



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0121341  
(43) 공개일자 2018년11월07일

(51) 국제특허분류(Int. Cl.)  
H04N 19/61 (2014.01) H04N 19/142 (2014.01)  
H04N 19/176 (2014.01) H04N 19/593 (2014.01)  
(52) CPC특허분류  
H04N 19/61 (2015.01)  
H04N 19/142 (2015.01)  
(21) 출원번호 10-2018-0026066  
(22) 출원일자 2018년03월06일  
심사청구일자 2018년03월06일  
(30) 우선권주장  
1020170055277 2017년04월28일 대한민국(KR)

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
최윤식  
서울특별시 마포구 마포대로24길 16, 115동 204호 (아현동, 공덕자이 아파트)  
홍은기  
서울특별시 서대문구 연희로10길 24-15, 205호 (연희동)  
황현수  
서울특별시 서대문구 연세로7길 13, 310호 (창천동)  
(74) 대리인  
특허법인우인

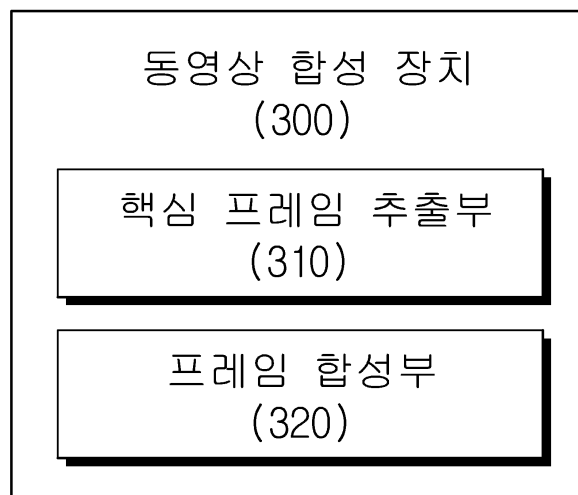
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 압축 영역에서 핵심 프레임이 포함된 동영상 합성 방법 및 장치

(57) 요약

본 실시예들은 매크로 블록 또는 예측 유닛을 그룹화한 지역 블록을 기반으로 부호화 예측 모드를 이용하여 장면 전환을 검출하고 압축 영역에서 프레임을 합성함으로써, 디코딩 및 재 인코딩에 따른 소프트웨어 및 하드웨어의 추가적인 부담없이 장면 전환을 검출하고 동작 속도를 증가시킬 수 있는 동영상 합성 방법 및 장치를 제공한다.

대표도 - 도3



(52) CPC특허분류

**H04N 19/176** (2015.01)

**H04N 19/593** (2015.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1425104091

부처명 산업통상자원부

연구관리전문기관 중소기업기술정보진흥원

연구사업명 산학협력기술개발사업

연구과제명 대용량 영상콘텐츠 판매를 위한 압축영상의 편집 및 마킹시스템

기 여 율 1/1

주관기관 연세대학교산학협력단

연구기간 2016.05.01 ~ 2017.04.30

---

## 명세서

### 청구범위

#### 청구항 1

컴퓨팅 디바이스에 의한 동영상 합성 방법에 있어서,

압축 영역에서, 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 단계; 및

상기 압축 영역에서, 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 단계

를 포함하는 동영상 합성 방법.

#### 청구항 2

제1항에 있어서,

상기 동영상은 기설정된 압축 포맷에 맞게 트랜스 코딩된 영상인 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 3

제1항에 있어서,

상기 복수의 프레임은 시간적 상관성을 고려하지 않고 프레임 자체 정보를 이용하여 부호화하는 인트라 프레임인 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 4

제1항에 있어서,

상기 지역 블록은 기설정된 압축 포맷에서 정의된 매크로 블록 또는 예측 유닛을 그룹화한 것으로, 상기 프레임에서 수평 방향으로 설정된 복수의 매크로 블록 또는 예측 유닛을 포함하고 상기 프레임에서 수직 방향으로 설정된 복수의 매크로 블록 또는 예측 유닛을 포함하며,

상기 지역 블록은 하나의 프레임에서 영역을 중복하지 않도록 구분하여 설정되는 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 5

제1항에 있어서,

상기 핵심 프레임을 추출하는 단계는,

(i) 하나의 프레임에서 위치 및 순서에 따라 설정된 지역 블록과 (ii) 상기 설정된 지역 블록에 대응하는 다른 프레임의 지역 블록을 비교하여, 인트라 모드로 부호화된 블록의 변화를 기반으로 상기 핵심 프레임을 추출하는 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 6

제5항에 있어서,

상기 핵심 프레임을 추출하는 단계는,

상기 지역 블록을 단위로 하여 (i) 상기 인트라 모드로 부호화된 블록의 크기, (ii) 상기 인트라 모드로 부호화된 블록의 분포, 및 (iii) 상기 인트라 모드로 부호화된 블록의 인트라 모드의 예측 방향에 따라 장면 전환을 판단하는 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 7

제1항에 있어서,

상기 프레임을 합성하는 단계는,

프레임 단위로 부호화된 비트스트림에 대하여 매크로 블록 어드레스 정보 또는 예측 유닛 어드레스 정보를 변경하여 재할당하는 방식으로 복수의 프레임을 합성하는 것을 특징으로 하는 동영상 합성 방법.

#### 청구항 8

압축 영역에서, 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 핵심 프레임 추출부; 및

상기 압축 영역에서, 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 프레임 합성부

를 포함하는 동영상 합성 장치.

#### 청구항 9

제8항에 있어서,

상기 동영상은 기설정된 압축 포맷에 맞게 트랜스 코딩된 영상인 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 10

제8항에 있어서,

상기 복수의 프레임은 시간적 상관성을 고려하지 않고 프레임 자체 정보를 이용하여 부호화하는 인트라 프레임인 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 11

제8항에 있어서,

상기 지역 블록은 기설정된 압축 포맷에서 정의된 매크로 블록 또는 예측 유닛을 그룹화한 것으로, 상기 프레임에서 수평 방향으로 설정된 복수의 매크로 블록 또는 예측 유닛을 포함하고 상기 프레임에서 수직 방향으로 설정된 복수의 매크로 블록 또는 예측 유닛을 포함하며,

상기 지역 블록은 하나의 프레임에서 영역을 중복하지 않도록 구분하여 설정되는 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 12

제8항에 있어서,

상기 핵심 프레임 추출부는,

(i) 하나의 프레임에서 위치 및 순서에 따라 설정된 지역 블록과 (ii) 상기 설정된 지역 블록에 대응하는 다른 프레임의 지역 블록을 비교하여, 인트라 모드로 부호화된 블록의 변화를 기반으로 상기 핵심 프레임을 추출하는 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 13

제12항에 있어서,

상기 핵심 프레임 추출부는,

상기 지역 블록을 단위로 하여 (i) 상기 인트라 모드로 부호화된 블록의 크기, (ii) 상기 인트라 모드로 부호화된 블록의 분포, 및 (iii) 상기 인트라 모드로 부호화된 블록의 인트라 모드의 예측 방향에 따라 장면 전환을 판단하는 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 14

제8항에 있어서,

상기 프레임 합성부는,

프레임 단위로 부호화된 비트스트림에 대하여 매크로 블록 어드레스 정보 또는 예측 유닛 어드레스 정보를 변경하여 재할당하는 방식으로 복수의 프레임을 합성하는 것을 특징으로 하는 동영상 합성 장치.

