



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2020년12월04일

(11) 등록번호 10-2186767

(24) 등록일자 2020년11월30일

(51) 국제특허분류(Int. Cl.)

G06K 9/00 (2006.01) G06K 9/48 (2006.01)

G06N 3/04 (2006.01) G06N 3/08 (2006.01)

G06T 5/20 (2006.01) G06T 7/11 (2017.01)

(52) CPC특허분류

G06K 9/00268 (2013.01)

G06K 9/481 (2013.01)

(21) 출원번호 10-2018-0049414

(22) 출원일자 2018년04월27일

심사청구일자 2018년04월27일

(65) 공개번호 10-2019-0134865

(43) 공개일자 2019년12월05일

(56) 선행기술조사문헌

US20110091115 A1*

US20160379044 A1*

KR101449744 B1*

KR1020180025093 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상윤

서울특별시 서초구 청계산로7길 43, 504동 1302호(신원동, 서초포레스타5단지)

박성주

서울특별시 서대문구 독립문길 12-24 401호
(뒷면에 계속)

(74) 대리인

민영준

전체 청구항 수 : 총 7 항

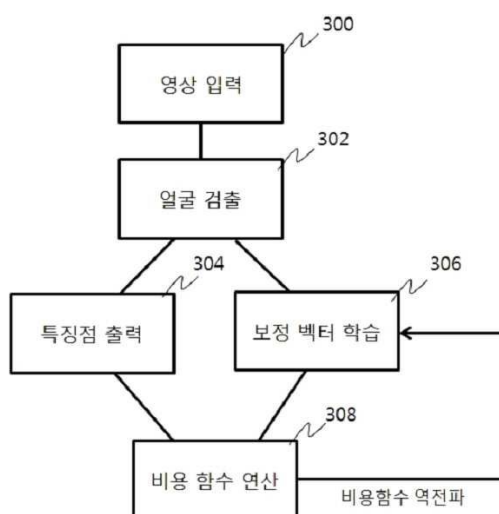
심사관 : 윤지영

(54) 발명의 명칭 학습을 이용한 얼굴 특징점 검출 방법 및 장치

(57) 요약

학습을 이용한 얼굴 특징점 검출 방법 및 장치가 개시된다. 개시된 장치는, 대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성하는 얼굴 검출부; 상기 얼굴 검출 영상에 대해 미리 설정된 알고리즘을 이용하여 복수의 특징점을 출력하는 특징점 출력부; 상기 특징점 출력부에서 출력하는 복수의 특징점의 좌표를 보정하도록 학습되어 있으며 상기 얼굴 검출 영상을 입력받아 상기 복수의 특징점의 좌표를 보정하기 위한 보정 벡터를 출력하는 보정 벡터 출력 네트워크; 및 상기 보정 벡터를 상기 복수의 특징점별로 반영하여 최종적인 특징점을 결정하는 특징점 결정부를 포함한다. 개시된 장치 및 방법에 의하면, 비교적 적은 양의 학습 데이터로도 효과적으로 특징점을 검출할 수 있는 장점이 있다.

대표도 - 도3



(52) CPC특허분류

G06N 3/04 (2013.01)
G06N 3/08 (2013.01)
G06T 5/20 (2013.01)
G06T 7/11 (2017.01)
G06T 2207/20081 (2013.01)
G06T 2207/20084 (2013.01)
G06T 2207/30201 (2013.01)

도진경

서울특별시 송파구 올림픽로4길 42, 2동 307호(잠실동, 우성아파트)

(72) 발명자

전태재

인천광역시 연수구 컨벤시아대로130번길 80, 102동 302호(송도동, 송도 푸르지오 하버뷰)

배한별

서울특별시 강남구 언주로29길 34, 502동 907호(도곡동, 우성5차아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	2016-0-00562
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터(NIPA산하)
연구사업명	정보통신방송연구개발사업
연구과제명	[이지바로][주관/한국과학기술원]상대방의 감성을 추론, 판단하여 그에 맞추어 대화
하고 대응할 수 있는 감성지능 기술 연구개발(2/5)	
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2017.09.01 ~ 2018.06.30

명세서

청구범위

청구항 1

대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성하는 얼굴 검출부;

상기 얼굴 검출 영상에 대해 미리 설정된 알고리즘을 이용하여 복수의 특징점을 출력하는 특징점 출력부;

상기 특징점 출력부에서 출력하는 복수의 특징점의 좌표를 보정하도록 학습되어 있으며 상기 얼굴 검출 영상을 입력받아 상기 복수의 특징점의 좌표를 보정하기 위한 보정 벡터를 출력하는 보정 벡터 출력 네트워크; 및

상기 보정 벡터를 상기 복수의 특징점별로 반영하여 최종적인 특징점을 결정하는 특징점 결정부를 포함하되,

상기 미리 설정된 알고리즘은 핸드 크래프트 알고리즘이고,

상기 보정 벡터 출력 네트워크는 합성곱 신경망(Convolutional Neural Network)를 포함하며, 컨볼루션 연산을 적용하는 필터의 계수를 갱신하도록 학습되며,

