



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년12월16일
(11) 등록번호 10-2192017
(24) 등록일자 2020년12월10일

(51) 국제특허분류(Int. Cl.)
H04N 17/02 (2006.01) G06N 3/08 (2006.01)
G06N 5/04 (2006.01)
(52) CPC특허분류
H04N 17/02 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2019-0085841
(22) 출원일자 2019년07월16일
심사청구일자 2019년07월16일
(56) 선행기술조사문헌
KR1020190063451 A*
KR101327709 B1
KR1020070061855 A
KR1020080107042 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이상훈
서울특별시 강남구 삼성로 151, 3동 504호(대치동, 선경아파트)
김재경
서울특별시 서대문구 성산로 394, 606호(창천동, JK캠퍼스)
(74) 대리인
특허법인우인

전체 청구항 수 : 총 4 항

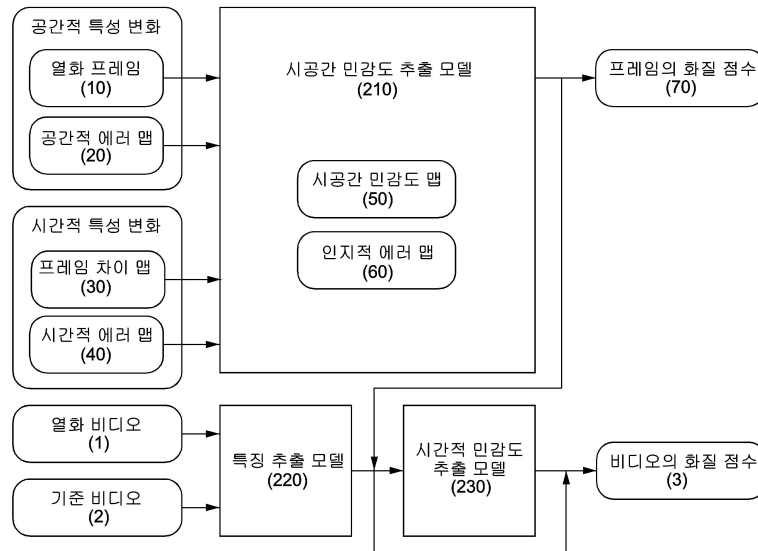
심사관 : 진민숙

(54) 발명의 명칭 인간의 시각 특성을 반영한 비디오의 화질 평가 장치 및 방법

(57) 요약

본 실시예들은 기준 비디오와 열화 비디오를 비교한 프레임 특징 집합에 프레임의 공간적 특성 변화와 시간적 특성 변화에 따른 프레임의 화질 점수를 반영하고 시간적 가중치를 적용하여 비디오의 화질 점수를 산출함으로써, 인간의 시공간적 인지 특성이 반영된 주관적 화질 평가가 가능한 비디오 화질 평가 장치를 제공한다.

대표도



(52) CPC특허분류
G06N 5/04 (2013.01)

공지예외적용 : 있음

명세서

청구범위

청구항 1

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 비디오 화질 평가 장치에 있어서,

상기 프로세서는 기준 비디오와 열화 비디오로부터 추출한 시계열 특징 집합에 프레임의 화질 점수를 반영하여 프레임별 화질 점수를 산출하고, 시간적 민감도 추출 모델을 통해 산출한 시간적 가중치를 이용하여 비디오의 화질 점수를 산출하며,

상기 시간적 민감도 추출 모델은 메모리 커널을 갖는 메모리 컨볼루션 네트워크를 포함하며, 상기 프레임별 화질 점수의 특징 영역에 집중하는 상기 메모리 커널을 통해 상기 프레임별 화질 점수의 분포를 분석하여 시계열적 패턴을 도출하고, 상기 시계열적 패턴에 따라 상기 시간적 가중치를 산출하는 것을 특징으로 하는 비디오 화질 평가 장치.

청구항 2

제1항에 있어서,

상기 프로세서는 프레임의 공간적 특성 변화와 시간적 특성 변화를 기반으로 시공간 민감도 추출 모델을 통해 상기 프레임의 화질 점수를 산출하는 것을 특징으로 하는 비디오 화질 평가 장치.

청구항 3

제2항에 있어서,

상기 시공간 민감도 추출 모델은 상기 프레임의 공간적 특성 변화가 반영된 열화 프레임 및 공간적 에러 맵을 입력받고, 상기 프레임의 시간적 특성 변화가 반영된 프레임 차이 맵 및 시간적 에러 맵을 입력받고, 상기 공간적 특성 변화와 시간적 특성 변화를 학습하여 시공간 민감도 맵을 재구성하고, 상기 시공간 민감도 맵과 상기 공간적 에러 맵을 연산하여 인지적 에러 맵을 예측하고, 상기 인지적 에러 맵으로부터 상기 프레임의 화질 점수를 산출하는 것을 특징으로 하는 비디오 화질 평가 장치.

청구항 4

제3항에 있어서,

상기 프로세서는 상기 프레임을 복수의 패치로 분할하는 전처리 과정을 수행한 후 상기 패치에 대해서 상기 시공간 민감도 맵을 재구성하고, 상기 패치와 상기 인지적 에러 맵 간의 크기 비율을 고려하여 슬라이딩 윈도우를 설정하여 상기 인지적 에러 맵의 일부 영역의 중첩을 방지하는 것을 특징으로 하는 비디오 화질 평가 장치.

청구항 5

삭제

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 비디오의 화질 평가 장치 및 방법에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 네트워크 기술의 발달 및 서버 확장에 따른 미디어 플랫폼 산업 규모가 증가하는 추세이다. 스마트 디바이스의 보급으로 사용자들은 모바일 인터넷을 통해 다양한 콘텐츠에 언제든지 접근 가능하다.

- [0004] 비디오 스트리밍 서비스는 양적으로 증가하고 있고, 예측 곤란한 통신 네트워크의 상태를 고려하여 양질의 비디오를 공급할 필요가 있다.
- [0005] 기존의 비디오 화질 평가 방법은 공간적 에러 신호에 따른 지역적 왜곡을 중심으로 화질을 평가한다. 비디오 콘텐츠의 화질 평가는 인간의 시각 특성에 기반한 인지적 요소를 반영할 필요가 있다.

