



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년08월02일
(11) 등록번호 10-1873609
(24) 등록일자 2018년06월26일

- (51) 국제특허분류(Int. Cl.)
HO4N 19/00 (2014.01)
- (21) 출원번호 10-2013-7012366
- (22) 출원일자(국제) 2011년10월10일
심사청구일자 2016년10월06일
- (85) 번역문제출일자 2013년05월13일
- (65) 공개번호 10-2013-0139285
- (43) 공개일자 2013년12월20일
- (86) 국제출원번호 PCT/US2011/055562
- (87) 국제공개번호 WO 2012/051094
국제공개일자 2012년04월19일
- (30) 우선권주장
61/393,186 2010년10월14일 미국(US)
- (56) 선행기술조사문현
US07734652 B2
US20090326880 A1
KR1020120115303 A

- (73) 특허권자
톰슨 라이센싱
프랑스 이씨래뮬리노 뤼 잔 다르크 1-5 (우:
92130)
- (72) 발명자
모가담, 압둘레자
미국 미시건주 48864 오케모스 아파트먼트 에이9
매디슨 애비뉴 5218
솔, 조엘
미국 캘리포니아주 92037 라 졸라 유닛 106 빌라
라 졸라 드라이브 8722
루, 시아오안
미국 뉴저지주 08540 프린스턴 케네디 코트 30
- (74) 대리인
특허법인아주

전체 청구항 수 : 총 33 항

심사관 : 김영태

(54) 발명의 명칭 모션 매트릭스를 사용하여 비디오 인코딩 및 디코딩하기 위한 방법 및 장치

(57) 요약

모션 매트릭스를 사용하여 비디오 인코딩 및 디코딩하기 위한 방법 및 장치가 제공된다. 장치는 모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하기 위한 비디오 인코더(200)를 포함한다. 모션 매트릭스는 주어진 기준점 아래의 등급 및 딕셔너리에 관한 희박 표현을 가진다. 딕셔너리는 아톰의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 화상이 파생되도록 허용하기 위한 기본 벡터를 포함한다. 딕셔너리는 상기 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

명세서

청구범위

청구항 1

모션 매트릭스(motion matrix)를 사용하여 비디오 시퀀스에서 화상을 인코딩하기 위한 비디오 인코더를 포함하여, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스는 주어진 값 아래의 등급 및 딕셔너리(dictionary)에 관한 희박 표현(sparse representation)을 가지며, 상기 딕셔너리는 인코딩될 화상의 함수가 아닌 적어도 하나의 기준 프레임의 함수로서 생성된 복수의 딕셔너리를 중에서 선택되고, 상기 선택은 목표 등급 및 목표 희박도를 달성하는 것에 기초하며,

상기 모션 매트릭스는, 상기 선택된 딕셔너리에 관하여 상기 화상의 표현(representation)을 반복적으로 찾는 것, 상기 화상의 표현의 가장 큰 엔트리의 상기 목표 희박도의 수(number)만을 유지하는 것, 상기 모션 매트릭스를 개선하는 것, 상기 화상의 표현이 특정 레벨로 수렴할 때까지 상기 모션 매트릭스의 고유값의 최대 목표 등급 수를 상기 목표 등급의 수까지 유지하는 것에 의해 찾아지는 장치.

청구항 2

제1항에 있어서, 상기 딕셔너리는 상기 화상과 기준 화상의 세트 사이에서의 공간적 및 시간적 상관관계로부터 형성되는 장치.

청구항 3

제1항에 있어서, 상기 비디오 인코더는 임베딩된 매트릭스(embedded matrix) 내로 상기 모션 매트릭스를 임베딩하기 위한 임베딩 디바이스, 상기 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 상기 임베딩된 매트릭스를 샘플링하기 위한 샘플러, 샘플의 양자화된 서브세트를 제공하도록 상기 샘플의 서브세트를 양자화하기 위한 양자화기, 및 압축된 심볼 내로 상기 샘플의 양자화된 서브세트를 엔트로피 코딩하기 위한 엔트로피 코더를 포함하는 장치.

청구항 4

제3항에 있어서, 상기 임베딩된 매트릭스는 상기 모션 매트릭스가 곱해지는 2개의 매트릭스를 사용하여 결정되고, 상기 2개의 매트릭스는 상기 임베딩된 매트릭스에서 특정량(particular amount)의 랜덤 고유 벡터를 제공하도록 결정되는 장치.