상기 보정 벡터 출력 네트워크는 학습 단계에서 레퍼런스 영상의 특징점 참값과 미리 설정된 알고리즘에 의해 검출되는 특징점에 상기 보정 벡터 출력 네트워크의 출력인 보정 벡터를 반영한 특징점 좌표의 차에 상응하는 비용 함수를 역전파받아 상기 필터의 계수를 갱신하고,

상기 필터의 계수의 갱신은 상기 레퍼런스 영상의 참값과 상기 보정 벡터 출력 네트워크에서 출력한 보정 벡터를 상기 미리 설정된 알고리즘에 의해 검출되는 특징점에 반영한 특징점 좌표와의 차이가 최소화되는 방향으로 이루어지는 것을 특징으로 하는 학습을 이용한 특징점 검출 장치.

청구항 2

삭제

청구항 3

제1 항에 있어서,

상기 보정 벡터는 상기 복수의 특징점 좌표 각각을 x축으로 이동시키기 위한 Δx 성분 및 y축으로 이동시키기 위한 Δy 성분을 포함하는 것을 특징으로 하는 학습을 이용한 특징점 검출 장치.

청구항 4

삭제

청구항 5

삭제

청구항 6

제1항에 있어서,

상기 얼굴 검출부는 HOG(Histogram of Gradient)+SVM(Support Vector Machine) 방식에 의해 얼굴 영역을 검출하는 것을 특징으로 하는 학습을 이용한 특징점 검출 장치.

청구항 7

대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성하는 단계(a);

상기 얼굴 검출 영상에 대해 미리 설정된 알고리즘을 이용하여 복수의 특징점을 출력하는 단계(b);

상기 단계(b)에서 출력하는 복수의 특징점의 좌표를 보정하도록 학습되어 있는 보정 벡터 출력 네트워크를 이용하여 상기 얼굴 검출 영상을 입력받아 상기 복수의 특징점의 좌표를 보정하기 위한 보정 벡터를 출력하는 단계(c); 및

상기 보정 벡터를 상기 복수의 특징점별로 반영하여 최종적인 특징점을 결정하는 단계(d)를 포함하되,

상기 미리 설정된 알고리즘은 핸드 크래프트 알고리즘이고,

상기 보정 벡터 출력 네트워크는 합성곱 신경망(Convolutional Neural Network)를 포함하며, 컨볼루션 연산을 적용하는 필터의 계수를 갱신하도록 학습되며,

상기 보정 벡터 출력 네트워크는 학습 단계에서 레퍼런스 영상의 특징점 참값과 미리 설정된 알고리즘에 의해 검출되는 특징점에 상기 보정 벡터 출력 네트워크의 출력인 보정 벡터를 반영한 특징점 좌표의 차에 상응하는 비용 함수를 역전파받아 상기 필터의 계수를 갱신하고,

상기 필터의 계수의 갱신은 상기 레퍼런스 영상의 참값과 상기 보정 벡터 출력 네트워크에서 출력한 보정 벡터를 상기 미리 설정된 알고리즘에 의해 검출되는 특징점에 반영한 특징점 좌표와의 차이가 최소화되는 방향으로 이루어지는 것을 특징으로 하는 학습을 이용한 특징점 검출 방법.

청구항 8

삭제

청구항 9

제7 항에 있어서,

상기 보정 벡터는 상기 복수의 특징점 좌표 각각을 x축으로 이동시키기 위한 Δx 성분 및 y축으로 이동시키기 위한 Δy 성분을 포함하는 것을 특징으로 하는 학습을 이용한 특징점 검출 방법.

청구항 10

삭제

청구항 11

삭제

청구항 12

제7항에 있어서,

상기 단계(a)는 HOG(Histogram of Gradient)+SVM(Support Vector Machine) 방식에 의해 얼굴 영역을 검출하는 것을 학습을 이용한 특징점 검출 방법

청구항 13

제7항의 방법을 실행하기 위한 프로그램이 유형적으로 기록되어 있으며 컴퓨터에 의해 판독 가능한 프로그램이 기록된 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 얼굴의 특징점 검출 방법 및 장치에 관한 것으로서, 더욱 상세하게는 학습을 이용한 얼굴 특징점 검출 방법 및 장치에 관한 것이다.

배경 기술

[0003] 얼굴 인식 및 얼굴 영상을 위한 다양한 응용을 위해 얼굴의 특징점 검출이 이루어진다. 특징점은 얼굴의 형상을 표현할 수 있는 지점을 선택하여 검출되며, 얼굴의 윤곽선, 눈, 코, 입과 같은 영역에 주로 지정된다. 정확한 얼굴 인식을 위해서 정확한 특징점의 검출은 매우 중요한 사항이다.