선행기술문헌

특허문헌

- [0006] (특허문헌 0001) 한국등록특허공보 제10-0434162호 (2004.05.24.)

발명의 내용

해결하려는 과제

- [0007] 본 발명의 실시예들은 기존 비디오와 열화 비디오를 비교한 프레임 특징 집합에 프레임의 공간적 특성 변화와 시간적 특성 변화에 따른 프레임의 화질 점수를 반영하고 시간적 가중치를 적용하여 비디오의 화질 점수를 산출함으로써, 인간의 시공간적 인지 특성이 반영된 주관적 화질 평가를 가능하게 하는데 주된 목적이 있다.
- [0008] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

- [0009] 본 실시예의 일 측면에 의하면, 하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 비디오 화질 평가 장치에 있어서, 상기 프로세서는 기존 비디오와 열화 비디오로부터 추출한 시계열 특징 집합에 프레임의 화질 점수를 반영하여 프레임별 화질 점수를 산출하고, 시간적 민감도 추출 모델을 통해 산출한 시간적 가중치를 이용하여 비디오의 화질 점수를 산출하는 것을 특징으로 하는 비디오 화질 평가 장치를 제공한다.

발명의 효과

- [0010] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 기존 비디오와 열화 비디오를 비교한 프레임 특징 집합에 프레임의 공간적 특성 변화와 시간적 특성 변화에 따른 프레임의 화질 점수를 반영하고 시간적 가중치를 적용하여 비디오의 화질 점수를 산출함으로써, 인간의 시공간적 인지 특성이 반영된 주관적 화질 평가를 할 수 있는 효과가 있다.
- [0011] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

- [0012] 도 1은 시간적 화질 변화에 따른 비디오 화질 평가 점수를 예시한 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치를 예시한 블록도이다.
- 도 3은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 공간적 특성 및 시간적 특성이 반영된 맵을 예시한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치의 학습 모델을 예시한 블록도이다.
- 도 5는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치의 전체 동작을 예시한 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 시간적 가중치를 산출하는 동작을 예시한 도면이다.
- 도 7은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 공간적 특성 및 시간적 특성이 반영된 맵을 예시한 도면이다.

도 8은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 화질 점수 및 시간적 가중치를 예시한 도면이다.

도 9는 본 발명의 다른 실시예에 따른 비디오 화질 평가 방법을 예시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0013] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0014] 도 1은 시간적 화질 변화에 따른 비디오 화질 평가 점수를 예시한 도면이다.
- [0015] 도 1은 인간의 시공간적 인지 특성을 보여 준다. 비디오 A와 비디오 B에 대해서 전체 프레임의 화질 점수를 평균 연산한 객관적 화질 점수는 서로 동일하게 나타난다. 이와 달리 전체 프레임의 화질 점수가 동일한 비디오 A는 특정 프레임에서 화질 점수가 작은 값을 갖는 비디오 B보다 주관적 화질 점수가 높게 나타난다.
- [0016] 화질 열화가 특정 프레임에서 크게 발생하면 인간은 열화 여부를 상대적으로 더 인지하는 경향이 있고, 이미지의 색깔, 구도, 내용 등에 따라 특정 영역에서 열화 여부를 더 인지하는 경향이 있다. 따라서 비디오의 화질 열화를 평가하기 위해서는 인간의 인지 특성에 따른 공간적 시간적 민감도를 반영할 필요가 있다.
- [0017] 도 2는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치를 예시한 블록도이다. 비디오 화질 평가 장치(110)는 인간의 시공간적 인지 특성을 반영한 학습 모델을 통해 주관적 화질 점수를 산출하고, 시공간적 점수를 시각 화할 수 있다.
- [0018] 비디오 화질 평가 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장 매체(130) 및 통신 버스(170)를 포함한다.
- [0019] 프로세서(120)는 비디오 화질 평가 장치(110)가 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 비디오 화질 평가 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0020] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독한 가능 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 비디오 화질 평가 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0021] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 비디오 화질 평가 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0022] 비디오 화질 평가 장치(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 비디오 화질 평가 장치(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0023] 도 3은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 공간적 특성 및 시간적 특성이 반영된 맵을 예시한 도면이다.
- [0024] 도 3의 (a) 및 (c)는 연속하는 열화 프레임 세트이다. 도 3의 (a)는 무선 네트워크 왜곡에 의한 다수의 아티팩트를 갖는다. 비디오를 시청하는 과정에서 보행자 주변의 아티팩트를 인지하는 것은 쉽지 않다. 큰 움직임이 존재할 때, 색조, 휘도, 및 크기의 변화가 덜 시각적으로 인식되는 마스킹 효과 때문이다.
- [0025] 도 3의 (b) 및 (d)는 열화 프레임 세트의 원본 프레임 세트이다.

