으로 수행되는 것이 바람직하다.
- [0071] 계속하여, 단계(S624)에서 제2 산출부(124)는 단계(S622)에서 산출된 확률 값들 및 이들의 배열 순서가 대화형 표정들에 각각 포함된 2 이상의 미리 설정된 특징 패턴의 확률 값들 및 배열 순서와 일치할 확률 값을 산출한다.
- [0072] 이는 추출된 특징 신호가 어떠한 부표정에 관한 특징 신호 패턴들을 어떠한 배열 순서로 포함하고 있는지를 판단함으로써, 모델링된 대화형 표정들 중에서 어디에 더 가까운지를 판단하는 과정에 해당한다.
- [0073] 앞서 설명한 예에서와 같이, 추출된 특징 신호에 α 가 70%, β 가 70%, γ 80%, δ 가 10% 일치할 확률로 포함되어 있고, 이들의 배열 순서가 α , β , γ , δ 순으로 나타나는 경우를 예로 들면, 다음과 같이 모델링된 대화형 표정과의 일치할 확률 값이 산출될 수 있다.
- [0074] 즉, 모델링된 대화형 표정 중 어느 하나의 대화형 표정이 미리 설정된 특징 신호에 관한 패턴들을 α , β , γ , δ 순으로 각각 70%, 70%, 80%, 10% 이상 포함하는 조합으로 이루어지는 대화형 표정인 경우, 제2 산출부(124)는 100%의 확률 값을 산출할 수 있다.
- [0075] 이때, 모델링된 대화형 표정 중 다른 하나의 대화형 표정이 미리 설정된 특징 신호에 관한 패턴들을 β , α , γ , δ 순으로 각각 70%, 70%, 80%, 10% 이상 포함하는 조합으로 이루어지는 대화형 표정인 경우에는, 비록, 각각의 미리 설정된 특징 패턴들이 포함된 확률은 그대로이나, 배열 순서가 바뀌었으므로, 제2 산출부(124)는 100%보다 낮은 확률 값을 산출하게 된다.
- [0076] 다음으로, 단계(S626)에서 판단부(126)는 단계(S624)에서 산출된 적어도 하나의 확률 값 중 가장 높은 값으로 일치하는 대화형 표정을 단계(S610)에서 추출된 특징 신호에 대응하는 것으로 판단한다.
- [0077] 앞서 설명한 예에서, 판단부(126)는 가장 높은 확률 값으로 일치하는 어느 하나의 대화형 표정을 추출된 특징 신호에 대응하는 것으로 판단하며, 이러한 단계(S626)에 의해, 가장 높은 확률로 일치하는 대화형 표정이 입력 영상에 대응하는 대화형 표정으로 결정될 수 있다.
- [0078] 상기의 단계들(S622 내지 S626)에 따라, 추출된 특징 신호에 대응하는 대화형 표정이 결정되게 되면, 대화형 표정 인식 장치(100)는 대응하는 대화형 표정을 입력 영상을 인식한 결과로서 출력한다.
- [0079] 이와 같이, 본 발명에 따르면, 복잡한 움직임으로 구성되고 이에 따라 긴 시퀀스 길이를 갖게 되는 대화형 표정을 보다 신속하고 정확하게 인식할 수 있다.
- [0080] 이는 대화 과정에서 나타나는 의사의 표현이나 보조를 위하여 사용되는 다양한 얼굴 표정의 인식을 가능하게 함으로써, 사람과 기계 사이에 보다 자유로운 인터랙션이 이루어질 수 있도록 한다.
- [0081] 또한, 본 발명의 실시예들은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 본 발명의 일 실시예들의 동작을 수행하기 위해 적어도 하나의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0082] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및