청구항 5

제3항에 있어서, 상기 샘플의 서브세트는 상기 임베딩된 매트릭스로부터 랜덤하게 선택되는 장치.

청구항 6

제5항에 있어서, 동일 랜덤 시드와 동일 랜덤 수 발생기는, 상기 임베딩된 매트릭스로부터 상기 샘플의 서브세트에 있는 샘플이 상기 비디오 인코더와 대응 비디오 디코더에서 동일한 각각의 위치를 가지는 것을 보장하도록 상기 비디오 인코더와 상기 대응 비디오 디코더에 의해 사용되는 장치.

청구항 7

제3항에 있어서, 상기 샘플러는 상기 샘플의 서브세트에 있는 샘플과 다른, 상기 임베딩된 매트릭스에서의 위치를 가진 상기 임베딩된 매트릭스의 부분을 버리는 장치.

청구항 8

비디오 인코더에서의 방법으로서,

모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하는 단계를 포함하며, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스는 주어진 값 아래의 등급 및 딕셔너리에 관한 희박 표현을 가지며, 상기 딕셔

너리는 인코딩될 화상의 함수가 아닌 적어도 하나의 기준 프레임의 함수로서 생성된 복수의 딕셔너리들 중에서 선택되고, 상기 선택은 목표 등급 및 목표 희박도를 달성하는 것에 기초하며,

상기 모션 매트릭스는, 상기 선택된 딕셔너리에 관하여 상기 화상의 표현을 반복적으로 찾는 것, 상기 화상의 표현의 가장 큰 엔트리의 상기 목표 희박도의 수만을 유지하는 것, 상기 모션 매트릭스를 갱신하는 것, 상기 화상의 표현이 특정 레벨로 수렴할 때까지 상기 모션 매트릭스의 고유값의 최대 목표 등급 수를 상기 목표 등급의 수까지 유지하는 것에 의해 찾아지는 방법.

청구항 9

제8항에 있어서, 상기 딕셔너리는 상기 화상과 기준 화상의 세트 사이에서의 공간적 및 시간적 상관관계로부터 형성되는 방법.

청구항 10

제8항에 있어서, 상기 인코딩 단계는, 임베딩된 매트릭스 내로 상기 모션 매트릭스를 임베딩하는 단계;

상기 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 상기 임베딩된 매트릭스를 샘플링하는 단계;

샘플의 양자화된 서브세트를 제공하도록 상기 샘플의 서브세트를 양자화하는 단계; 및

압축된 심볼 내로 상기 샘플의 양자화된 서브세트를 엔트로피 코딩하는 단계를 포함하는 방법.

청구항 11

제10항에 있어서, 상기 임베딩된 매트릭스는 상기 모션 매트릭스가 곱해지는 2개의 매트릭스를 사용하여 결정되고, 상기 2개의 매트릭스는 상기 임베딩된 매트릭스에서 특정량의 랜덤 고유 벡터를 제공하도록 결정되는 것인 방법.

청구항 12

제10항에 있어서, 상기 샘플의 서브세트는 상기 임베딩된 매트릭스로부터 랜덤하게 선택되는 것인 방법.

청구항 13

제12항에 있어서, 동일 랜덤 시드와 동일 랜덤 수 발생기는, 상기 임베딩된 매트릭스로부터 상기 샘플의 서브세트에 있는 샘플이 상기 비디오 인코더와 대응 비디오 디코더에서 동일한 각각의 위치를 가지는 것을 보장하도록 상기 비디오 인코더와 상기 대응 비디오 디코더에 의해 사용되는 방법.

청구항 14

제10항에 있어서, 상기 샘플링 단계는 상기 샘플의 서브세트에 있는 샘플과 다른, 상기 임베딩된 매트릭스에서의 위치를 가지는 상기 임베딩된 매트릭스의 부분을 버리는 단계를 포함하는 방법.