- [0026] 도 3의 (e)는 도 3의 (c)의 열화 프레임과 도 3의 (d)의 원본 프레임 간의 차이에 따른 공간적 에러 맵이다.
- [0027] 도 3의 (f)는 도 3의 (c)의 열화 프레임과 도 3의 (a)의 열화 프레임 간의 차이에 따른 모션 맵이다. 모션 맵은 프레임 차이 맵으로 칭할 수 있다.
- [0028] 도 3의 (g)는 도 3의 (a) 및 (c)의 열화 프레임 세트의 모션 맵과 도 3의 (b) 및 (d)의 원본 프레임 세트의 모션 맵 간의 차이에 따른 시간적 에러 맵이다. 모션 맵에서 다수의 에러가 발생하면, 공간적 에러는 더 시각적으로 인식된다. 공간적 에러와 움직임 간의 복잡한 상호 관계 때문이다.
- [0029] 도 3의 (h)는 도 3의 (c)의 열화 프레임에 대해 본 실시예에 따른 비디오 화질 평가 장치가 예측한 시공간 민감도 맵이다. 어두운 영역은 보행자의 큰 움직임에 따라 덜 민감한 것을 나타내고, 밝은 영역은 시간적 에러 요소에 따라 높은 가중치를 갖는 것을 나타낸다.
- [0030] 게다가 본 실시예에 따른 비디오 화질 평가 장치는 프레임의 화질 점수를 판단할 때, 인간의 회고적인 화질 판단 패턴을 적용한다. 이러한 인지적 패턴은 도 1에 도시되어 있다.
- [0031] 도 4는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치의 학습 모델을 예시한 블록도이다.
- [0032] 비디오 화질 평가 장치는 시공간 민감도 추출 모델(210), 특징 추출 모델(220), 및 시간적 민감도 추출 모델(230)을 포함한다.
- [0033] 시공간 민감도 추출 모델(210)은 프레임의 공간적 특성 변화가 반영된 열화 프레임(10) 및 공간적 에러 맵(20)을 입력받는다. 시공간 민감도 추출 모델(210)은 프레임의 시간적 특성 변화가 반영된 프레임 차이 맵(30) 및 시간적 에러 맵(40)을 입력받는다. 열화 프레임(10), 공간적 에러 맵(20), 프레임 차이 맵(30), 및 시간적 에러 맵(40)는 결합되어 입력될 수 있다.
- [0034] 시공간 민감도 추출 모델(210)은 공간적 특성 변화와 시간적 특성 변화를 학습하여 시공간 민감도 맵(50)을 재구성한다. 시공간 민감도 추출 모델(210)은 시공간 민감도 맵(50)과 공간적 에러 맵(20)을 연산하여 인지적 에러 맵(60)을 예측한다. 시공간 민감도 추출 모델(210)은 인지적 에러 맵으로부터 프레임의 화질 점수(70)를 산출한다.
- [0035] 시공간 민감도 추출 모델(210)은 다수의 레이어가 네트워크로 연결되며 히든 레이어를 포함한다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습 가능한 필터 집합을 포함한다. 필터는 컨볼루션 필터를 적용할 수 있다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다.
- [0036] 특징 추출 모델(220)은 열화 비디오(1)와 기준 비디오(2)를 입력받고 열화 비디오(1)와 기준 비디오(2)의 비교한 결과를 학습하여 시계열 특징 집합을 추출한다. 시계열 특징 집합은 특징 벡터로 표현될 수 있다.
- [0037] 특징 추출 모델은 다수의 레이어가 네트워크로 연결되며 히든 레이어를 포함한다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습 가능한 필터 집합을 포함한다. 필터는 컨볼루션 필터를 적용할 수 있다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다.
- [0038] 시간적 민감도 추출 모델(230)은 시계열 특징 집합을 입력받고 학습하여 시간적 가중치를 출력한다.
- [0039] 시간적 민감도 추출 모델(230)은 메모리 커널을 갖는 메모리 컨볼루션 네트워크를 포함하며, 관심 모델(Attention Model)을 적용한다.
- [0040] 도 5는 본 발명의 일 실시예에 따른 비디오 화질 평가 장치의 전체 동작을 예시한 도면이다.
- [0041] 프로세서는 프레임을 정규화하는 전처리 과정을 수행할 수 있다. 예측된 민감도는 주파수 대역에 영향을 받을 수 있기 때문이다. 그레이 스케일로 변환된 프레임을 기준으로 저주파수 대역 필터 등을 적용한 프레임을 차감하는 방식으로 프레임을 정규화할 수 있다.
- [0042] 프로세서는 프레임을 복수의 패치로 분할하는 전처리 과정을 수행할 수 있다. 패치로 분할한 후 패치에 대해서 시공간 민감도 맵을 재구성하고, 패치와 인지적 에러 맵 간의 크기 비율을 고려하여 슬라이딩 윈도우를 설정하여 인지적 에러 맵의 일부 영역의 중첩을 방지할 수 있다. 윈도우 간격은 $size_{patch} - (N_{ign} \times 2 \times R)$ 로 표현될 수 있다. $size_{patch}$ 는 패치의 크기이고, N_{ign} 는 제외된 픽셀 수이고, R은 입력된 패치와 인지적 에러 맵 간의 크기 비율을 의미한다.
- [0043] 프로세서는 프레임의 공간적 특성 변화와 시간적 특성 변화를 기반으로 시공간 민감도 추출 모델을 통해 프레임

의 화질 점수를 산출한다.

[0044] 시공간 민감도 추출 모델은 프레임의 공간적 특성 변화가 반영된 열화 프레임 및 공간적 에러 맵을 입력받고, 프레임의 시간적 특성 변화가 반영된 프레임 차이 맵 및 시간적 에러 맵을 입력받고, 공간적 특성 변화와 시간적 특성 변화를 학습하여 시공간 민감도 맵을 재구성하고, 시공간 민감도 맵과 공간적 에러 맵을 연산하여 인지적 에러 맵을 예측하고, 인지적 에러 맵으로부터 프레임의 화질 점수를 산출한다.

[0045] 공간적 에러 맵은 정규화된 로그 차이로 설정될 수 있고, 수학식 1과 같이 표현될 수 있다.

수학식 1

$$e_s^t = \frac{\log(1/((\hat{I}_r^t - \hat{I}_d^t)^2 + \epsilon/255^2))}{\log(255^2/\epsilon)}$$

[0046]

[0047] \hat{I}_r^t 는 정규화된 기준 프레임이고, \hat{I}_d^t 는 정규화된 열화 프레임이고, I_d^t 는 기준 프레임이고, I_r^t 는 계수 ϵ 는 1로 설정될 수 있다.

[0048] 프레임 차이 맵은 열화 프레임 간의 정규화된 로그 차이로 설정될 수 있다. 즉, 프레임 차이 맵은 $f_d^t = |I_d^{t+\delta} - I_d^t|$ 로 표현될 수 있다. $\delta = \lfloor fps/25 \rfloor$ 로 설정될 수 있다.

[0049] 시간적 에러 맵은 열화 프레임과 기준 프레임 간의 정규화된 로그 차이로 설정될 수 있다. 즉, 시간적 에러 맵은 $e_T^t = |f_d^t - f_r^t|$ 로 표현될 수 있다.

[0050] 시공간 민감도 맵은 컨볼루션 네트워크를 갖는 시공간 민감도 추출 모델을 통해 획득되며, 수학식 2와 같이 표현된다.