변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

100 : 대화형 표정 인식 장치

110 : 추출부

120 : 결정부

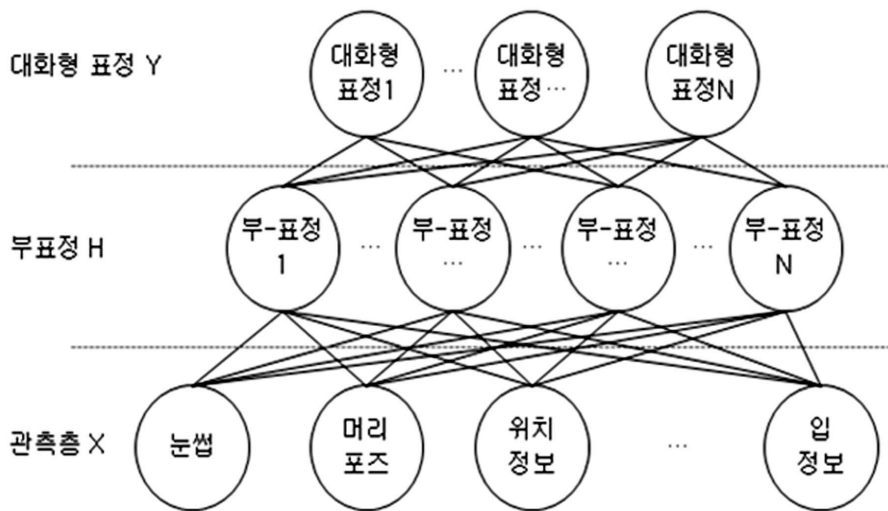
122 : 제1 산출부

124 : 제2 산출부

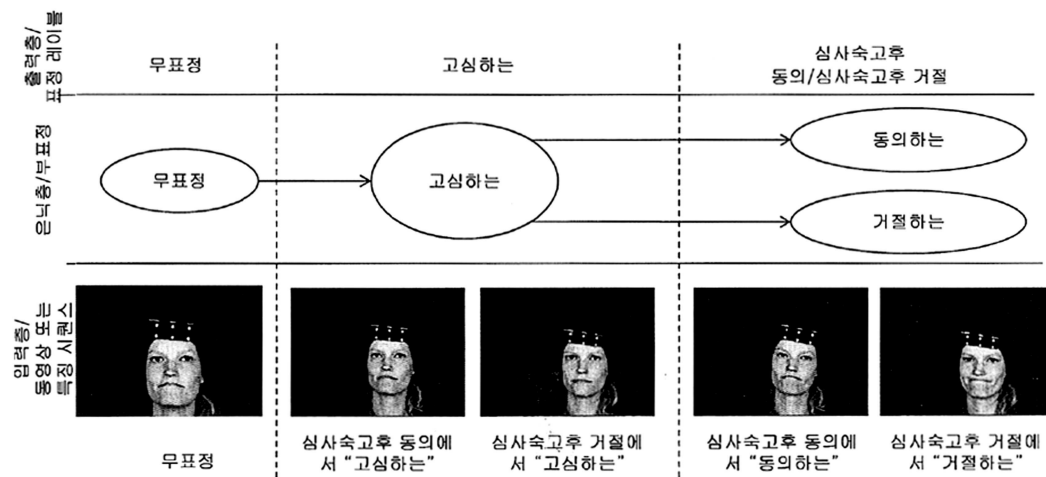
126 : 판단부

도면