청구항 15

모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하기 위한 비디오 디코더를 포함하되, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스의 근사는 주어진 값 아래의 등급 및 딕셔너리에 관한 희박 표현을 가지며, 상기 딕셔너리는 인코딩될 화상의 함수가 아닌 적어도 하나의 기준 프레임의 함수로서 생성된 복수의 딕셔너리들 중에서 선택되고, 상기 선택은 목표 등급 및 목표 희박도를 달성하는 것에 기초하며,

상기 모션 매트릭스는, 상기 선택된 딕셔너리에 관하여 상기 화상의 표현을 반복적으로 찾는 것, 상기 화상의 표현의 가장 큰 엔트리의 상기 목표 희박도의 수만을 유지하는 것, 상기 모션 매트릭스를 갱신하는 것, 상기 화상의 표현이 특정 레벨로 수렴할 때까지 상기 모션 매트릭스의 고유값의 최대 목표 등급 수를 상기 목표 등급의 수까지 유지하는 것에 의해 찾아지는 장치.

청구항 16

제15항에 있어서, 상기 딕셔너리는 상기 화상과 기준 화상의 세트 사이에서의 공간적 및 시간적 상관관계로부터

형성되는 장치.

청구항 17

제15항에 있어서, 상기 비디오 디코더는 임베딩된 매트릭스로부터 샘플의 양자화된 서브세트를 제공하도록 압축된 심볼을 엔트로피 디코딩하기 위한 엔트로피 디코더, 상기 임베딩된 매트릭스의 회수된 버전(recovered version)으로부터 샘플의 서브세트의 근사를 제공하도록 상기 샘플의 양자화된 서브세트를 역양자화하는 역양자화기, 상기 임베딩된 매트릭스의 완료된 회수 버전(completed recovered version)을 제공하도록 상기 임베딩된 매트릭스의 회수된 버전의 누락 부분을 결정하기 위한 매트릭스 완료 디바이스, 및 상기 임베딩된 매트릭스의 완료된 회수 버전으로부터 상기 모션 매트릭스의 근사를 재구성하기 위한 재구성 디바이스를 포함하는 장치.

청구항 18

제17항에 있어서, 상기 임베딩된 매트릭스의 회수된 버전의 누락 부분은 상기 샘플의 서브세트에 있는 샘플과 다른, 상기 임베딩된 매트릭스에서의 다른 위치를 가지는 상기 임베딩된 매트릭스의 이전에 버려진 부분에 대응하는 장치.

청구항 19

제17항에 있어서, 상기 모션 매트릭스의 근사는 상기 임베딩된 매트릭스와, 상기 임베딩된 매트릭스가 곱해지는 2개의 매트릭스를 사용하여 결정되는 장치.

청구항 20

제17항에 있어서, 동일 랜덤 시드와 동일 랜덤 수 발생기는, 상기 임베딩된 매트릭스로부터의 상기 샘플의 서브세트에 있는 샘플이 상기 비디오 디코더와 대응 비디오 인코더에서 동일한 각각의 위치를 가지는 것을 보장하도록 상기 비디오 디코더와 상기 대응 비디오 인코더에 의해 사용되는 장치.

청구항 21

제15항에 있어서, 상기 모션 매트릭스의 희박 표현은 임베딩된 매트릭스의 회수된 버전에 대해 결정되며, 상기 화상의 벡터화된 형태는 상기 희박 표현과 상기 딕셔너리에 반응하여 결정되고, 상기 벡터화된 형태는 상기 화상의 재구성된 버전을 연도록 재형상화되는 장치.

청구항 22

비디오 디코더에서의 방법으로서,

모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하는 단계를 포함하며, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스의 근사는 주어진 값 아래의 등급 및 딕셔너리에 관한 희박 표현을 가지며, 상기 딕셔너리는 인코딩될 화상의 함수가 아닌 적어도 하나의 기준 프레임의 함수로서 생성된 복수의 딕셔너리들 중에서 선택되고, 상기 선택은 목표 등급 및 목표 희박도를 달성하는 것에 기초하며,

상기 모션 매트릭스는, 상기 선택된 딕셔너리에 관하여 상기 화상의 표현을 반복적으로 찾는 것, 상기 화상의 표현의 가장 큰 엔트리의 상기 목표 희박도의 수만을 유지하는 것, 상기 모션 매트릭스를 갱신하는 것, 상기 화상의 표현이 특정 레벨로 수렴할 때까지 상기 모션 매트릭스의 고유값의 최대 목표 등급 수를 상기 목표 등급의 수까지 유지하는 것에 의해 찾아지는 방법.

청구항 23

제22항에 있어서, 상기 딕셔너리는 상기 화상과 기준 화상의 세트 사이에서의 공간적 및 시간적 상관관계로부터 형성되는 방법.