수학식 2

$$s^t = CNN_{s1}(\hat{I}_d^t, e_s^t, f_d^t, e_T^t; \theta_{s1})$$

[0051]

[0052] 시공간 민감도 맵은 계수 θ_{s1} 에 의해 조절될 수 있다.

[0053] 인지적 에러 맵은 시공간 민감도 맵과 공간적 에러 맵을 연산한다. 예컨대, 인지적 에러 맵은 $p^t = s^t \odot e_s^t$ 로 표현되며 \odot 는 원소별 곱을 의미한다.

[0054] 프레임의 화질 점수 μ_p^t 는 공간적 점수로 일부 영역의 인지적 에러 맵을 평균하여 산출된다. 프레임의 화질 점수는 수학식 3과 같이 표현된다.

수학식 3

$$\mu_p^t = \frac{1}{(H-8) \cdot (W-8)} \sum_{(i,j) \in \Omega} p^t$$

[0055]

[0056] H와 W는 인지적 에러 맵의 높이와 너비이고, (i, j)는 일부 영역 Ω의 픽셀 인덱스이다.

[0057] 공간적 점수는 평균 풀링을 통해 산출된다. 공간적 점수는 $\mu_{s1} = \sum_t \mu_{\mathbf{p}}^t$ 로 표현될 수 있다.

[0058] 시간적 민감도 추출 모델은 수학식 4로 표현된 목적 함수를 최적화할 수 있다.

수학식 4

$$\mathcal{L}_{step1}(\hat{\mathbf{I}}_d, \mathbf{e}_s, \mathbf{f}_d, \mathbf{e}_T; \theta_{s1}, \phi_1) = \|f_{\phi_1}(\mu_{s1}) - \mathbf{s}_{sub}\|_2^2 + \lambda_1 TV + \lambda_2 L_2$$

[0059]

[0060] $\hat{\mathbf{I}}_d$ 는 입력의 시퀀스이고, f()는 회귀 함수이고, \mathbf{s}_{sub} 는 열화 비디오의 주관적 점수에 대한 검증자료이다. 고주파 수 노이즈를 저감하고 오버피팅을 피하기 위해서 TV(Total Variation) 및 L_2 놈을 적용할 수 있다. λ_1 및 λ_2 는 가중치 파라미터이다.

[0061] 프로세서는 기준 비디오와 열화 비디오로부터 추출한 시계열 특징 집합에 프레임의 화질 점수를 반영하여 프레임별 화질 점수를 산출하고, 시간적 민감도 추출 모델을 통해 산출한 시간적 가중치를 이용하여 비디오의 화질 점수를 산출한다.

[0062] 도 6은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 시간적 가중치를 산출하는 동작을 예시한 도면이다.

[0063] 프로세서는 기준 비디오와 열화 비디오로부터 추출한 시계열 특징 집합에 프레임의 화질 점수를 반영하여 프레임별 화질 점수를 산출하고, 시간적 민감도 추출 모델을 통해 산출한 시간적 가중치를 이용하여 비디오의 화질 점수를 산출한다.

[0064] 시간적 민감도 추출 모델은 메모리 커널을 갖는 메모리 컨볼루션 네트워크를 포함하며, 프레임별 화질 점수의 특정 영역에 집중하는 메모리 채널을 통해 프레임별 화질 점수의 분포를 분석하여 시계열적 패턴을 도출한다. 시계열적 패턴에 따라 시간적 가중치를 산출하고, 프레임별 화질 점수에 시간적 가중치를 적용한다.

[0065] 시공간 민감도 추출 모델을 통해 산출된 전체 프레임의 화질 점수는 단일 벡터 μ_p 로 표현될 수 있다. 메모리 커널 m을 사용하여 관심 블록의 대응하는 중요도 e의 세트를 정의한다. 중요도 e는 메모리 커널 m과 단일 벡터 μ_p 의 곱으로 연산할 수 있다. 입력된 단일 벡터 μ_p 의 경계에서 제로 패딩을 수행한다.

[0066] 시간적 가중치 ω_t 는 관심 블록에 의해 추출되며, $\sum_t \omega_t = 1$ 조건에서 수학식 5와 같이 표현된다.

수학식 5

$$\omega_t = \frac{\exp(e_t)}{\sum_j \exp(e_j)}$$

[0067]

[0068] 비디오의 화질 점수는 $\mu_{s2} = \sum_t \omega_t \mu_{\mathbf{p}}^t$ 로 표현될 수 있다.

[0069] 시간적 민감도 추출 모델은 수학식 6으로 표현된 목적 함수를 최적화한다.

수학식 6

$$\mathcal{L}_{step2}(\hat{\mathbf{I}}_d, \mathbf{e}_s, \mathbf{f}_d, \mathbf{e}_T; \theta_{s1}, \phi_2) = \|f_{\phi_2}(\mu_{s2}) - \mathbf{s}_{sub}\|_2^2$$

$\hat{\mathbf{I}}_d$ 는 입력의 시퀀스이고, $f()$ 는 회귀 함수이고, \mathbf{s}_{sub} 는 열화 비디오의 주관적 점수에 대한 검증자료이다.

도 7은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 공간적 특성 및 시간적 특성이 반영된 맵을 예시한 도면이다. 도 7에서 어두운 영역은 픽셀 값이 작은 것을 나타낸다.

도 7의 (a), (f), (k), 및 (p)는 Wireless, IP, H.264, 및 MPEG-2에 해당하는 4가지 유형으로 왜곡된 열화 프레임 세트이다.

도 7의 (b), (g), (l), 및 (q)는 도 7의 (a), (f), (k), 및 (p)에 각각 대응하는 객관적 에러 맵이다.

도 7의 (c), (h), (m), 및 (r)은 도 7의 (a), (f), (k), 및 (p)에 각각 대응하는 프레임 차이 맵이다.

도 7의 (d), (i), (n), 및 (s)는 도 7의 (a), (f), (k), 및 (p)에 각각 대응하는 시간적 에러 맵이다.

도 7의 (e), (j), (o), 및 (t)는 도 7의 (a), (f), (k), 및 (p)에 각각 대응하는 예측된 시공간 민감도 맵이다.