#### 청구항 15

프로세서에 의해 실행 가능한 컴퓨터 프로그램 명령어들을 포함하는 비일시적(Non-Transitory) 컴퓨터 판독 가능한 매체에 기록되어 동영상 합성을 위한 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램 명령어들이 컴퓨팅 디바이스의 적어도 하나의 프로세서에 의해 실행되는 경우에,

압축 영역에서, 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 단계; 및

상기 압축 영역에서, 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 단계

를 포함한 동작들을 수행하는 컴퓨터 프로그램.

#### 발명의 설명

##### 기술 분야

[0001] 본 발명이 속하는 기술 분야는 압축 영상에서 핵심 프레임이 포함된 동영상을 합성하는 방법 및 장치에 관한 것이다.

##### 배경 기술

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

[0003] 기존의 H.264 또는 HEVC 기반의 영상 합성 방식은 각각의 독립된 영상들을 하나로 합하여 합성 영상을 생성한다. 이러한 방식은 공간영역(픽셀영역)에서 주요 장면을 추출한다. 주요 장면을 추출하기 위해 동영상을 디코딩하고 다시 인코딩을 수행하므로 전처리(Pre-Processing)를 위하여 비디오 인코더의 복잡도가 증가하고 동작 속도가 저하되는 문제가 있다.

#### 발명의 내용

##### 해결하려는 과제

[0004] 본 발명의 실시예들은 압축 영역에서 이미 생성된 여러 영상들의 비트스트림으로부터 필요한 부분만 추출하고 추출된 부분을 합하여 합성 영상의 비트스트림을 생성함으로써, 압축 영역에서 장면 전환을 검출하고 압축 영역에서 매끄러운 영상을 합성하는 데 주된 목적이 있다.

[0005] 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

[0006] 본 실시예의 일 측면에 의하면, 컴퓨팅 디바이스에 의한 동영상 합성 방법에 있어서, 압축 영역에서 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 단계, 및 상기 압축 영역에서 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 단계를 포함하는 동영상 합성 방법을 제공한다.

[0007] 본 실시예의 다른 측면에 의하면, 압축 영역에서 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 핵심 프레임 추출부, 및 상기 압축 영역에서 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 프레임 합성부를 포함하는 동영상 합성 장치를 제공한다.

[0008] 본 실시예의 또 다른 측면에 의하면, 프로세서에 의해 실행 가능한 컴퓨터 프로그램 명령어들을 포함하는 비일시적(Non-Transitory) 컴퓨터 판독 가능한 매체에 기록되어 동영상 합성을 위한 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램 명령어들이 컴퓨팅 디바이스의 적어도 하나의 프로세서에 의해 실행되는 경우에, 압축 영역에서 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출하는 단계, 및 상기 압축 영역에서 상기 추출된 핵심 프레임을 이용하여 상기 복수의 프레임 중에서 일부의 프레임을 합성하는 단계를 포함한 동작들을 수행하는 컴퓨터 프로그램을 제공한다.

### 발명의 효과

[0009] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 매크로 블록 또는 예측 유닛을 그룹화한 지역 블록을 기반으로 부호화 예측 모드를 이용한 후처리(Post-Processing) 방식으로 장면 전환을 검출하고 압축 영역에서 프레임을 합성함으로써, 디코딩 및 재 인코딩에 따른 소프트웨어 및 하드웨어의 추가적인 부담없이 장면 전환을 검출하고 동작 속도를 향상시킬 수 있는 효과가 있다.

[0010] 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

[0011] 도 1은 영상 부호화 장치를 예시한 도면이다.

도 2는 영상 복호화 장치를 예시한 도면이다.

도 3은 본 발명의 일 실시예에 따른 동영상 합성 장치를 예시한 블록도이다.

도 4는 본 발명의 일 실시예에 따른 동영상 합성 장치가 프레임 간에 유사도를 측정하는 동작을 예시한 도면이다.

도 5a는 프레임에서 매크로 블록을 예시한 도면이고, 도 5b는 프레임에서 지역 블록을 예시한 도면이다.

도 6은 본 발명의 다른 실시예에 따른 동영상 합성 방법을 예시한 흐름도이다.

### 발명을 실시하기 위한 구체적인 내용

[0012] 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.

[0013] 도 1은 영상 부호화 장치를 예시한 도면이다.

[0014] 도 1을 참조하면, 상기 영상 부호화 장치(100)는 움직임 예측부(111), 움직임 보상부(112), 인트라 예측부(120), 스위치(115), 감산기(125), 변환부(130), 양자화부(140), 엔트로피 부호화부(150), 역양자화부(160), 역변환부(170), 가산기(175), 필터부(180) 및 참조영상 버퍼(190)를 포함한다.

[0015] 영상 부호화 장치(100)는 입력 영상에 대해 인트라(Intra) 모드 또는 인터(Inter) 모드로 부호화를 수행하고 비트스트림을 출력한다. 이하 본 발명의 실시예에서는 인트라 예측은 화면 내 예측, 인터 예측은 화면 간 예측과 동일한 의미로 사용될 수 있다. 예측 단위에 대한 최적의 예측 방법을 결정하기 위해 예측 단위에 대해 화면 내 예측 방법 및 화면 간 예측 방법이 선택적으로 사용될 수 있다. 영상 부호화 장치(100)는 입력 영상의 원본 블록에 대한 예측 블록을 생성한 후, 원본 블록과 예측 블록의 차분을 부호화한다.

[0016] 화면 내 예측 모드인 경우, 인트라 예측부(120)(또는 화면 내 예측부도 동일한 의미를 가지는 용어로 사용될 수 있음)는 현재 블록 주변의 이미 부호화된 블록의 픽셀값을 이용하여 공간적 예측을 수행하여 예측 블록을 생성한다.

[0017] 화면 간 예측 모드인 경우, 움직임 예측부(111)는, 움직임 예측 과정에서 참조 영상 버퍼(190)에 저장되어 있는 참조 영상에서 입력 블록과 가장 매치가 잘 되는 영역을 찾아 움직임 벡터를 구한다. 움직임 보상부(112)는 움직임 벡터를 이용하여 움직임 보상을 수행함으로써 예측 블록을 생성한다.

[0018] 감산기(125)는 입력 블록과 생성된 예측 블록의 차분에 의해 잔여 블록(Residual Block)을 생성한다. 변환부(130)는 잔여 블록에 대해 변환(Transform)을 수행하여 변환 계수(Transform Coefficient)를 출력한다. 그리고

양자화부(140)는 입력된 변환 계수를 양자화 파라미터에 따라 양자화하여 양자화된 계수(Quantized Coefficient)를 출력한다. 엔트로피 부호화부(150)는 입력된 양자화된 계수를 확률 분포에 따라 엔트로피 부호화하여 비트스트림(Bit Stream)을 출력한다.