청구항 24

제22항에 있어서, 임베딩된 매트릭스로부터 샘플의 양자화된 서브세트를 제공하도록 압축된 심볼을 엔트로피 디코딩하는 단계;

상기 임베딩된 매트릭스의 회수된 버전으로부터 샘플의 서브세트의 근사를 제공하도록 상기 샘플의 양자화된 서

브세트를 역양자화하는 단계;

상기 임베딩된 매트릭스의 완료된 회수 버전을 제공하도록 상기 임베딩된 매트릭스의 회수된 버전의 누락 부분을 결정하는 단계; 및

상기 임베딩된 매트릭스의 완료된 회수 버전으로부터 상기 모션 매트릭스의 근사를 재구성하는 단계를 추가로 포함하는 방법.

청구항 25

제24항에 있어서, 상기 임베딩된 매트릭스의 회수된 버전의 누락 부분은 상기 샘플의 서브세트에 있는 샘플과 다른, 상기 임베딩된 매트릭스에서의 위치를 가지는 상기 임베딩된 매트릭스의 이전에 벼려진 부분에 대응하는 방법.

청구항 26

제24항에 있어서, 상기 모션 매트릭스의 근사는 상기 임베딩된 매트릭스와, 상기 임베딩된 매트릭스가 곱해지는 2개의 매트릭스를 사용하여 결정되는 방법.

청구항 27

제24항에 있어서, 동일 랜덤 시드와 동일 랜덤 수 발생기는, 상기 임베딩된 매트릭스로부터 상기 샘플의 서브세트에 있는 샘플이 상기 비디오 디코더와 대응 비디오 인코더에서 동일한 각각의 위치를 가지는 것을 보장하도록 상기 비디오 디코더와 상기 대응 비디오 인코더에 의해 사용되는 방법.

청구항 28

제22항에 있어서, 상기 모션 매트릭스의 희박 표현은 임베딩된 매트릭스의 회수된 버전에 대해 결정되며, 상기 화상의 벡터화된 형태는 상기 희박 표현과 상기 딕셔너리에 반응하여 결정되고, 상기 벡터화된 형태는 상기 화상의 재구성된 버전을 얻도록 재형상화되는 방법.

청구항 29

인코딩된 비디오 신호 데이터를 가지는 컴퓨터 판독 가능한 저장 매체로서,

모션 매트릭스를 사용하여 인코딩된 비디오 시퀀스에서의 화상을 포함하며, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스는 주어진 값 아래의 등급 및 딕셔너리에 관한 희박 표현을 가지며, 상기 딕셔너리는 인코딩될 화상의 함수가 아닌 적어도 하나의 기준 프레임의 함수로서 생성된 복수의 딕셔너리를 중에서 선택되고, 상기 선택은 목표 등급 및 목표 희박도를 달성하는 것에 기초하며,

상기 모션 매트릭스는, 상기 선택된 딕셔너리에 관하여 상기 화상의 표현을 반복적으로 찾는 것, 상기 화상의 표현의 가장 큰 엔트리의 상기 목표 희박도의 수만을 유지하는 것, 상기 모션 매트릭스를 갱신하는 것, 상기 화상의 표현이 특정 레벨로 수렴할 때까지 상기 모션 매트릭스의 고유값의 최대 목표 등급 수를 상기 목표 등급의 수까지 유지하는 것에 의해 찾아지는, 컴퓨터 판독 가능한 저장 매체.

청구항 30

모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하기 위한 비디오 인코더를 포함하되, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스는 주어진 등급 및 벡터를 가지며, 상기 벡터는 딕셔너리와 희박 벡터(sparse vector)의 곱으로 구성되는 상기 모션 매트릭스의 행을 적층하는 것에 의해 형성되며,

상기 딕셔너리는 아톰(atom)의 세트와, 상기 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 상기 화상이 디코딩되도록 허용하기 위한 기본 벡터를 포함하고,

상기 딕셔너리는 영이 아닌(non-zero) 잔여값의 수를 최소화하기 위해 상기 비디오 시퀀스 내의 기준 화상의 세트로부터 형성되며,

상기 모션 매트릭스는,

제1벡터를 찾는 제1단계로서, 상기 딕셔너리와 상기 제1벡터의 곱은 초기값이 상기 화상인 제1매트릭스의 행을 적층하여 형성된 벡터를 나타내는 제1단계;

상기 제1벡터의 가장 큰 엔트리의 주어진 수만을 유지하고, 나머지 엔트리에는 0을 유지하는 제2벡터를 생성하는 제2단계;