[0004] 종래에 있어서 특징점의 검출은 주로 알고리즘에 의해 이루어졌으며 가장 일반적인 알고리즘으로는 핸드 크래프트(Hand-craft) 알고리즘이 있다. 그런데, 알고리즘에 의한 특징점 검출은 얼굴 영상의 특정 포즈에서는 얼굴과 전혀 관련 없는 영역에 특징점이 검출되는 big failure가 종종 발생하는 문제점이 있었다.

[0005] 이와 같은 big failure의 발생을 방지하기 위해 신경망을 이용하여 학습에 의해 특징점을 검출하는 방법이 연구되었다.

[0006] 이러한 신경망 학습 모델을 입력 데이터로 얼굴 영상이 입력되고, 미리 알고 있는 특징점 참값(Ground Truth)과 신경망이 출력하는 특징점의 차이를 역전파하는 방식으로 학습이 이루어졌다.

[0007] 신경망을 이용한 특징점 검출은 기존의 알고리즘에 의한 특징점 검출에 비해 big failure의 발생은 저하시킬 수 있는 장점이 있으나 안정적인 특징점 검출이 이루어지려면 충분한 학습이 이루어져야 했다. 결국, 신경망 학습 모델을 이용하여 특징점을 검출하려면 충분한 학습 데이터를 확보한 후 신경망 학습을 수행하여야 했다.

[0008] 그런데, 얼굴의 특징점의 참값은 일일이 수작업에 의해 지정되어야 하기에 충분한 학습 데이터를 확보하기 어려운 문제점이 있다.

[0009] 따라서, 기존의 학습을 이용한 특징점 검출 방식은 현재까지 알려진 학습 데이터를 이용하여 안정적으로 특징점을 검출하기 어려운 문제점이 있었다.

발명의 내용

해결하려는 과제

[0011] 본 발명은 비교적 적은 양의 학습 데이터로도 효과적으로 특징점을 검출할 수 있는 학습을 이용한 특징점 검출 장치 및 방법을 제안한다.

과제의 해결 수단

[0013] 상기와 같은 목적을 달성하기 위해, 본 발명의 일 측면에 따르면, 대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성하는 얼굴 검출부; 상기 얼굴 검출 영상에 대해 미리 설정된 알고리즘을 이용하여 복수의 특징점을 출력하는 특징점 출력부; 상기 특징점 출력부에서 출력하는 복수의 특징점의 좌표를 보정하도록 학습되어 있으며 상기 얼굴 검출 영상을 입력받아 상기 복수의 특징점의 좌표를 보정하기 위한 보정 벡터를 출력하는 보정 벡터 출력 네트워크; 및 상기 보정 벡터를 상기 복수의 특징점별로 반영하여 최종적인 특징점을 결정하는 특징점 결정부를 포함하는 학습을 이용한 특징점 검출 장치가 제공된다.

[0014] 상기 미리 설정된 알고리즘은 핸드 크래프트 알고리즘을 포함한다

[0015] 상기 보정 벡터는 상기 복수의 특징점 좌표 각각을 x축으로 이동시키기 위한 Δx 성분 및 y축으로 이동시키기 위한 Δy 성분을 포함한다.

[0016] 상기 보정 벡터 출력 네트워크는 합성곱 신경망(Convolutional Neural Network)를 포함하며, 컨볼루션 연산을 적용하는 필터의 계수를 갱신하도록 학습된다.

[0017] 상기 보정 벡터 출력 네트워크는 학습 단계에서 레퍼런스 영상의 특징점 참값과 미리 설정된 알고리즘에 의해 검출되는 특징점에 상기 보정 벡터 출력 네트워크의 출력인 보정 벡터를 반영한 특징점 좌표의 차에 상응하는 비용 함수를 역전파받아 상기 필터의 계수를 갱신한다.

[0018] 상기 얼굴 검출부는 HOG(Histogram of Gradient)+SVM(Support Vector Machine) 방식에 의해 얼굴 영역을 검출

한다.

[0019] 본 발명의 다른 측면에 따르면, 대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성하는 단계(a); 상기 얼굴 검출 영상에 대해 미리 설정된 알고리즘을 이용하여 복수의 특징점을 출력하는 단계(b); 상기 단계(b)에서 출력하는 복수의 특징점의 좌표를 보정하도록 학습되어 있는 보정 벡터 출력 네트워크를 이용하여 상기 얼굴 검출 영상을 입력받아 상기 복수의 특징점의 좌표를 보정하기 위한 보정 벡터를 출력하는 단계(c); 및 상기 보정 벡터를 상기 복수의 특징점별로 반영하여 최종적인 특징점을 결정하는 단계(d)를 포함하는 학습을 이용한 특징점 검출 방법이 제공된다.

발명의 효과

[0021] 본 발명에 의하면, 비교적 적은 양의 학습 데이터로도 효과적으로 특징점을 검출할 수 있는 장점이 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 일 실시예에 따른 학습을 이용한 얼굴 특징점 검출 장치의 개략적인 구조를 도시한 블록도.
 도 2는 핸드-크래프트 알고리즘에 의한 특징점 검출 시 발생하는 big failure의 일례를 나타낸 도면.
 도 3은 본 발명의 일 실시예에 따른 보정 벡터 출력 네트워크의 학습 구조를 나타낸 도면.
 도 4는 본 발명의 일 실시예에 따른 보정 벡터 네트워크의 상세 구조를 나타낸 도면.
 도 5는 본 발명의 일 실시예에 따른 얼굴 특징점 검출 방법의 전체적인 흐름을 도시한 순서도.