- [0019] HEVC 역시 인트라 예측 부호화(화면 내 예측 부호화) 또는 인터 예측 부호화(화면 간 예측 부호화)를 수행한다. 인터 예측 부호화시 현재 부호화된 영상은 참조 영상으로 사용되기 위해 복호화되어 저장될 필요가 있다. 따라서 양자화된 계수는 역양자화부(160)에서 역양자화되고 역변환부(170)에서 역변환된다. 역양자화, 역변환된 계수는 가산기(175)를 통해 예측 블록과 더해지고 복원 블록이 생성된다.
- [0020] 복원 블록은 필터부(180)를 거치고, 필터부(180)는 디블록킹 필터(Deblocking Filter) 및 SAO(Sample Adaptive Offset) 중 적어도 하나 이상을 복원 블록 또는 복원 픽처에 적용할 수 있다. 필터부(180)는 적응적 인루프(in-loop) 필터로 불릴 수도 있다. 디블록킹 필터는 블록 간의 경계에 생긴 블록 왜곡을 제거할 수 있다. SAO는 코딩 에러를 보상하기 위해 픽셀값에 적정 오프셋(offset) 값을 더해줄 수 있다. 필터부(180)를 거친 복원 블록은 참조 영상 버퍼(190)에 저장된다.
- [0021] 도 2는 영상 복호화 장치를 예시한 도면이다.
- [0022] 도 2를 참조하면, 영상 복호화 장치(200)는 엔트로피 복호화부(210), 역양자화부(220), 역변환부(230), 인트라 예측부(240), 움직임 보상부(250), 필터부(260) 및 참조 영상 버퍼(270)를 포함한다.
- [0023] 영상 복호화 장치(200)는 부호화기에서 출력된 비트스트림을 입력 받아 인트라 모드 또는 인터 모드로 복호화를 수행하고 재구성된 영상, 즉 복원 영상을 출력한다. 인트라 모드인 경우 화면 내 예측 모드를 사용하여 예측 블록을 생성하고 인터 모드인 경우 화면 간 예측 방법을 사용하여 예측 블록을 생성한다. 영상 복호화 장치(200)는 입력 받은 비트스트림으로부터 잔여 블록(Residual Block)을 얻고 예측 블록을 생성한 후 잔여 블록과 예측 블록을 더하여 재구성된 블록, 즉 복원 블록을 생성한다.
- [0024] 엔트로피 복호화부(210)는 입력된 비트스트림을 확률 분포에 따라 엔트로피 복호화하여 양자화된 계수(Quantized Coefficient)를 출력한다. 양자화된 계수는 역양자화부(220)에서 역양자화되고 역변환부(230)에서 역변환되며, 양자화된 계수가 역양자화/역변환된 결과, 잔여 블록(Residual Block)이 생성된다.
- [0025] 화면 내 예측 모드인 경우, 인트라 예측부(240)(또는 화면 간 예측부)는 현재 블록 주변의 이미 부호화된 블록의 픽셀값을 이용하여 공간적 예측을 수행하여 예측 블록을 생성한다.
- [0026] 화면 간 예측 모드인 경우, 움직임 보상부(250)는 움직임 벡터 및 참조 영상 버퍼(270)에 저장되어 있는 참조 영상을 이용하여 움직임 보상을 수행함으로써 예측 블록을 생성한다.
- [0027] 잔여 블록과 예측 블록은 가산기(255)를 통해 더해지고, 더해진 블록은 필터부(260)를 거친다. 필터부(260)는 디블록킹 필터 및 SAO 중 적어도 하나 이상을 복원 블록 또는 복원 픽처에 적용할 수 있다. 필터부(260)는 재구성된 영상, 즉 복원 영상을 출력한다. 복원 영상은 참조 영상 버퍼(270)에 저장되어 화면 간 예측에 사용될 수 있다.
- [0028] 부호화/복호화 장치의 예측 성능을 향상시키기 위한 방법에는 보간(Interpolation) 영상의 정확도를 높이는 방법과 차신호를 예측하는 방법이 있다. 여기서 차신호란 원본 영상과 예측 영상과의 차이를 나타내는 신호이다.
- [0029] 본 발명에서 '차신호'는 문맥에 따라 '차분 신호', '잔여 블록' 또는 '차분 블록'으로 대체되어 사용될 수 있으며, 해당 기술분야에서 통상의 지식을 가진 자는 발명의 사상, 본질에 영향을 주지 않는 범위 내에서 이를 구분할 수 있을 것이다.
- [0030] 본 발명의 실시예에서는 설명의 편의상 코딩 유닛(Coding Unit, CU)을 부호화 단위라는 용어로 사용하지만, 부호화뿐만 아니라 복호화를 수행하는 단위가 될 수도 있다. 이하, 본 발명의 실시예에서 설명하는 영상 부호화 방법은, 각 모듈의 기능에서 맞게 구현될 수 있고 이러한 부호화기 및 복호화기는 본 발명의 권리범위에 포함된다. 즉, 본 발명의 실시예에서 후술할 영상 부호화/복호화 방법은 영상 부호화기 및 영상 복호화기에 포함된 각 구성부에서 수행될 수 있다. 구성부의 의미는 하드웨어적인 의미뿐만 아니라 알고리즘을 통해 수행될 수 있는 소프트웨어적인 처리 단위도 포함할 수 있다.
- [0031] 도 1, 도 2의 실시예에 따른 영상 부호화/복호화에서는, 영상의 효율적인 부호화를 위해, 표준에 따른 압축의 기본 단위를 기준으로 부호화를 진행한다. 예컨대, H.264의 부호화 단위는 매크로블록이고, HEVC에서는 단일 크기의 부호화 단위가 다양한 크기로 확장된 CU 구조가 정의될 수 있다.