상기 제1매트릭스의 행을 적층하여 형성된 벡터가 상기 딕셔너리와 상기 제2벡터의 곱인 제2매트릭스를 생성하는 제3단계;

상기 제2매트릭스의 가장 큰 고유값의 주어진 수만을 상기 주어진 등급으로 유지하고, 나머지 고유값에는 0을 유지하는 제3매트릭스를 생성하는 제4단계; 및

상기 제1매트릭스와 상기 제3매트릭스 사이의 각 고유값의 차이가 주어진 값보다 작을때까지 상기 제1매트릭스에 상기 제3매트릭스를 대입하여 상기 제1단계 내지 제4단계를 반복하고, 상기 각 고유값의 차이가 상기 주어진 값보다 작으면 상기 제1벡터를 상기 희박 벡터로 간주하고 상기 제1매트릭스를 상기 모션 매트릭스로 간주하는 제5단계에 의해 찾아지는 장치.

청구항 31

비디오 인코더에서의 방법으로서,

모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하는 단계를 포함하며, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스는 주어진 등급 및 벡터를 가지며, 상기 벡터는 딕셔너리와 희박 벡터(sparse vector)의 곱으로 구성되는 상기 모션 매트릭스의 행을 적층하는 것에 의해 형성되며,

상기 딕셔너리는 아톰(atom)의 세트와, 상기 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 상기 화상이 디코딩되도록 허용하기 위한 기본 벡터를 포함하고,

상기 딕셔너리는 영이 아닌(non-zero) 잔여값의 수를 최소화하기 위해 상기 비디오 시퀀스 내의 기준 화상의 세트로부터 형성되며,

상기 모션 매트릭스는,

제1벡터를 찾는 제1단계로서, 상기 딕셔너리와 상기 제1벡터의 곱은 초기값이 상기 화상인 제1매트릭스의 행을 적층하여 형성된 벡터를 나타내는 제1단계;

상기 제1벡터의 가장 큰 엔트리의 주어진 수만을 유지하고, 나머지 엔트리에는 0을 유지하는 제2벡터를 생성하는 제2단계;

상기 제1매트릭스의 행을 적층하여 형성된 벡터가 상기 딕셔너리와 상기 제2벡터의 곱인 제2매트릭스를 생성하는 제3단계;

상기 제2매트릭스의 가장 큰 고유값의 주어진 수만을 상기 주어진 등급으로 유지하고, 나머지 고유값에는 0을 유지하는 제3매트릭스를 생성하는 제4단계; 및

상기 제1매트릭스와 상기 제3매트릭스 사이의 각 고유값의 차이가 주어진 값보다 작을때까지 상기 제1매트릭스에 상기 제3매트릭스를 대입하여 상기 제1단계 내지 제4단계를 반복하고, 상기 각 고유값의 차이가 상기 주어진 값보다 작으면 상기 제1벡터를 상기 희박 벡터로 간주하고 상기 제1매트릭스를 상기 모션 매트릭스로 간주하는 제5단계에 의해 찾아지는 방법.

청구항 32

모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하기 위한 비디오 디코더를 포함하되, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스의 근사는 주어진 등급 및 벡터를 가지며, 상기 벡터는 딕셔너리와 희박 벡터(sparse vector)의 곱으로 구성되는 상기 모션 매트릭스의 행을 적층하는 것에 의해 형성되며,

상기 딕셔너리는 아톰(atom)의 세트와, 상기 화상을 표현하고 단지 상기 세트만을 사용하여 상기 비디오 디코더

에서 상기 화상이 디코딩되도록 허용하기 위한 기본 벡터를 포함하고,

상기 딕셔너리는 영이 아닌(non-zero) 잔여값의 수를 최소화하기 위해 상기 비디오 시퀀스 내의 기준 화상의 세트로부터 형성되며,

상기 모션 매트릭스는,

제1벡터를 찾는 제1단계로서, 상기 딕셔너리와 상기 제1벡터의 곱은 초기값이 상기 화상인 제1매트릭스의 행을 적층하여 형성된 벡터를 나타내는 제1단계;

상기 제1벡터의 가장 큰 엔트리의 주어진 수만을 유지하고, 나머지 엔트리에는 0을 유지하는 제2벡터를 생성하는 제2단계;