발명을 실시하기 위한 구체적인 내용

[0024] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태로 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.

[0025] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.

[0026] 도 1은 본 발명의 일 실시예에 따른 학습을 이용한 얼굴 특징점 검출 장치의 개략적인 구조를 도시한 블록도이다.

[0027] 도 1을 참조하면, 본 발명의 일 실시예에 따른 학습을 이용한 얼굴 특징점 검출 장치는 영상 입력부(100), 얼굴 검출부(110), 핸드-크래프트 특징점 출력부(120), 보정 벡터 출력 네트워크(130) 및 특징점 결정부(140)를 포함한다.

[0028] 본 발명의 특징점 검출 장치는 얼굴 영상을 입력 받아 특징점을 검출하는 장치이다. 특징점은 얼굴의 형태를 파악하기 위한 점들로서 복수의 특징점이 얼굴의 주요 영역에 검출된다. 예를 들어, 얼굴의 윤곽선 영역, 눈 영역, 코 영역, 입 영역 등과 같이 얼굴의 형상을 파악할 수 있는 중요한 부분에 특징점이 지정된다.

[0029] 검출되는 특징점의 개수는 미리 설정되며, 예를 들어 68개의 특징점이 입력 영상으로부터 검출될 수 있다.

[0030] 본 발명에서 제안되는 얼굴 특징점 검출 장치는 기존의 핸드-크래프트 알고리즘을 이용한 특징점 검출과 학습을 이용한 특징점 검출이 함께 이용된다 다만, 본 발명에서 학습 대상은 특징점이 아니라 보정 벡터이며, 보정 벡터의 상세한 의미에 대해서는 추후 설명하기로 한다.

[0031] 영상 입력부(100)는 특징점을 검출할 대상 영상을 입력받는다. 대상 영상은 얼굴이 촬영된 영상으로서, 얼굴과 배경이 함께 존재하는 영상이다.

[0032] 얼굴 검출부(102)는 영상 입력부로 입력된 대상 영상으로부터 얼굴 영역을 검출하는 기능을 한다. 대상 영상으로부터 배경 영역을 제거하고 얼굴 영역만을 검출하는 것이다. 영상으로부터 배경을 제거하고 얼굴 영역만을 검출하는 기술로 다양한 기술들이 알려져 있으며 어떠한 방법에 의해 얼굴 영역을 검출하여도 무방하다.

[0033] 본 발명의 일 실시예에 따르면, HOG(Histogram of Gradient)+SVM(Support Vector Machine) 방법을 이용하여 얼굴 영역을 검출할 수 있다. HOG는 일정 크기의 셀 내부에서 모든 픽셀 값의 그래디언트(Gradient)와 방향(Orientation)을 계산한뒤 이 값들을 사용하여 히스토그램을 생성하며, 이를 SVM의 특징 벡터로 이용하여 얼굴

영역을 검출하는 방법이다.