Wireless 및 IP 왜곡에 해당하는 도 7의 (d) 및 (i)는 전체 영역에서 시간적 에러가 큰 것을 알 수 있고, Wireless 및 IP 왜곡에 해당하는 도 7의 (e) 및 (j)는 전체 영역에서 민감도가 높은 것을 알 수 있다.

H.264 및 MPEG-2에 해당하는 도 7의 (d) 및 (i)는 시간적 에러가 작은 것을 알 수 있다. 이러한 경우는 모션 정보를 갖는 프레임 차이 맵이 시공간 민감도 맵에서 지배적인 요소로 작용한다.

도 7에서 전경 객체는 비디오에서 추적되고, 도 7의 (m) 및 (r)에서 배경 영역은 객체보다 큰 값을 갖는다. 도 7의 (o) 및 (t)의 예측된 시공간 민감도 맵에서 배경 영역은 상대적으로 작은 값을 갖는다. 이러한 차이는 시간적 마스킹 효과로 인한 결과이다.

도 8은 본 발명의 일 실시예에 따른 비디오 화질 평가 장치가 처리하는 프레임의 화질 점수 및 시간적 가중치를 예시한 도면이다.

도 8의 (a)를 참조하면, 화질 점수가 5 프레임마다 상승과 하강을 반복한다. 시간적 가중치는 화질 점수와 반대된 형상으로 진행하며, 화질 점수가 낮을 때 높은 값을 갖는다. 전체적인 평균 화질보다 낮은 피크에 영향을 더 받는 인지적 요소를 반영한 결과이다.

도 8의 (b)를 참조하면, 30번째 프레임에서 화질 점수가 급격한 하강으로 진행하고 시간적 가중치는 상승하는 것을 알 수 있다. 즉, 비디오 화질 평가 장치가 산출한 시간적 가중치가 화질 점수의 변화를 적응적으로 반영하는 것을 쉽게 파악할 수 있다.

도 9는 본 발명의 다른 실시예에 따른 비디오 화질 평가 방법을 예시한 흐름도이다. 비디오 화질 평가 방법은 비디오 화질 평가 장치에 의하여 수행될 수 있다.

단계 S310에서 프로세서는 프레임의 공간적 특성 변화와 시간적 특성 변화를 기반으로 시공간 민감도 추출 모델을 통해 프레임의 화질 점수를 산출한다.

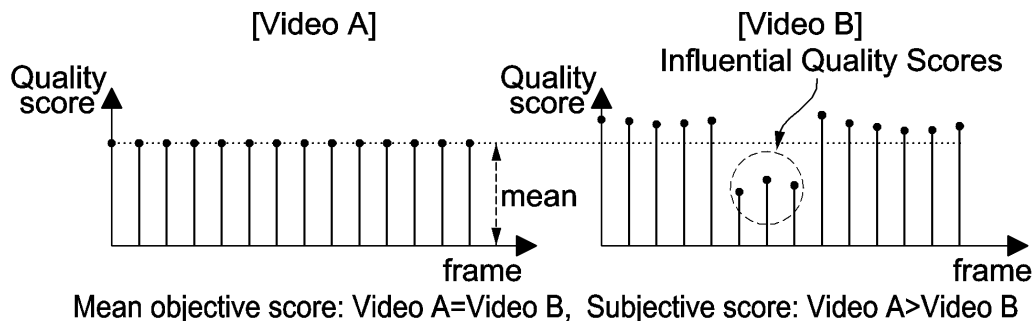
단계 S320에서 프로세서는 기준 비디오와 열화 비디오로부터 추출한 시계열 특징 집합에 프레임의 화질 점수를 반영하여 프레임별 화질 점수를 산출하고, 시간적 민감도 추출 모델을 통해 산출한 시간적 가중치를 이용하여 비디오의 화질 점수를 산출한다.

비디오 화질 평가 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그래밍 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.

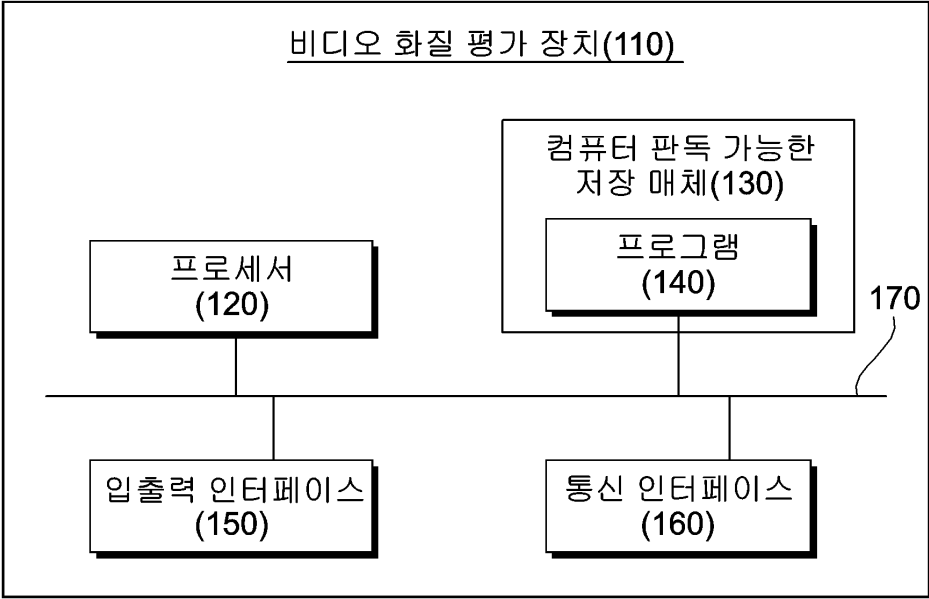
- [0088] 비디오 화질 평가 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0089] 도 9에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 9에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0090] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.
- [0091] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