상기 제1매트릭스의 행을 적층하여 형성된 벡터가 상기 딕셔너리와 상기 제2벡터의 곱인 제2매트릭스를 생성하는 제3단계;

상기 제2매트릭스의 가장 큰 고유값의 주어진 수만을 상기 주어진 등급으로 유지하고, 나머지 고유값에는 0을 유지하는 제3매트릭스를 생성하는 제4단계; 및

상기 제1매트릭스와 상기 제3매트릭스 사이의 각 고유값의 차이가 주어진 값보다 작을때까지 상기 제1매트릭스에 상기 제3매트릭스를 대입하여 상기 제1단계 내지 제4단계를 반복하고, 상기 각 고유값의 차이가 상기 주어진 값보다 작으면 상기 제1벡터를 상기 희박 벡터로 간주하고 상기 제1매트릭스를 상기 모션 매트릭스로 간주하는 제5단계에 의해 찾아지는 장치.

청구항 33

비디오 디코더에서의 방법으로서,

모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하는 단계를 포함하며, 상기 모션 매트릭스는 디코더가 상기 모션 매트릭스 및 기준 프레임을 사용하여 상기 화상을 특유의 형태로 재구성 할 수 있게하는 데이터를 포함하고, 상기 모션 매트릭스의 근사는 주어진 등급 및 벡터를 가지며, 상기 벡터는 딕셔너리와 희박 벡터(sparse vector)의 곱으로 구성되는 상기 모션 매트릭스의 행을 적층하는 것에 의해 형성되며,

상기 딕셔너리는 아톰(atom)의 세트와, 상기 화상을 표현하고 단지 상기 세트만을 사용하여 상기 비디오 디코더에서 상기 화상이 디코딩되도록 허용하기 위한 기본 벡터를 포함하고,

상기 딕셔너리는 영이 아닌(non-zero) 잔여값의 수를 최소화하기 위해 상기 비디오 시퀀스 내의 기준 화상의 세트로부터 형성되며,

상기 모션 매트릭스는,

제1벡터를 찾는 제1단계로서, 상기 딕셔너리와 상기 제1벡터의 곱은 초기값이 상기 화상인 제1매트릭스의 행을 적층하여 형성된 벡터를 나타내는 제1단계;

상기 제1벡터의 가장 큰 엔트리의 주어진 수만을 유지하고, 나머지 엔트리에는 0을 유지하는 제2벡터를 생성하는 제2단계;

상기 제1매트릭스의 행을 적층하여 형성된 벡터가 상기 딕셔너리와 상기 제2벡터의 곱인 제2매트릭스를 생성하는 제3단계;

상기 제2매트릭스의 가장 큰 고유값의 주어진 수만을 상기 주어진 등급으로 유지하고, 나머지 고유값에는 0을 유지하는 제3매트릭스를 생성하는 제4단계; 및

상기 제1매트릭스와 상기 제3매트릭스 사이의 각 고유값의 차이가 주어진 값보다 작을때까지 상기 제1매트릭스에 상기 제3매트릭스를 대입하여 상기 제1단계 내지 제4단계를 반복하고, 상기 각 고유값의 차이가 상기 주어진 값보다 작으면 상기 제1벡터를 상기 희박 벡터로 간주하고 상기 제1매트릭스를 상기 모션 매트릭스로 간주하는 제5단계에 의해 찾아지는 방법.

발명의 설명

기술 분야

[0001] 관련 출원의 상호 참조

[0002] 본 출원은 2010년 10월 14일자 출원된 미국 가출원 제61/393,186호에 대해 우선권을 주장하며, 이 기초 출원은 그의 전문이 참조로 본원에 병합된다.

[0003] 기술분야

[0004] 본 발명은 대체로 비디오 디코딩 및 인코딩에 관한 것이고, 특히 모션 매트릭스(motion matrix)를 사용하여 비디오 인코딩 및 디코딩하기 위한 방법 및 장치에 관한 것이다.

배경 기술

[0005] 비디오 코덱은 디지털 비디오를 위한 비디오 압축 및/또는 압축 해제할 수 있는 디바이스이다. 비디오 품질, 비디오 품질을 나타내는데 필요한 데이터의 양(또한 "비트율(bit rate)"로서 공지된), 인코딩 및 디코딩 알고리즘의 복잡성, 데이터 손실 및 에러의 견고성, 편집의 용이성, 랜덤 액세스(random access), 압축 알고리즘 디자인 기술의 상태, 종대종 지연(end-to-end delay), 및 다수의 다른 팩터들 사이의 복합적인 밸런스(complex balance)가 있다.