- [0032] 본 명세서에서 블록 또는 예측 부호화 단위는 부호화/복호화의 단위를 의미한다. 부호화/복호화 과정에서, 영상은 소정의 크기로 분할되어 부호화/복호화된다. H.264에서 블록(예측 부호화 단위)은 매크로블록으로 불릴 수 있고, HEVC에서 블록(예측 부호화 단위)은 CU, PU, TU 등으로도 불릴 수도 있으며, 하나의 블록은 더 작은 크기의 하위 블록으로 분할될 수도 있다.
- [0033] 압축의 기본 단위인 기존의 H.264에서의 매크로 블록(Macro Block, MB)이 HEVC(High-Efficiency Video Coding)에서 코딩 트리 유닛(Coding Tree Unit, CTU)으로 대체되었다. 이때, MB의 크기는 16x16크기를 가지는 것에 반해, CTU의 크기는 16x16, 32x32, 64x64 중 하나를 선택할 수 있다. 그리고 하나의 CTU는 다시 하위 코딩 유닛(Coding Unit, CU)들로 나누어질 수 있고, 이러한 CU는 다시 예측 유닛(Prediction Unit, PU)와 변환 유닛(Transform Unit, TU)으로 나누어진다.
- [0034] 각 CU는 64x64 내지 8x8의 크기를 가질 수 있으며, PU는 64x64 내지 4x4의 크기를 가질 수 있고, TU는 32x32 내지 4x4의 크기를 가질 수 있다. 따라서 HEVC의 경우 H.264에 비해 다양한 크기의 블록 및 주파수 변환을 지원한다. H.264에서는 인트라 예측(Intra-Prediction, 화면 내 예측)에 9개의 예측 방향을 가지는 것에 반해, HEVC는 35가지의 예측 방향을 사용한다. 인터 예측(Inter-Prediction, 화면 간 예측)의 경우에도 기존 H.264의 인터 예측의 경우 최대 블록 사이즈가 16x16까지 제한되어 있는 반면, HEVC에서는 4x4에서 64x64까지 다양한 블록 사이즈를 취급할 수 있도록 확장되었다.
- [0035] CU는 비디오 부호화기에서 부호화가 수행되는 하나의 단위로서, 쿼드 트리 구조(Quad Tree Structure)를 기초로 깊이(Depth) 정보를 가지고 계층적으로 분할될 수 있다. CU는 8X8, 16X16, 32X32, 64X64 등 다양한 크기를 가질 수 있다. 또한 가장 큰 크기의 CU를 LCU(Largest Coding Unit), 가장 작은 크기의 CU를 SCU(Smallest Coding Unit)라 한다. SCU를 제외한 모든 CU는 split\_flag 정보를 할당하여 그 값에 따라 해당 CU가 분할된 영역인지 아닌지를 지시하게 된다. 부호화기는 다양한 비디오 신호 특성에 따라, 부호화 과정에서 LCU의 크기를 조절할 수도 있다. CU는 인트라 또는 인터 예측에 사용되도록 PU(Prediction Unit)으로 분할될 수 있으며, 변환 및 양자화를 위해 TU(Transform Unit)으로 분할될 수 있다.
- [0036] PU는 예측 및/또는 움직임 보상 수행의 기본 단위를 의미한다. PU는 복수의 파티션(Partition)으로 분할될 수 있으며, 각각의 파티션은 PU 파티션(Prediction Unit Partition)으로 불린다. PU가 복수의 파티션으로 분할된 경우, PU 파티션은 예측 및/또는 움직임 보상 수행의 기본단위가 될 수 있다. 이하, 본 발명의 실시예에서는 PU는 PU 파티션을 의미할 수도 있다.
- [0037] HEVC(High Efficiency Video Coding)에서는 향상된 움직임 벡터 예측(Advanced Motion Vector Prediction, AMVP)에 기반한 움직임 벡터 예측(Motion Vector Prediction) 방법을 사용한다.
- [0038] 향상된 움직임 벡터 예측에 기반한 움직임 벡터 예측 방법에서는 부호화/복호화 대상 블록의 주변에 위치하는 복원 블록의 움직임 벡터(Motion Vector, MV)뿐만 아니라, 참조 픽처(Reference Picture) 내에서 부호화/복호화 대상 블록과 동일한 위치 또는 대응되는 위치에 존재하는 블록의 움직임 벡터를 이용할 수 있다. 이때, 참조 픽처 내에서 부호화/복호화 대상 블록과 동일한 위치 또는 공간적으로 대응되는 위치에 존재하는 블록을 동등 위치 블록(Collocated Block), 동등 위치 블록의 움직임 벡터를 동등 위치 움직임 벡터(Collocated Motion Vector) 또는 시간적 움직임 벡터(Temporal Motion Vector)라고 부른다. 그러나, 동등 위치 블록(Collocated Block)은 참조 픽처의 부호화/복호화 대상 블록과 반드시 동일한 위치에 존재하는 블록만이 아니라, 부호화/복호화 대상 블록과 위치가 유사한, 즉 대응되는 위치에 존재하는 블록일 수도 있다.
- [0039] 움직임 정보 병합(Motion Information Merge) 방법에서는 움직임 정보를 주변에 위치하는 복원 블록뿐만 아니라 동등 위치 블록으로부터도 유추하여, 부호화/복호화 대상 블록의 움직임 정보로 이용한다. 이때, 움직임 정보는 인터 예측 시에 필요한 참조 픽처 인덱스(Reference Picture Index), 움직임 벡터, 단방향(Uni-Direction) 또는 양방향(Bi-Direction) 등을 나타내는 인터 예측 모드 정보, 참조 픽처 리스트(Reference Picture List), 인트라 예측 모드로 부호화되었는지 인터 예측 모드로 부호화되었는지에 관한 예측 모드(Prediction Mode) 정보 중 적어도 하나 이상을 포함하는 정보이다.
- [0040] 부호화/복호화 대상 블록에서 예측된 움직임 벡터(Predicted Motion Vector)는 부호화/복호화 대상 블록과 공간적으로 인접한 주변 블록의 움직임 벡터뿐만 아니라, 부호화/복호화 대상 블록과 시간적으로 인접한 블록인 동등 위치 블록의 움직임 벡터일 수도 있다.
- [0041] 도 3은 동영상 합성 장치를 예시한 블록도이다.