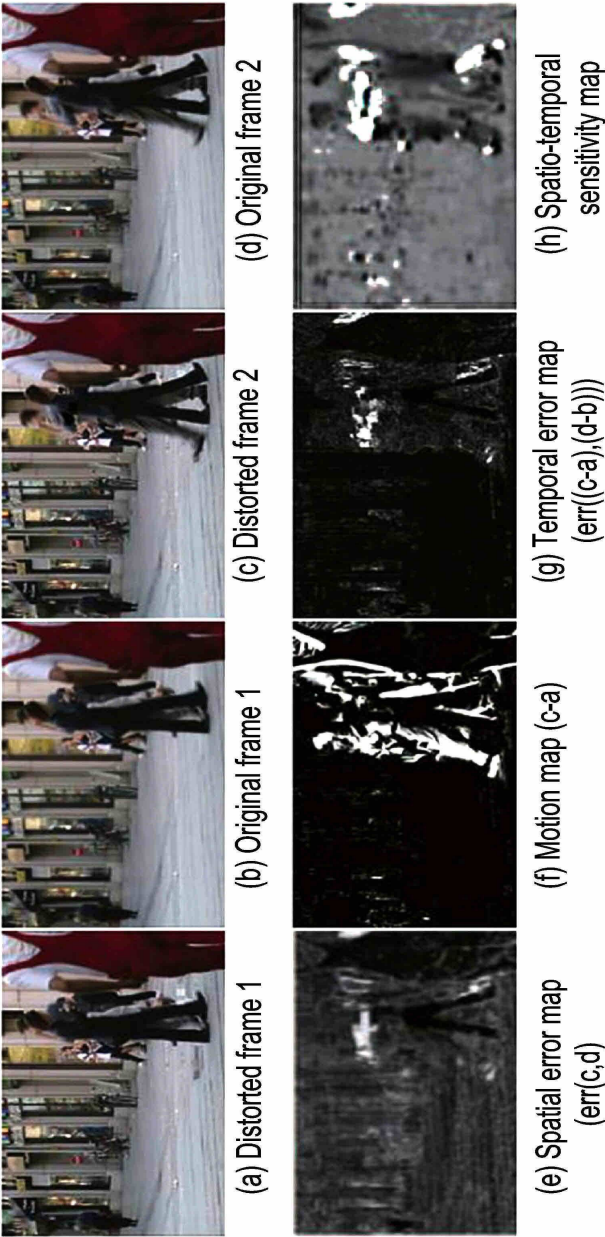
도면1



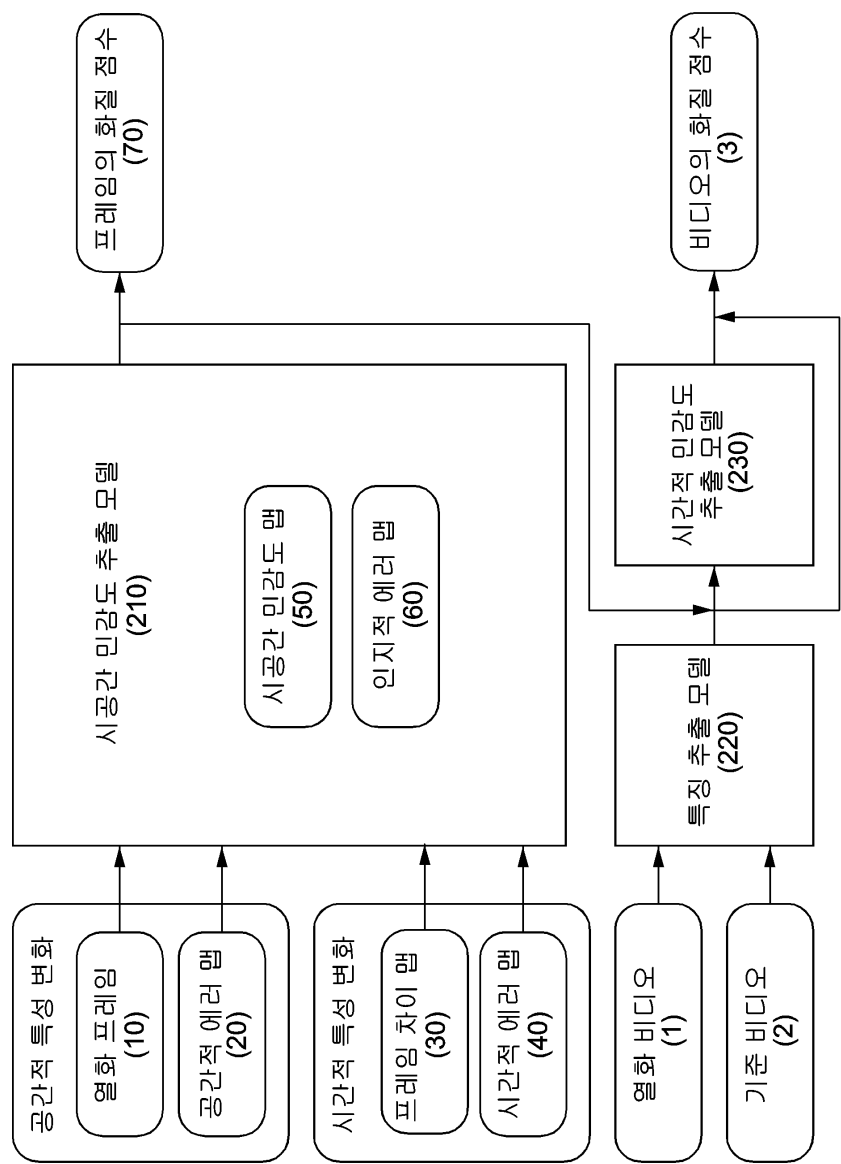
도면2