[0006] 비디오 코덱은 디지털 포맷에서 근본적으로 아날로그의 데이터 세트를 나타내도록 추구한다. 아날로그 비디오 신호의 디자인이 휘도(luma)와 컬러 정보를 분리하여 나타내기 때문에, 코덱 디자인에서 이미지 압축 시 통상의 첫 단계는 YCbCr 색 공간에서 이미지를 나타내고 저장하는 것이다. YCbCr로의 변환은 다음의 두 가지 장점을 제공한다: 먼저, 변환은 컬러 신호의 비상관관계를 제공하는 것에 의해 압축률을 개선하며; 두 번째로, 변환은, 지각적으로(perceptually) 덜 중요하고 보다 효율적인 데이터 압축을 달성하도록 보다 낮은 해상도로 표현될 수 있는 색 신호(chroma signal)로부터, 지각적으로 훨씬 더 중요한 휘도 신호를 분리한다.

[0007] 디코딩 프로세스는 가능한 최대로 인코딩 프로세스의 각 스테이지의 반전(inversion)을 수행하는 것으로 구성된다. 정확하게 반전될 수 없는 하나의 스테이지는 양자화 스테이지이다. 여기에서, 반전의 최선 노력 근사(best-effort approximation)가 수행된다. 프로세스의 이러한 부분은 양자화가 고유하게 비반전성 프로세스일지라도 흔히 "반전 양자화(inverse quantization)" 또는 "역양자화(de-quantization)"로 지칭된다.

[0008] 비디오를 인코딩하는 전통적인 방법은 기준 및 임의의 나머지(residue)에서 간단한 움직임에 대해 프레임을 분석하는 것을 포함한다. 도 1을 참조하여, 전통적인 인코딩 방법은 대체로 도면부호 (100)으로 지시된다. 인코딩 방법(100)은 인코딩될 프레임(2, 시간(t_2)에서 f_2), (기준) 프레임(1, 시간(t_1)에서 f_1), 및 임의의 나머지를 수반한다. 프레임(2, 시간(t_2)에서 f_2)은 프레임(1, 시간(t_1)에서 f_1)에 대한 기준 및 나머지를 사용하여 인코딩된다. International Organization for Standardization/International Electrotechnical Commission(ISO/IEC) Moving Picture Experts Group-4(MPEG-4) Part 10 Advanced Video Coding(AVC) Standard/International Telecommunication Union, Telecommunication Sector (ITU-T) H.264 Recommendation(이후에, "MPEG-4 AVC Standard"으로서 지칭됨)과 같은 전통적인 하이브리드 비디오 코덱은 이러한 원리에 기초한다. 블록들은 그 안 또는 상호 간에(intra or inter) 예측되고, 그런 다음 변환, 양자화, 엔트로피 코딩, 및 잡재적으로 일부 필터링의 전형적인 프로세스가 적용된다.

[0009] 비디오 코딩은 매우 광범위한 분야이며, 그 주제에 대해 수십년의 연구가 있었다. 다양한 기술이 비디오 코딩에 적용되었다. 압축 센싱(compressive sensing)의 최근의 출현은 이러한 문제에서 사용하는 또 다른 툴을 제공하였다. 예를 들어, 하나의 종래의 접근에서, 압축 센싱은 데이터 블록을 인코딩하는 모드로서 사용된다. 그러나, 이러한 접근은 종래의 체계에 여전히 의지하며, 시스템 내에서 측면 방법(side method)으로서 압축 센싱을 도입한다.

발명의 내용

[0010] 종래의 기술의 이들 및 다른 결함 및 단점은 모션 매트릭스를 사용하여 비디오 인코딩 및 디코딩하는 방법 및 장치에 관한 본 발명에 의해 다루어진다.