- [0034] 물론, 전술한 바와 같이, HOG+SVM 방법 이외에도 다양한 얼굴 검출 방법이 사용될 수 있을 것이다.
- [0035] 얼굴 검출부(110)는 검출된 얼굴 영역만이 존재하고 배경이 제거된 얼굴 검출 영상을 출력한다.
- [0036] 핸드-크래프트 특징점 출력부(120)는 얼굴 검출부(110)에서 출력되는 얼굴 영상에 대해 핸드-크래프트 알고리즘을 이용하여 해당 얼굴 영상의 특징점을 출력한다. 핸드-크래프트 알고리즘은 특징점 추출을 위한 변수와 해당 변수의 연산을 통해 특징점을 검출하는 방식이며 예를 들어, TREE 알고리즘이 있다. TREE 알고리즘은 random ferm에 의해 추출된 특징을 regression tree 분류기를 사용하여 수렴 속도를 빠르게 하는 알고리즘이다. TREE 알고리즘은 캐스케이드(Cascade) 방식으로 여러 번 보정이 가능한 장점이 있는 알고리즘이다. 물론, TREE 알고리즘 이외에도 다양한 방식의 특징점 추출 알고리즘이 사용될 수 있다는 점은 당업자라면 용이하게 이해할 수 있을 것이다.
- [0037] 핸드-크래프트 알고리즘은 우수한 성능을 가지는 알고리즘이기는 하나 사용자의 포즈에 따라 big failure를 유발하기도 한다.
- [0038] 도 2는 핸드-크래프트 알고리즘에 의한 특징점 검출 시 발생하는 big failure의 일례를 나타낸 도면이다.
- [0039] 도 2를 참조하면, 얼굴의 윤곽선 및 눈/코/입과 관련이 없는 영역에 특징점이 지정되어 있는 것을 확인할 수 있다.
- [0040] 본 발명은 이와 같은 big failure의 발생을 최소화하고 보다 정확한 특징점 검출을 위해 학습에 의해 형성되는 네트워크인 보정 벡터 출력 네트워크를 포함한다.
- [0041] 여기서, 보정 벡터는 핸드-크래프트 특징점 출력부(120)에서 출력되는 특징점의 좌표를 보정하기 위한 벡터이다. 핸드-크래프트 특징점 출력부(120)에서 출력되는 특징점은 (x, y)와 같이 픽셀 좌표의 형태로 출력되며, 보정 벡터는 특징점의 좌표를 보정하기 위한 값의 형태를 가진다. 예를 들어, 보정 벡터는 $[\Delta x, \Delta y]$ 의 형태를 가질 수 있다.
- [0042] 여기서, Δx 는 특징점의 x축 값을 보정하기 위한 값이며, Δy 는 특징점의 y축 값을 보정하기 위한 값이다.
- [0043] 일례로, 핸드 크래프트 특징점 출력부(120)는 68개의 특징점을 출력할 수 있으며, 보정 벡터 출력 네트워크(130)는 각 68개의 특징점별로 특징점 좌표를 보정하기 위한 보정 벡터를 출력한다. 즉, 보정 벡터 출력 네트워크(130)는 각 특징점 별 보정 벡터를 출력하는 것이다.
- [0044] 보정 벡터 출력 네트워크(130)는 학습에 의해 결정되는 네트워크이며, 보정 벡터 출력 네트워크(130)의 입력 데이터는 얼굴 검출부(110)의 출력 영상이며, 출력 데이터는 보정 벡터 데이터이다. 보정 벡터 출력 네트워크의 학습 구조는 별도의 도면을 참조하여 설명하기로 한다.
- [0045] 핸드 크래프트 특징점 출력부(120)에서 출력되는 특징점 정보(특징점 좌표 정보)와 보정 벡터 출력 네트워크(130)에서 출력되는 보정 벡터 정보는 특징점 결정부(140)로 입력된다.
- [0046] 특징점 결정부(140)는 핸드 래프트 특징점 출력부(120)에서 출력되는 특징점 좌표에 보정 벡터 출력 네트워크(130)에서 출력되는 보정 벡터를 반영하여 최종적으로 특징점을 결정한다.
- [0047] 예를 들어, 제1 특징점의 픽셀 좌표가 (x1, y1)이고, 제1 특징점에 대한 보정 벡터가 $[\Delta x1, \Delta y1]$ 일 경우, 특징점 결정부(140)는 (x1+ $\Delta x1$, y1+ $\Delta y1$)을 제1 특징점에 대한 최종 좌표로 결정한다. 이러한 특징점 결정 작업은 모든 특징점에 대해 이루어진다.
- [0048] 이상과 같은 구조를 가지는 본 발명은 기존의 학습에 의한 특징점 결정 장치에 비해 작은 양의 학습 데이터로 효율적인 특징점 검출이 가능하다. 학습을 위한 특징점 데이터는 확보가 매우 어려운 측면이 있다. 따라서, 기존의 특징점을 학습에 의해 검출하는 방법의 경우 충분한 학습이 이루어지지 않아 정확한 특징점 검출이 어려운 문제점이 있었다.
- [0049] 그러나, 본 발명은 특징점 검출 자체를 학습하는 것이 아니라 특징점을 보정하기 위한 보정 벡터를 학습하는 것이므로 기존의 방법에 비해 작은 양의 학습 데이터로 정확한 특징점 검출을 가능하게 한다.
- [0050] 한편, 위에서는 1차적인 특징점 검출이 핸드 크래프트 알고리즘에 의해 이루어지는 경우에 대해 설명하였으나 알고리즘에 의한 특징점 검출이 핸드 크래프트 알고리즘에 한정되지 않는다는 점을 당업자라면 용이하게 이해할

수 있을 것이다.