- [0042] 도 3에 도시한 바와 같이, 동영상 합성 장치(300)는 핵심 프레임 추출부(310) 및 프레임 합성부(320)를 포함한다. 동영상 합성 장치(300)는 도 3에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0043] 동영상 합성 장치(300)는 압축 포맷의 압축 영역에서 핵심 프레임 및 주위 프레임을 추출하고 추출한 프레임들을 압축 영역에서 합성함으로써, 저복잡도에서 핵심 프레임을 추출할 수 있고 매끄러운 합성 영상을 생성할 수 있다.
- [0044] 핵심 프레임 추출을 압축 영역이 아닌 공간 영역에서 추출하면, 핵심 프레임 추출하기 위해 동영상을 디코딩하고 재 인코딩해야 하므로 많은 시간이 소요되는 문제가 있다. 따라서, 본 실시예에서는 H.264의 매크로 블록 또는 HEVC 코덱의 예측 유닛을 그룹화한 지역 블록의 크기 및 분포, 인트라 모드 블록의 모드 예측 방향 등을 통해 장면의 중요도를 측정하고 중요도를 통해 중요도의 기준점을 충족하는 모든 장면 전환점을 추출한다. 기준점은 동영상의 종류에 따라 다르게 측정될 수 있다. 영상의 복잡도가 높을수록 기준점은 더 높게 설정한다. 영상의 복잡도는 지역 블록의 크기 및 분포를 보고 결정한다.
- [0045] 핵심 프레임 추출부(310)는 압축 영역에서 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출한다. 동영상은 기설정된 압축 포맷에 맞게 트랜스 코딩된 영상일 수 있다. 즉, 핵심 프레임 추출부(310)는 압축 포맷의 압축 영역에서 트랜스 코딩 영상으로부터 핵심 프레임을 추출할 수 있다. 복수의 프레임은 시간적 상관성을 고려하지 않고 프레임 자체 정보를 이용하여 부호화하는 인트라 프레임일 수 있다. 편집용 코덱은 인트라 프레임을 사용하고, 재생용 코덱에는 압축 효율이 좋은 인터 프레임을 사용할 수 있다. 편집하기 전에 모든 프레임을 인트라 프레임으로 변환해서 작업하면 편집속도를 향상시킬 수 있다. 합성이나 보정 등의 작업시 인트라 프레임으로 변환하여 작업할 수 있다. Long GOP(Group Of Picture)를 인트라 프레임 코덱으로 재 인코딩하여 인트라 프레임만 갖는 Intra GOP로 변환할 수 있다.
- [0046] 동영상 합성 장치(300)는 추가로 트랜스 코딩 영상 생성부(미도시)를 포함할 수 있다. 트랜스 코딩 영상 생성부는 고화질 영상을 기 설정된 압축 포맷에 맞게 트랜스 코딩하여 트랜스 코딩 영상을 생성한다. 예컨대, 입력된 고화질 영상의 H.264 또는 HEVC 포맷으로 맞춰 트랜스 코딩한다. 트랜스 코딩은 입력된 고화질 영상의 형식을 맞추는 동작으로 입력 영상을 디코딩하여 목표하는 인코더로 재 인코딩한다.
- [0047] 트랜스 코딩은 압축되어 있는 동영상 데이터를 음성과 영상 데이터로 분리하고 각각의 압축을 해제한 상태에서 크기를 줄이고 압축해 하나의 데이터로 만드는 것이 트랜스코딩이다. 이 작업에서 동영상 크기를 변화시키고 다른 포맷으로 압축하는 작업을 처리하는 소프트웨어가 인코더이다. 원본 영상을 읽어서 비디오와 오디오로 나눈 다음 각각 디코딩 작업과 후처리, 인코딩 작업을 거친 다음 다시 하나로 합친다. 트랜스코더에서 Splitter(Parser, Demuxer 등)는 비디오와 오디오가 합쳐져 있는 동영상에서 비디오와 오디오를 분리해서 내보내는 역할을 수행하는 모듈이다. Muxer는 Splitter와 반대 역할을 수행하는 모듈로, 비디오와 오디오를 임의의 동영상 형식으로 합치는 역할을 수행한다.
- [0048] 프레임 합성부(320)는 압축 영역에서 추출된 핵심 프레임을 이용하여 복수의 프레임 중에서 일부의 프레임을 합성한다.
- [0049] 도 4는 동영상 합성 장치가 프레임 간에 유사도를 측정하는 동작을 예시한 도면이다.
- [0050] 도 4를 참조하면, 핵심 프레임 추출부(310)는 지역 블록을 비교 단위로 프레임 간의 유사도를 측정하며 지역 블록의 크기 및 분포, 인트라 모드 블록의 모드 예측 방향을 이용해 장면의 중요도를 측정하고 중요도의 기준점을 충족하는 장면 전환점을 추출한다. 각각의 인트라 프레임별로 인트라 예측 모드의 분포를 산출하고, 연속된 두 개의 인트라 프레임별 인트라 예측 모드 분포의 차이를 산출한다.
- [0051] 매크로 블록 또는 예측 유닛을 유사도 측정시의 비교 단위로 할 경우, 빠르게 움직이는 객체나 카메라의 급격한 움직임이 있는 동영상에서 예측 모드의 변화 특성을 잘못 추출할 수 있다. 즉, 프레임의 전체 부분이 아닌 국지적 부분에서 변화가 검출된 경우에도 장면 전환으로 인식되는 문제가 있다. 본 실시예는 매크로 블록 또는 예측 유닛 대신 각 프레임을 서로 중복되지 않는 복수의 지역 블록으로 나누고, 이를 비교 단위로 하여 프레임 별로 지역 블록 간의 변화를 측정하여 유사도를 산출한다.
- [0052] 도 5a는 프레임에서 매크로 블록을 예시한 도면이고, 도 5b는 프레임에서 지역 블록을 예시한 도면이다.
- [0053] 지역 블록(520)은 기설정된 압축 포맷에서 정의된 매크로 블록(510)을 그룹화한 것이다. 지역 블록(520)은 프레임을 임의로 구역별로 나눈 블록으로 매크로 블록(510)들의 집합이라고 볼 수 있다. 지역 블록(520)은 프레임에

서 수평 방향으로 설정된 복수의 매크로 블록을 포함하고 프레임에서 수직 방향으로 설정된 복수의 매크로 블록을 포함한다. 지역 블록(520)은 하나의 프레임에서 영역을 중복하지 않도록 구분하여 설정된다.

[0054] 다시 도 4를 참조하면, 상이한 크기를 갖는 매크로 블록이 아닌 균일한 크기를 갖는 복수개의 지역블록으로 분할하여 프레임 간의 유사도를 측정한다.

[0055] 핵심 프레임 추출부(310)는 (i) 하나의 프레임에서 위치 및 순서에 따라 설정된 지역 블록과 (ii) 설정된 지역 블록에 대응하는 다른 프레임의 지역 블록을 비교하여, 인트라 모드로 부호화된 블록의 변화를 기반으로 핵심 프레임을 추출한다.

[0056] 핵심 프레임 추출부(310)는 (i) 인트라 모드로 부호화된 블록의 크기, (ii) 인트라 모드로 부호화된 블록의 분포, 및 (iii) 인트라 모드로 부호화된 블록의 인트라 모드의 예측 방향에 따라 장면 전환을 판단한다.

[0057] 매크로 블록 대신 지역 블록을 비교 단위로 프레임 간의 유사도를 측정하는 방식에서, 인트라 예측 모드의 변화

( $\Delta f_i^L$ )는 수학식 1과 같이 표현된다.

### 수학식 1

$$\Delta f_i^L = \frac{1}{|S^L|} \left| \sum_{j \in S^L} Mode_i^j - \sum_{j \in S^L} Mode_{i+1}^j \right|$$

[0059] L은 한 프레임 내의 지역 블록을,  $S^L$ 는 지역 블록에 포함되는 매크로 블록의 집합을,  $|S^L|$ 은 지역 블록 내의 매크로 블록의 개수를 의미한다.

[0060] 수학식 1을 바탕으로 장면전환을 결정하는 함수  $d_{i,i+1}$ 는  $\Delta f_i^L$ 와 임계값  $\lambda$ 과 비교하여 결정되는  $q_i^L$ 에 의해 수학식 2와 같이 표현된다.

### 수학식 2

$$d_{i,i+1} = \frac{1}{N_{Local}} \sum_{\forall L} q_i^L$$

$$q_i^L = \begin{cases} 1, & \Delta f_i^L \geq \lambda \\ 0, & \Delta f_i^L < \lambda \end{cases}$$

[0062]  $N_{Local}$ 은 한 프레임 내의 지역 블록들의 개수이다. 이렇게 계산된 장면전환 결정 함수  $d_{i,i+1}$ 가 미리 결정된 임계값보다 클 경우, 연속된 프레임 i와 i+1에서 장면전환이 발생했다고 결정한다.

[0063] 프레임 합성부(320)는 압축 포맷의 압축 영역에서 추출한 핵심 프레임을 기준으로 복수의 프레임을 합성하여 합성 영상을 생성한다.

[0064] 프레임 합성부(320)는 하나 이상의 추출된 핵심 프레임들을 통해서 핵심 프레임이 포함된 영상을 만들기 위해 핵심 프레임에 연속된 후속 프레임들을 사용자의 선택, 본래 영상의 길이, 또는 추출된 핵심 프레임의 수에 따라 배치하여 자연스러운 영상 합성이 이루어지도록 한다. 영상 합성시, 시퀀스에 관리코드(국가, 계절, 지역, 영상번호 등)를 생성하여 함께 저장하고, 동시에 저작권 보호를 위해 자동으로 워터마크를 삽입할 수 있다.

[0065] 프레임 합성부(320)는 압축 영역에서 프레임 단위로 부호화된 비트스트림에 대하여 매크로 블록 어드레스 정보

또는 예측 유닛 어드레스 정보를 변경하여 재할당하는 방식으로 복수의 프레임을 합성할 수 있다.