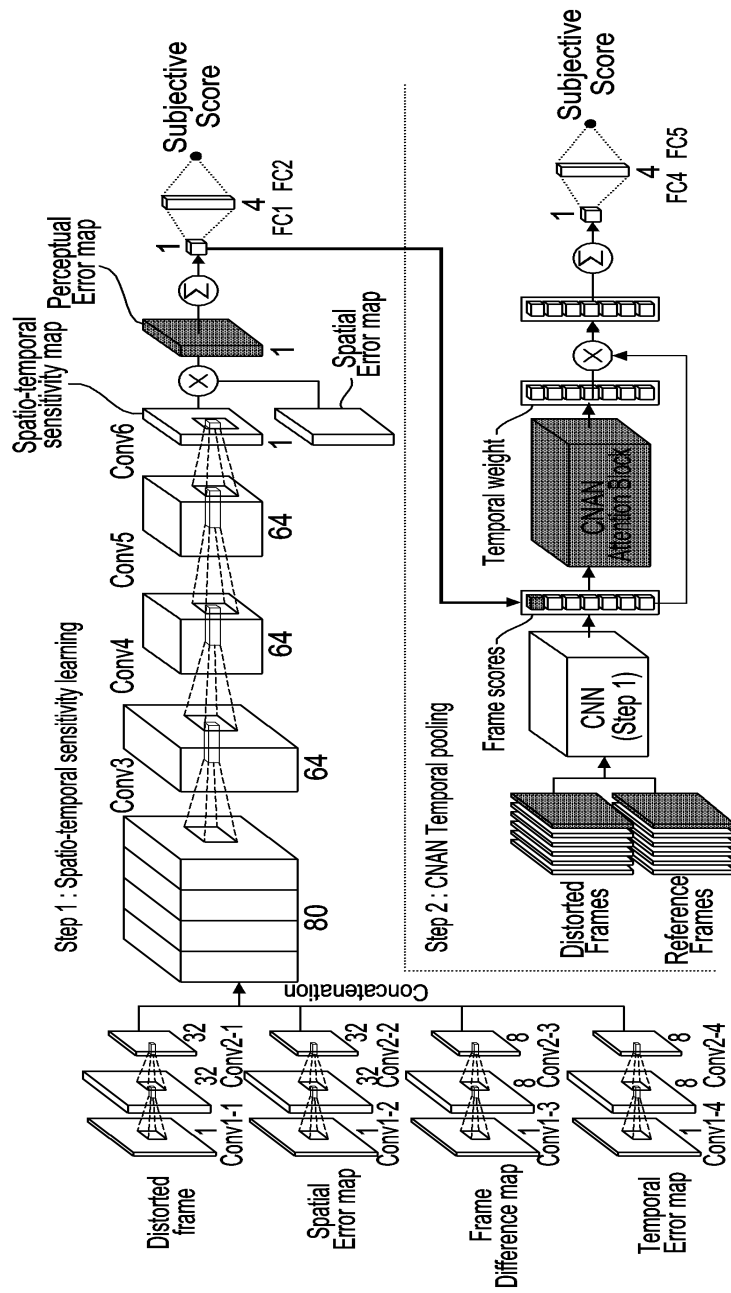
도면3



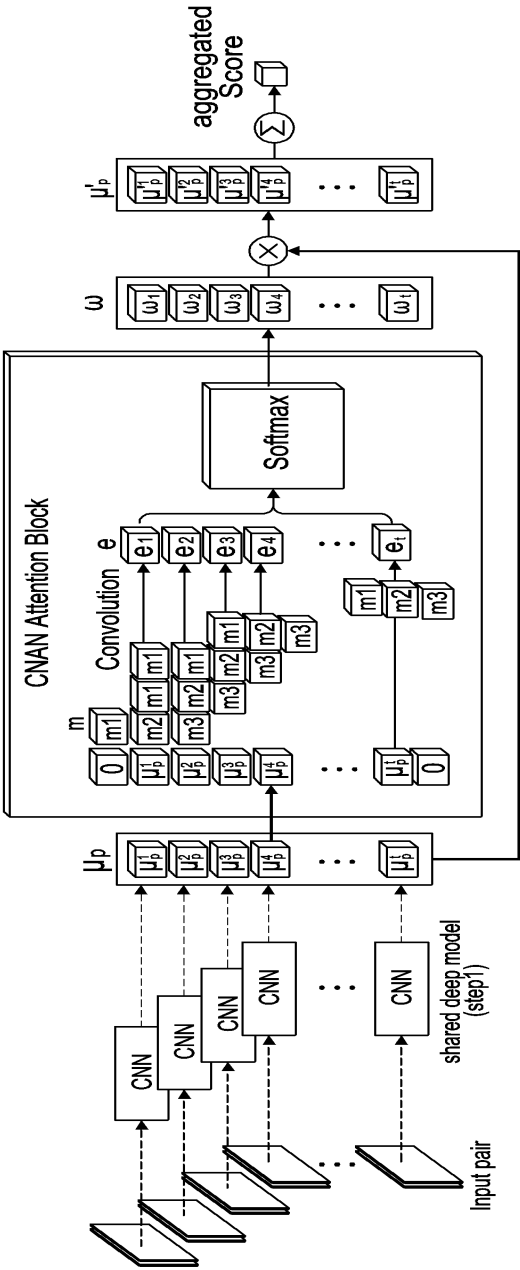
도면4