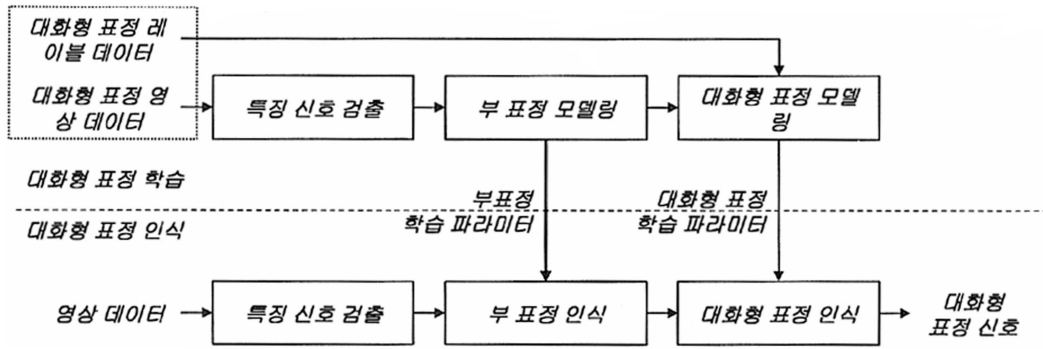
도면1



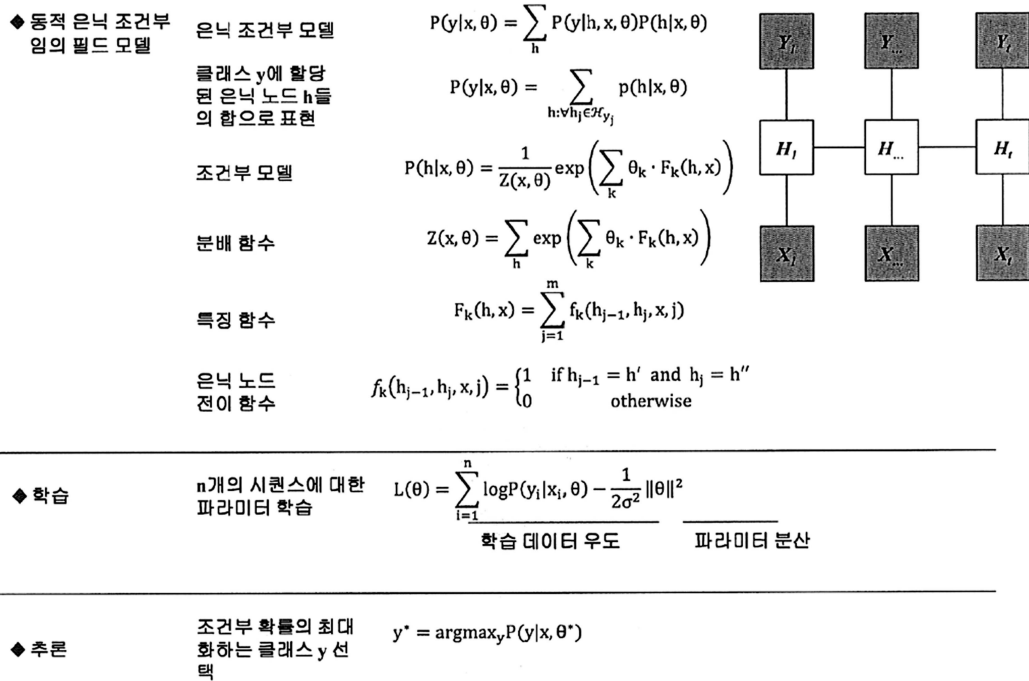
도면2



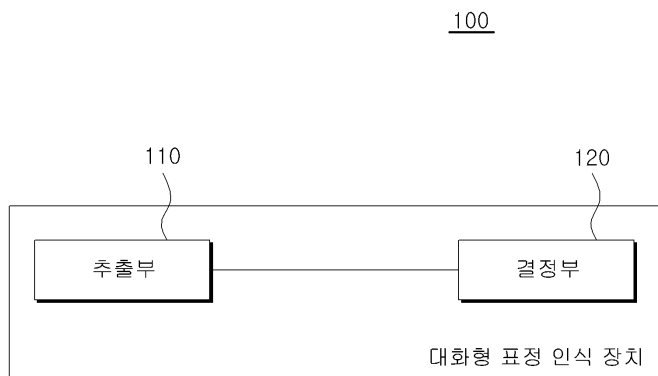
도면3



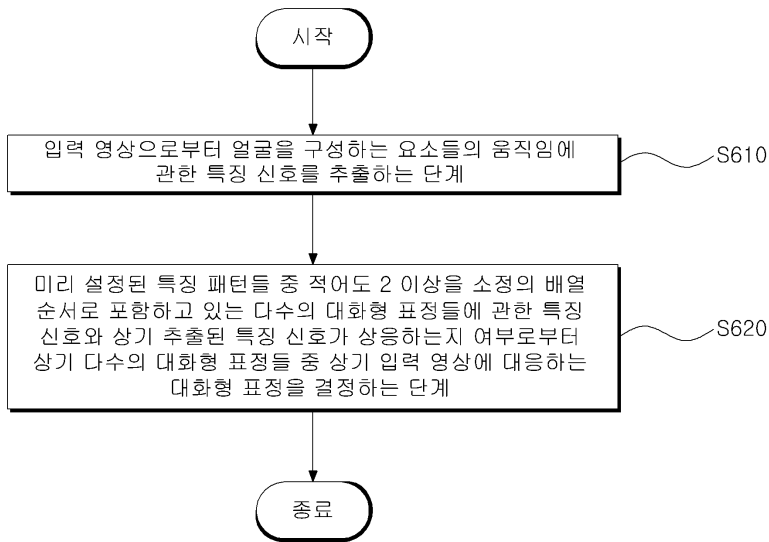
도면4



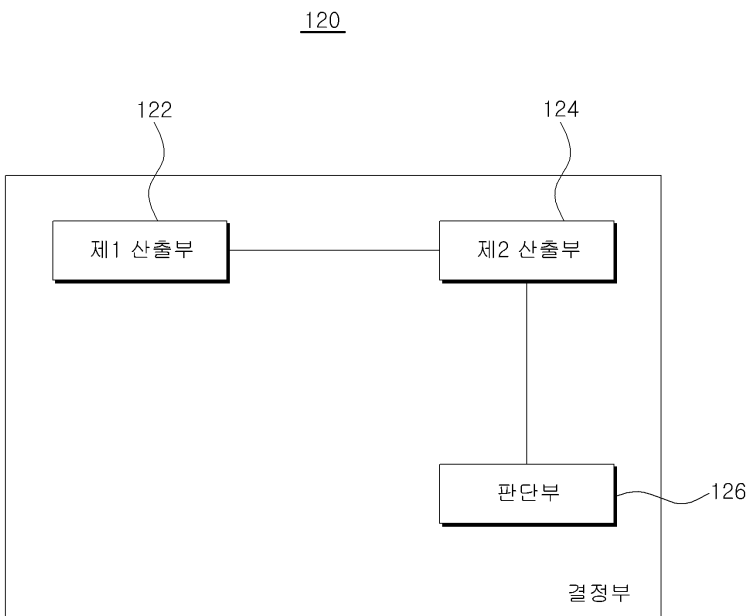
도면5



도면6



도면7



도면8