- [0051] 도 3은 본 발명의 일 실시예에 따른 보정 벡터 출력 네트워크의 학습 구조를 나타낸 도면이다.
- [0052] 도 3을 참조하면, 학습을 위한 레퍼런스 영상이 선택되어 입력된다(300). 여기서, 학습을 위한 레퍼런스 영상은 얼굴 영상의 특징점 참값(Ground Truth)을 가지고 있는 영상을 의미한다. 준비된 학습 영상 중 하나를 랜덤으로 선택하여 입력할 수 있다.
- [0053] 학습을 위한 레퍼런스 영상이 선택되어 입력되면, 레퍼런스 영상으로부터 얼굴 영역을 검출한다(302). 여기서 얼굴 검출은 도 1의 얼굴 검출부(102)의 동작과 동일하게 이루어진다. 예를 들어, 얼굴 검출부(102)가 SVM+HOG 방식에 의해 얼굴 영역을 검출할 경우, 학습 단계에서의 얼굴 검출 역시 동일하게 SVM+HOG 방식을 이용하여 얼굴을 검출한다. 보정 벡터 출력 네트워크가 적용되는 장치에서의 얼굴 검출 방식과 동일한 방식으로 얼굴을 검출하면 되며, 얼굴 검출 방식에는 특별한 제한이 없으며, 알려진 어떠한 얼굴 검출 방식을 사용하여도 무방하다.
- [0054] 레퍼런스 영상에 대한 얼굴 검출 영상에 대해 핸드-크래프트 알고리즘을 이용하여 특징점을 검출한다(304). 핸드-크래프트 알고리즘 역시 보정 벡터 출력 네트워크가 적용되는 얼굴 검출 장치와 동일한 알고리즘을 이용하여 특징점을 검출한다. 앞서 설명한 바와 같이, 일례로 TREE 알고리즘이 사용될 수 있을 것이다.
- [0055] 검출된 얼굴 영상은 보정 벡터 출력 네트워크로 입력되며, 보정벡터 출력 네트워크는 얼굴 검출 영상을 이용하여 학습을 수행한다(306).
- [0056] 핸드-크래프트 알고리즘을 이용하여 출력되는 특징점에 보정 벡터 출력 네트워크에서 출력되는 보정 벡터를 반영하여 특징점을 결정하며, 결정된 특징점은 입력된 레퍼런스 영상의 특징점 참값(Ground Truth)을 이용하여 비용 함수가 연산된다(308).
- [0057] 여기서, 비용 함수는 입력된 레퍼런스 영상의 특징점 참값과 보정 벡터가 반영된 특징점과의 차에 상응하는 값이다.
- [0058] 예를 들어, 비용함수는 다음이 수학식 1과 같이 연산될 수 있다.

수학식 1

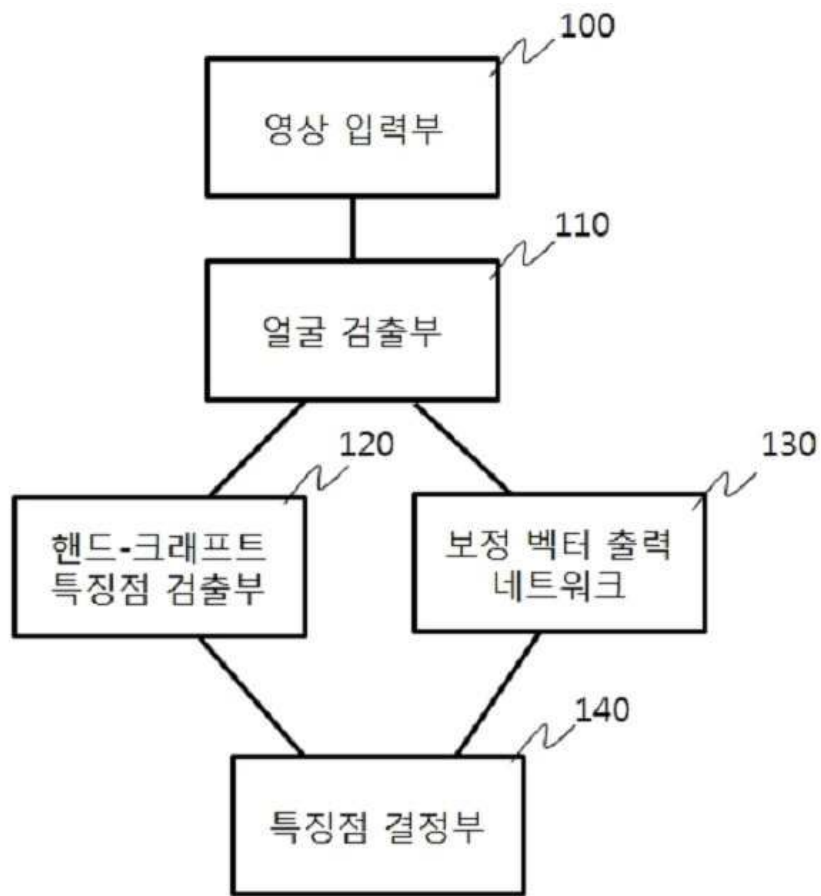
$$\frac{1}{2}[(\text{특징점 참값}) - (\text{보정벡터가 반영된 특징점})]^2$$