- [0066] 프레임 합성부(320)는 합성 영상을 압축 포맷의 인코더를 통해 저장한다. 예컨대, 합성된 영상을 H.264 또는 HEVC 인코더를 통해 저장할 수 있다.
- [0067] 동영상 합성 장치에 포함된 구성요소들이 도 3에서는 분리되어 도시되어 있으나, 복수의 구성요소들은 상호 결합되어 적어도 하나의 모듈로 구현될 수 있다. 구성요소들은 장치 내부의 소프트웨어적인 모듈 또는 하드웨어적인 모듈을 연결하는 통신 경로에 연결되어 상호 간에 유기적으로 동작한다. 이러한 구성요소들은 하나 이상의 통신 버스 또는 신호선을 이용하여 통신한다.
- [0068] 동영상 합성 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 프로그램 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.
- [0069] 동영상 합성 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.
- [0070] 도 6은 본 발명의 다른 실시예에 따른 동영상 합성 방법을 예시한 흐름도이다. 동영상 합성 방법은 컴퓨팅 디바이스에 의하여 수행될 수 있으며, 동영상 합성 장치가 수행하는 동작에 관한 상세한 설명과 중복되는 설명은 생략하기로 한다.
- [0071] 단계 S610에서, 컴퓨팅 디바이스는 압축 영역에서 동영상에 포함된 복수의 프레임 간에 지역 블록을 기준으로 유사도를 판단하고 장면의 중요도를 측정하여 핵심 프레임을 추출한다. 동영상은 기설정된 압축 포맷에 맞게 트랜스 코딩된 영상일 있다. 복수의 프레임은 시간적 상관성을 고려하지 않고 프레임 자체 정보를 이용하여 부호화하는 인트라 프레임일 수 있다. 지역 블록은 기설정된 압축 포맷에서 정의된 매크로 블록을 그룹화한 것으로, 프레임에서 수평 방향으로 설정된 복수의 매크로 블록을 포함하고 프레임에서 수직 방향으로 설정된 복수의 매크로 블록을 포함한다. 지역 블록은 하나의 프레임에서 영역을 중복하지 않도록 구분하여 설정된다.
- [0072] 핵심 프레임을 추출하는 단계(S610)는 (i) 하나의 프레임에서 위치 및 순서에 따라 설정된 지역 블록과 (ii) 설정된 지역 블록에 대응하는 다른 프레임의 지역 블록을 비교하여, 인트라 모드로 부호화된 블록의 변화를 기반으로 핵심 프레임을 추출한다.
- [0073] 핵심 프레임을 추출하는 단계(S610)는 (i) 인트라 모드로 부호화된 블록의 크기, (ii) 인트라 모드로 부호화된 블록의 분포, 및 (iii) 인트라 모드로 부호화된 블록의 인트라 모드의 예측 방향에 따라 장면 전환을 판단할 수 있다.
- [0074] 단계 S620에서, 컴퓨팅 디바이스는 압축 영역에서 추출된 핵심 프레임을 이용하여 복수의 프레임 중에서 일부의 프레임을 합성한다.
- [0075] 프레임을 합성하는 단계(S620)는 압축 영역에서 프레임 단위로 부호화된 비트스트림에 대하여 매크로 블록 어드레스 정보를 변경하여 재할당하는 방식으로 복수의 프레임을 합성할 수 있다.
- [0076] 도 6에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 6에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0077] 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.

[0078]

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.

### 부호의 설명

[0079]