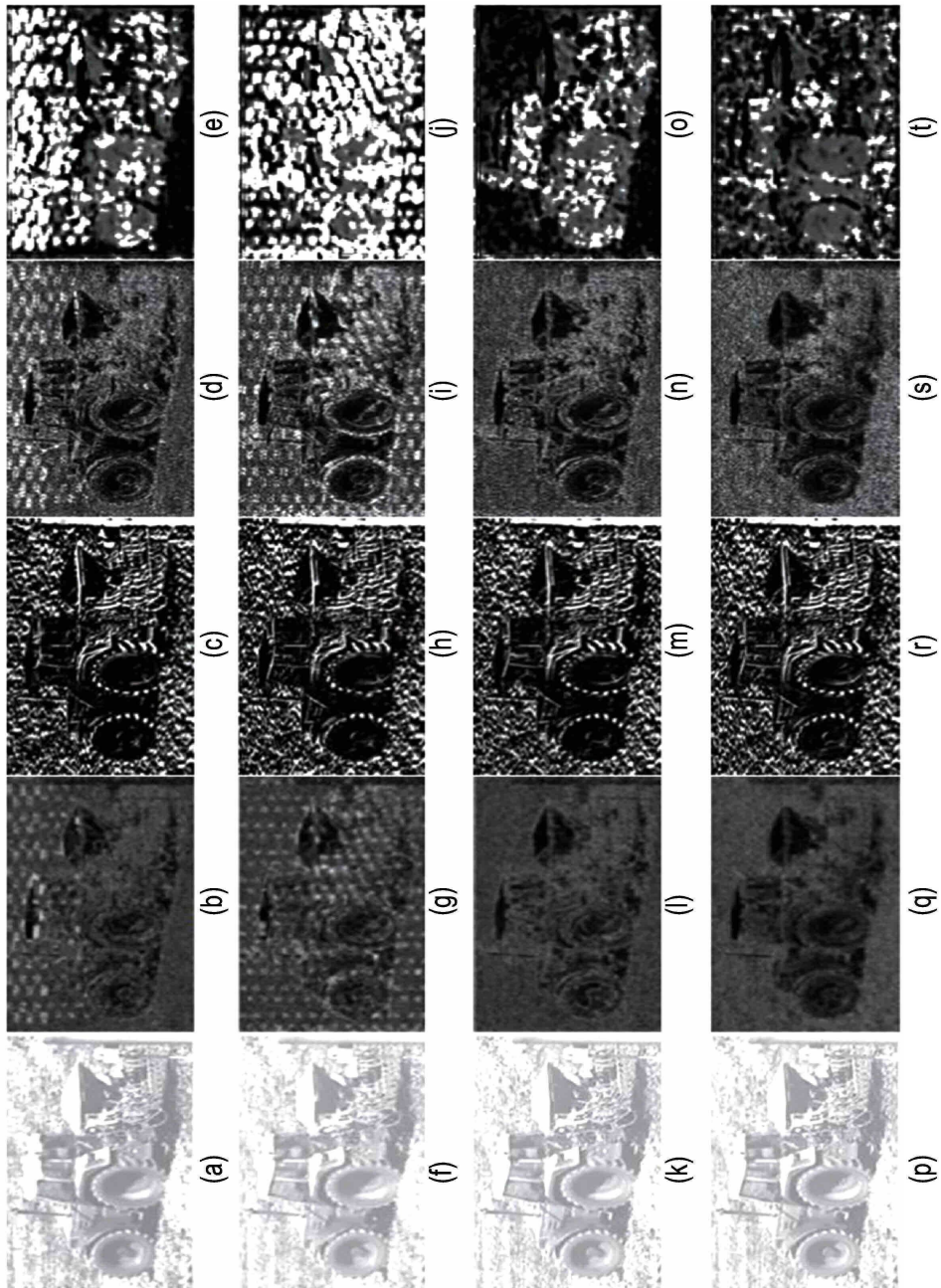
도면5



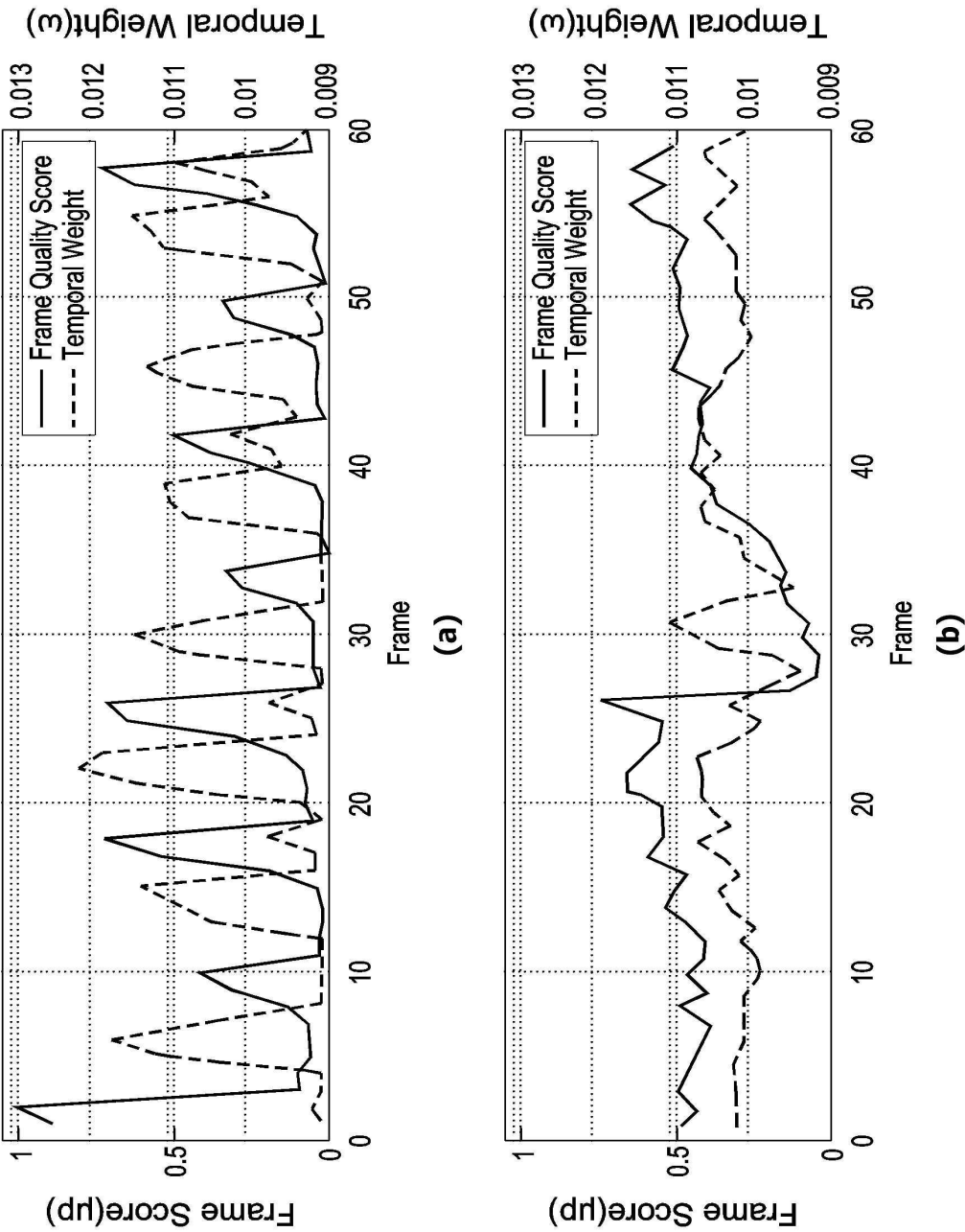
도면6



도면7



도면8



도면9