- [0059]
- [0060] 연산된 비용 함수는 다시 보정 벡터 출력 네트워크로 역전파(Back Propagation)된다.
- [0061] 보정 벡터 출력 네트워크는 합성곱 신경망(Convolutional Neural Network)의 형태를 가지며 보정 벡터 출력 네트워크는 역전파되는 비용함수 값에 기초하여 컨볼루션을 적용할 필터 계수를 갱신한다. 필터 계수의 갱신은 레퍼런스 영상의 참값과 보정 벡터를 반영한 특징점 좌표와의 차이가 최소화되는 방향으로 이루어진다.
- [0062] 이러한 보정 벡터 출력 네트워크의 필터 계수 갱신 작업은 미리 설정된 반복 학습 횟수(N)에 도달할 때까지 역전파되는 비용함수 값에 기초하여 반복적으로 이루어진다.
- [0063] 도 4는 본 발명의 일 실시예에 따른 보정 벡터 네트워크의 상세 구조를 나타낸 도면이다.
- [0064] 도 4를 참조하면, 입력 영상(400)에 대해 제1 컨볼루션 레이어(410)에서의 컨볼루션 필터링이 이루어진다. 일례로, 입력 영상의 사이즈가 60 X 60일 경우, 56 X 56 사이즈의 영상에 대해 컨볼루션 필터링이 이루어질 수 있다. 도 4에는 제1 컨볼루션 레이어(410)에서 20개의 필터를 적용하여 컨볼루션 필터링을 수행하는 경우가 도시되어 있다.
- [0065] 제1 컨볼루션 레이어에서 필터링이 이루어지면 제1 맥스 풀링 레이어(420)에서 사이즈를 축소하는 맥스 풀링이 이루어진다. 맥스 풀링을 위한 필터 역시 20개가 이용될 수 있다.
- [0066] 제1 맥스 풀링 레이어(420)에서 맥스 풀링이 이루어지면, 제2 컨볼루션 레이어(430)에서 컨볼루션 필터링이 이루어진다. 24 X 24 사이즈의 영상에 대해 컨볼루션 필터링이 이루어질 수 있다. 필터의 수는 일례로 48개로 설정될 수 있다.

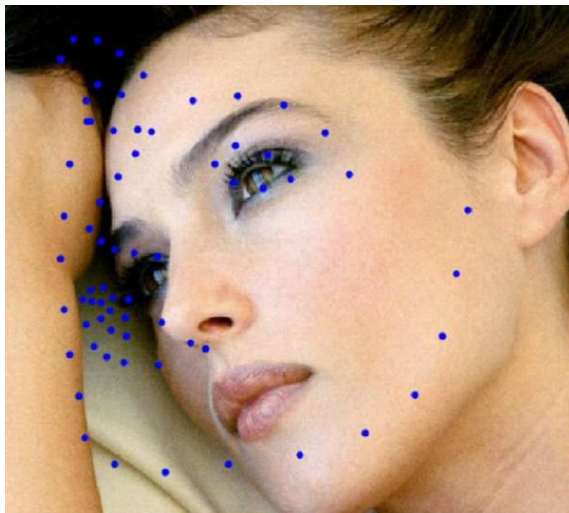
- [0067] 제2 컨볼루션 레이어(430)에서의 필터링 이후, 제2 맥스 풀링 레이어(440)에서의 맥스 풀링이 이루어져 사이즈가 축소된다.
- [0068] 이러한 컨볼루션 및 맥스 풀링은 제3 컨볼루션 레이어(450) 및 제4 맥스 풀링 레이어(460)에서도 동일하게 이루어진다.
- [0069] 제5 컨볼루션 레이어(470)에서는 3 X 3 사이즈의 데이터에 대해 80개의 필터를 적용하여 컨볼루션 필터링이 이루어진다. 제5 컨볼루션 레이어의 출력은 1차원 데이터(480)로 출력된다. 도 1에는 512개로 이루어진 1차원 데이터가 도시되어 있다.
- [0070] 512개의 1차원 데이터는 다시 136개의 1차원 데이터(490)로 필터링을 통해 축소되며, 이 데이터가 보정 벡터에 해당된다. 예를 들어, 68개의 특징점이 사용되는 경우, 보정 벡터는 각 특징점별로 두 개의 데이터인 Δx 와 Δy 를 필요로 하므로 총 136개의 데이터가 출력되는 것이다.
- [0071] 각 컨볼루션 레이어 및 맥스 풀링 레이어에 적용되는 다수의 필터의 필터 계수는 역전파되는 비용 함수에 기초하여 갱신된다.
- [0072] 도 5는 본 발명의 일 실시예에 따른 얼굴 특징점 검출 방법의 전체적인 흐름을 도시한 순서도이다.
- [0073] 도 5를 참조하면, 특징점을 검출할 대상 영상이 입력된다(단계 500). 입력 영상은 얼굴과 배경을 포함하는 영상이다.
- [0074] 대상 영상이 입력되면, 대상 영상으로부터 얼굴 영역을 검출하여 얼굴 검출 영상을 생성한다(단계 502). 얼굴 검출 영상은 대상 영상의 배경 영역에 널(Null) 값을 적용하여 배경이 제거된 영상이다.
- [0075] 대상 영상의 얼굴 검출 영상에 대해 핸드 크래프트 알고리즘을 이용하여 특징점을 검출한다(단계 504). 앞서 설명한 바와 같이, TREE 알고리즘과 같은 핸드 크래프트 알고리즘이 이용될 수 있으며, 미리 설정된 개수(예를 들어, 68개)의 특징점이 검출된다.
- [0076] 한편, 대상 영상의 얼굴 검출 영상은 보정 벡터 출력 네트워크로도 입력되며, 보정 벡터 출력 네트워크는 학습된 컨볼루션 필터를 이용하여 보정 벡터를 생성한다(단계 506). 보정 벡터는 각 특징점별로 생성된다.
- [0077] 단계 504에서 검출되는 각 특징점의 좌표는 단계 506에서 출력되는 보정 벡터를 이용하여 보정되며, 보정된 특징점을 최종적인 특징점으로 결정한다(단계 508).
- [0078] 이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래시 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록매체로서는 자기 기록매체, 광 기록매체 등이 포함될 수 있다.
- [0079] 또한, 기술적이거나 과학적인 용어를 포함한 모든 용어들은, 상세한 설명에서 다르게 정의되지 않는 한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 사전에 정의된 용어와 같이 일반적으로 사용되는 용어들은 관련 기술의 문맥상의 의미와 일치하는 것으로 해석되어야 하며, 본 발명에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0080] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구 범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