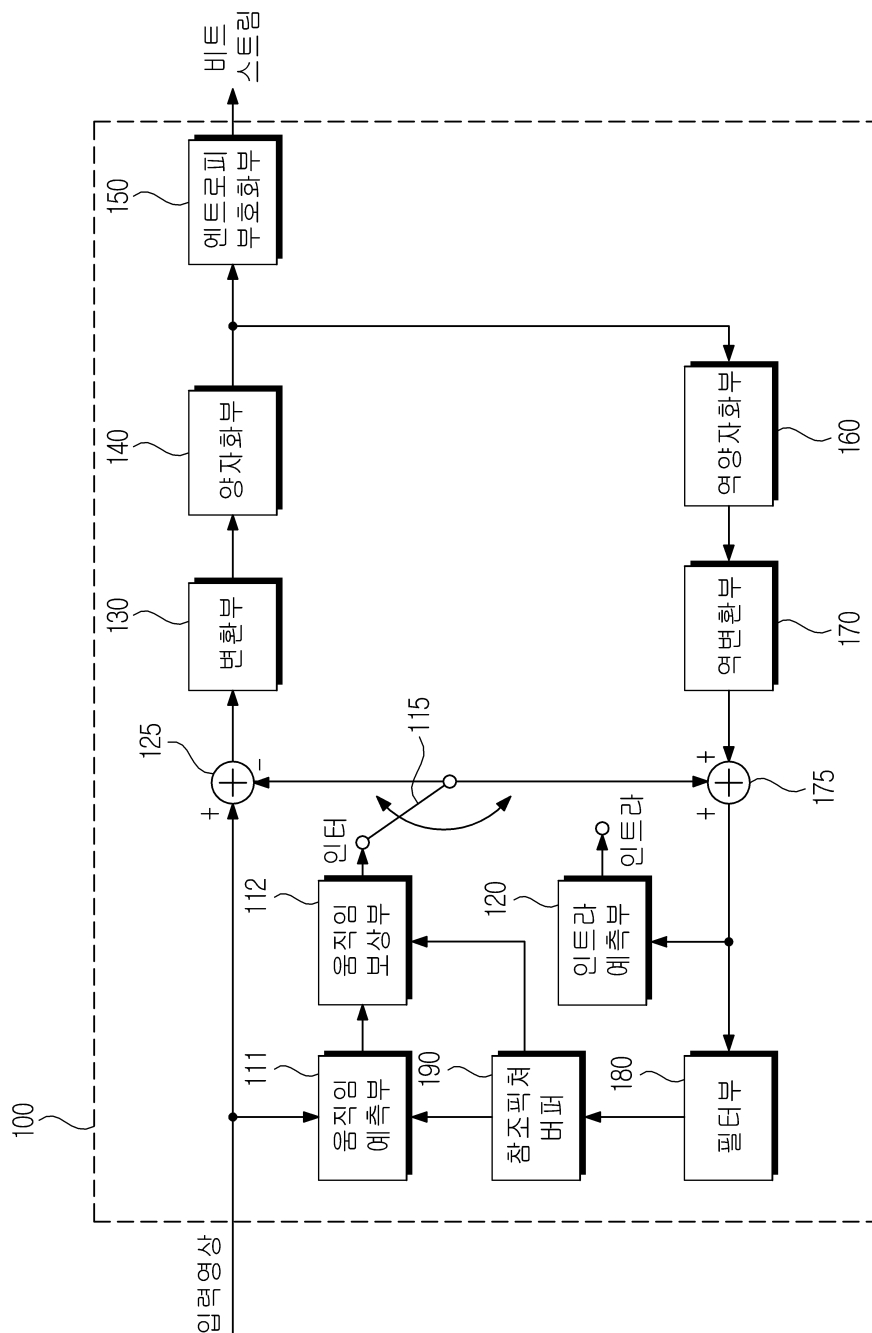
300: 동영상 합성 장치

310: 핵심 프레임 추출부

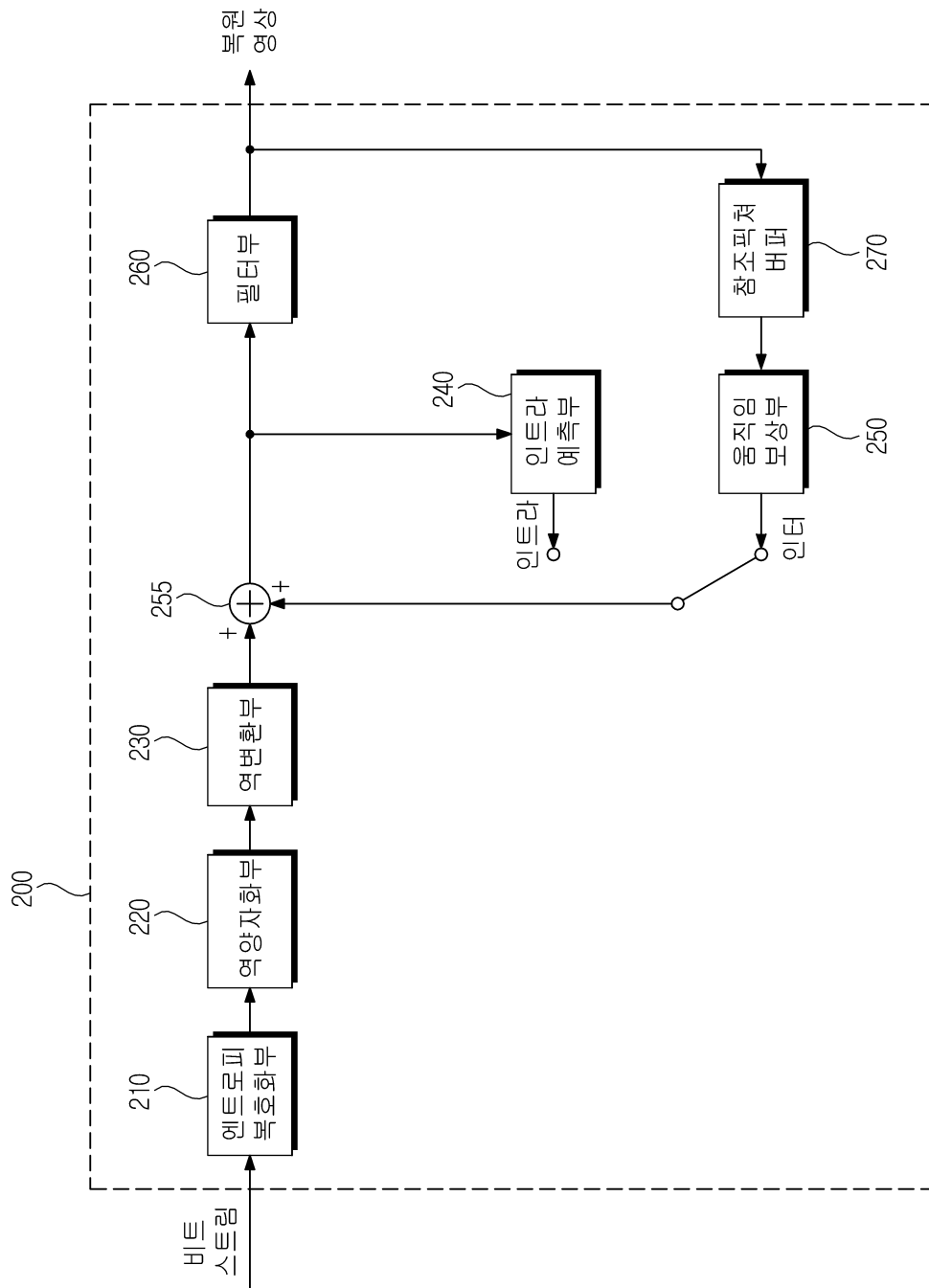
320: 프레임 합성부

### 도면

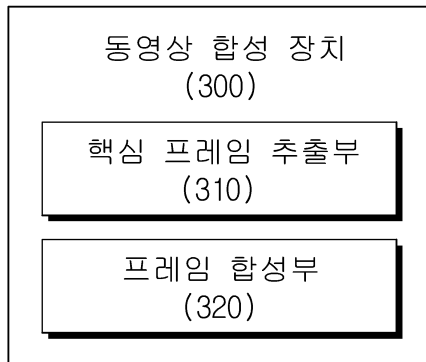