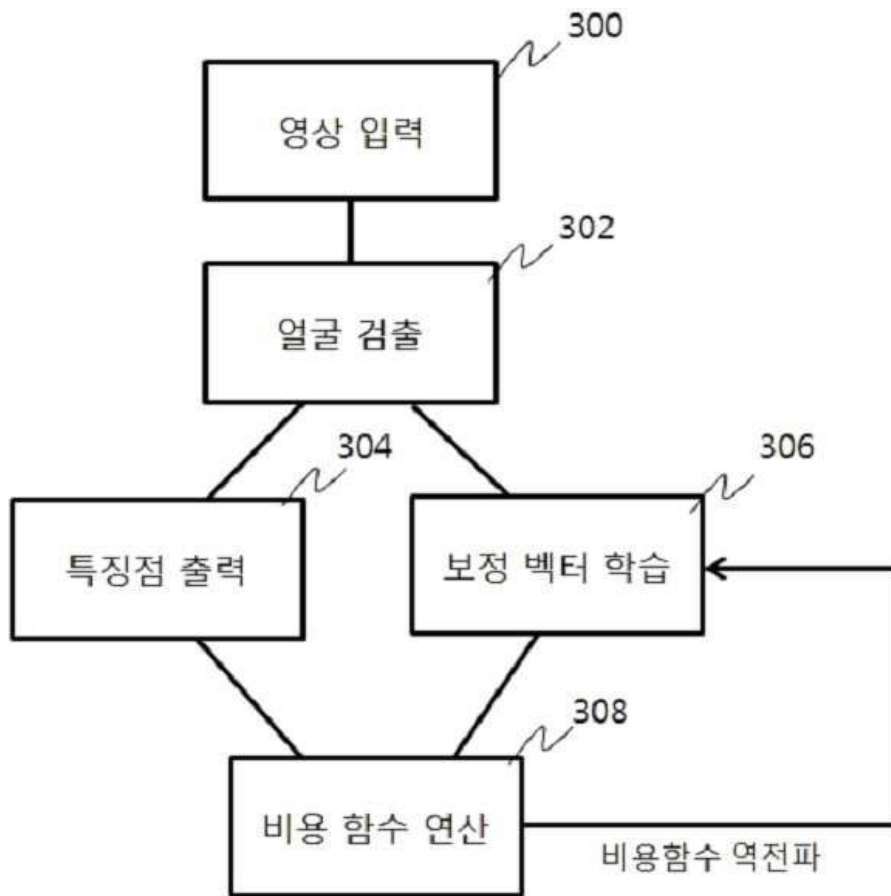
도면1



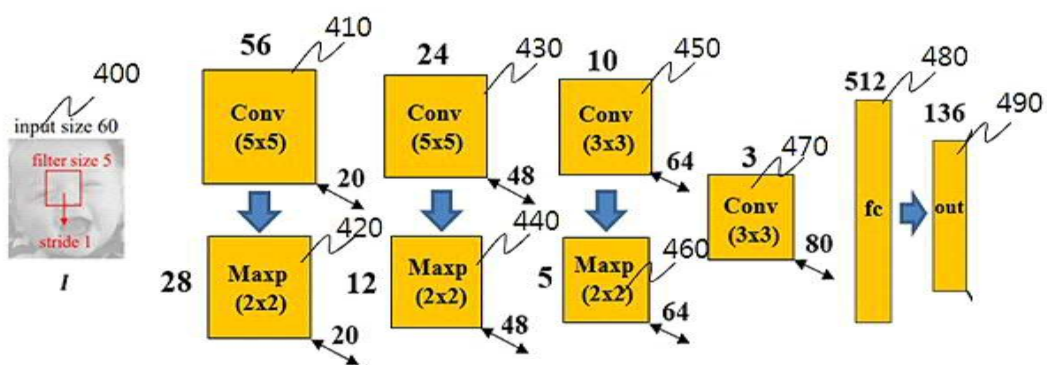
도면2



도면3



도면4



도면5