#### 도면1



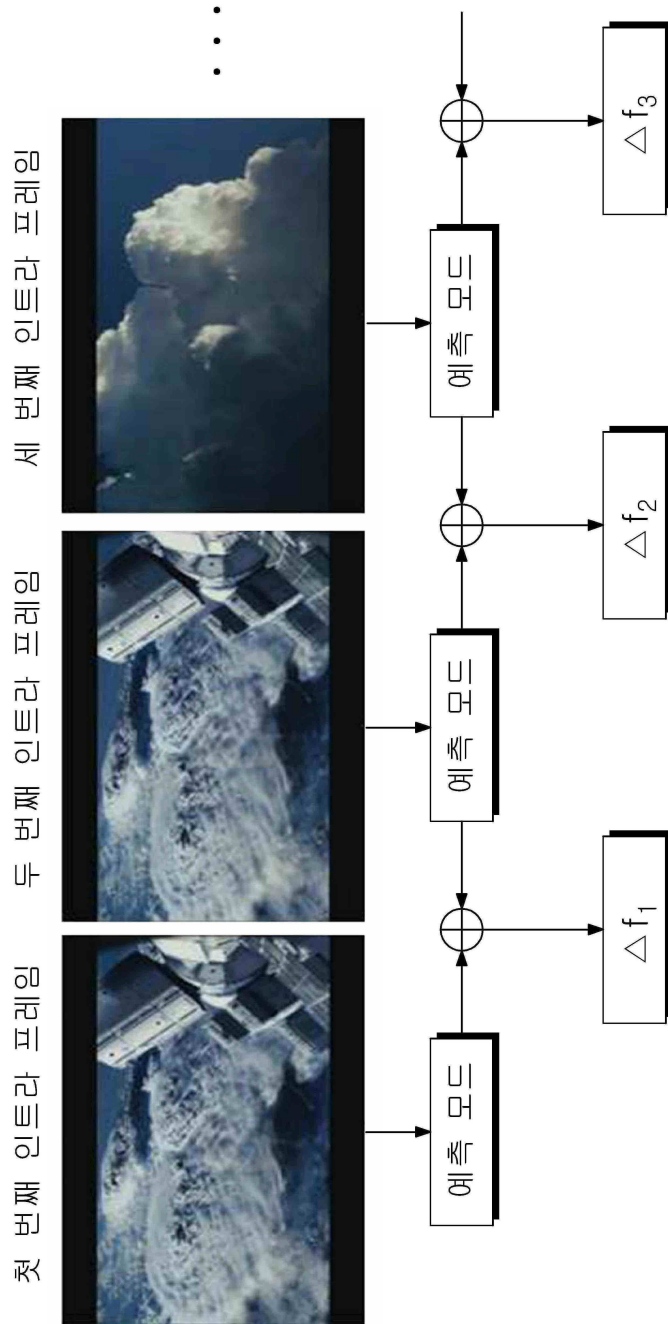
도면2



도면3



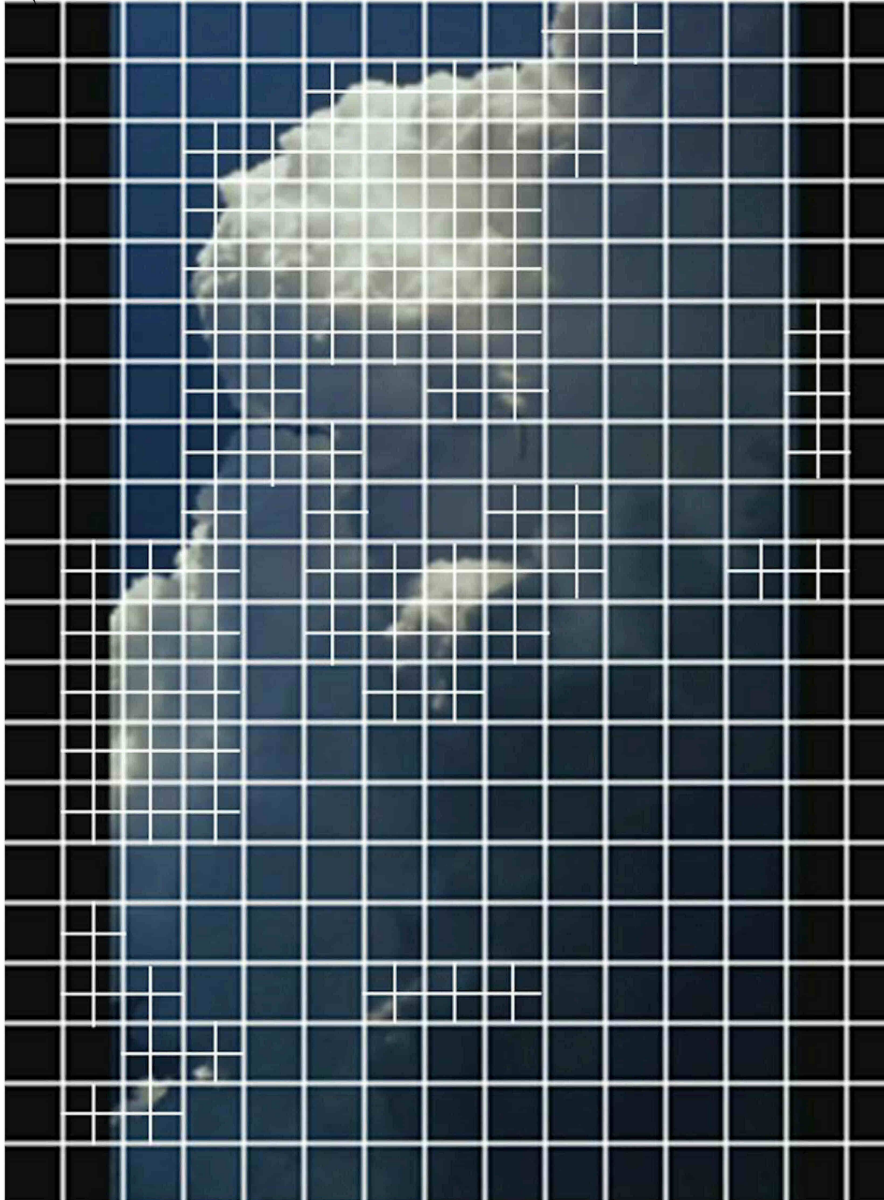
도면4





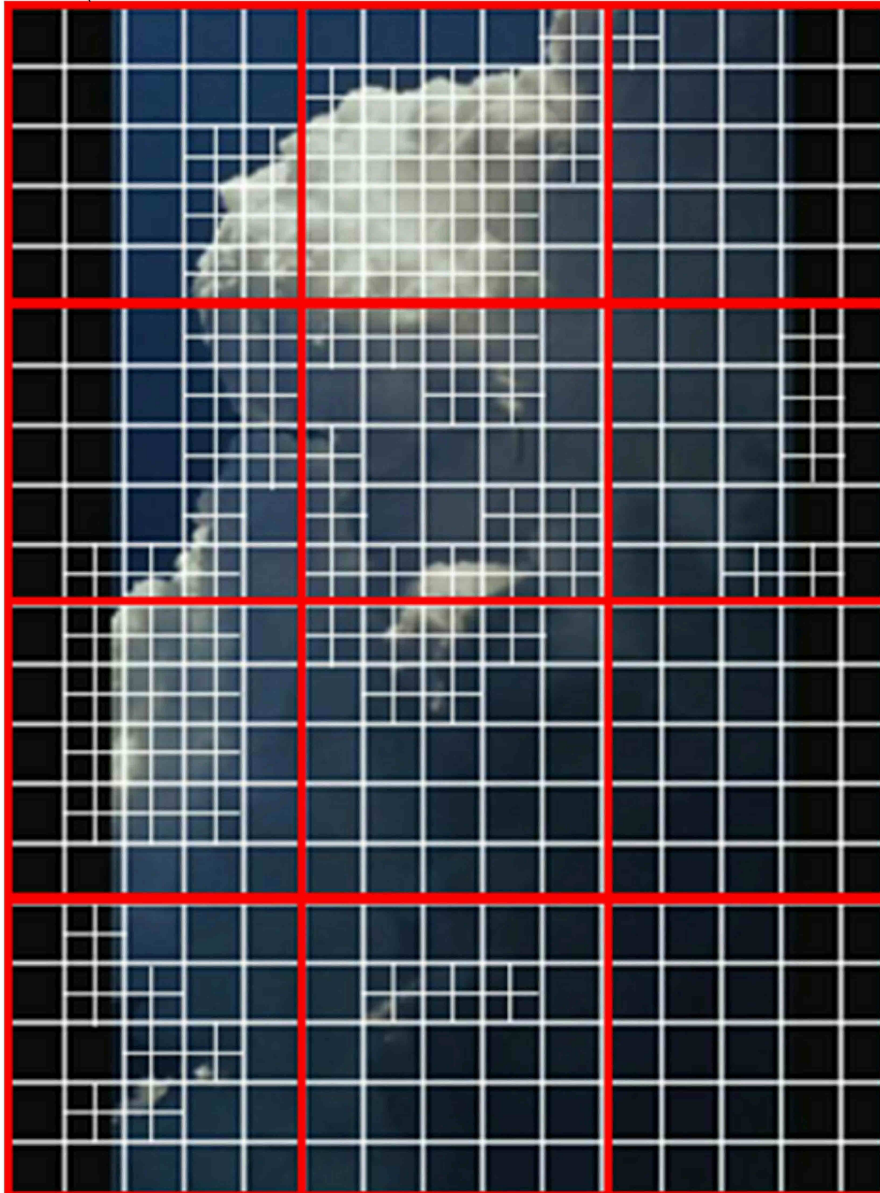
도면5a

510



도면5b

520



도면6