[0011] 본 발명의 한 양태에 따라서, 장치가 제공된다. 상기 장치는 모션 매트릭스를 사용하여 비디오 시퀀스에서 화상(picture)을 인코딩하기 위한 비디오 인코더를 포함한다. 모션 매트릭스는 주어진 기준점(threshold) 아래의 등급(rank) 및 딕셔너리(dictionary)에 관한 희박 표현(sparse representation)을 가진다. 딕셔너리는 아톰(atom)의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 화상이 파생되도록 허용하기

위한 기본 벡터를 포함한다. 딕셔너리는 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

[0012] 본 발명의 또 다른 양태에 따라서, 비디오 인코더에서 방법을 제공한다. 상기 방법은 모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하는 단계를 포함한다. 모션 매트릭스는 주어진 기준점 아래의 등급 및 딕셔너리에 관한 희박 표현을 가진다. 딕셔너리는 아톰의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 화상이 파생되도록 허용하기 위한 기본 벡터를 포함한다. 딕셔너리는 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

[0013] 본 발명의 또 다른 양태에 따라서, 장치가 제공된다. 상기 장치는 모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하기 위한 비디오 디코더를 포함한다. 모션 매트릭스의 근사는 주어진 기준점 아래의 등급 및 딕셔너리에 관한 희박 표현을 가진다. 딕셔너리는 아톰의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 비디오 디코더에서 화상이 파생되도록 허용하기 위한 기본 벡터를 포함한다. 딕셔너리는 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

[0014] 본 발명의 여전히 또 다른 양태에 따라서, 비디오 디코더에서 방법이 제공된다. 상기 방법은 모션 매트릭스의 근사를 사용하여 비디오 시퀀스에서 화상을 디코딩하는 단계를 포함한다. 모션 매트릭스의 근사는 주어진 기준점 아래의 등급 및 딕셔너리에 관한 희박 표현을 가진다. 딕셔너리는 아톰의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 비디오 디코더에서 화상이 파생되도록 허용하기 위한 기본 벡터를 포함한다. 딕셔너리는 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

[0015] 본 발명의 이들 및 다른 양태, 특징 및 장점들은 첨부된 도면과 관련하여 읽혀지는 예시적인 실시예의 다음의 상세한 설명으로부터 명백하게 될 것이다.

[0016] 본 발명은 다음의 예시적인 도면에 따라서 더욱 잘 이해될 수 있다.

도면의 간단한 설명

[0017] 도 1은 종래 기술에 따라서, 전통적인 인코딩 방법을 도시한 도면;

도 2는 본 발명의 실시예에 따라서, 본 발명이 적용될 수 있는 예시적인 비디오 인코더를 도시한 블록도;

도 3은 본 발명의 실시예에 따라서, 본 발명이 적용될 수 있는 예시적인 비디오 디코더를 도시한 블록도;

도 4는 본 발명의 실시예에 따라서, 모션 매트릭스를 벡터화하는 예시적인 작업 및 벡터화한 작업의 반전을 도시한 도면;

도 5는 본 발명의 실시예에 따라서, 모션 매트릭스 기반 비디오 코딩 체계를 사용하여 인코딩하기 위한 예시적인 방법을 도시한 흐름도;

도 6은 본 발명의 실시예에 따라서, 모션 매트릭스 기반 비디오 코딩 체계를 사용하여 디코딩하기 위한 예시적인 방법을 도시한 흐름도;

도 7은 본 발명의 실시예에 따라서, 저등급 조밀 매트릭스(Z, low-rank dense matrix) 내로 모션 매트릭스를 임베딩(embedding)하기 위한 예시적인 방법을 도시한 도면;

도 8은 본 발명의 실시예에 따라서, 모션 매트릭스 및 매트릭스 완료(matrix completion)로부터 샘플링하기 위한 예시적인 방법을 도시한 도면;

도 9는 본 발명의 실시예에 따라서, 대응 고유값(corresponding Eigen-value)이 충분히 랜덤할 때 저등급 매트릭스 완료의 예를 도시한 도면;

도 10은 본 발명의 실시예에 따라서, 충분히 랜덤하지 않은 대응 고유값으로 인하여 실패한 저등급 매트릭스의 예를 도시한 도면; 및

도 11은 본 발명의 실시예에 따라서, 변형된 프레임(f_q')의 희박 표현의 압축 샘플로부터 형성된 임베딩된 매트릭스(Z)의 실시예를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0018] 본 발명은 모션 매트릭스를 사용하여 비디오 인코딩 및 디코딩하기 위한 방법 및 장치에 관한 것이다.

- [0019] 본 설명은 본 발명을 예시한다. 그러므로, 비록 본 명세서에서 명시적으로 기술 또는 도시되지 않았을지라도, 당업자가 본 발명을 구현하고 그 사상 및 범위에 포함되는 다양한 장치를 고안할 수 있다는 것이 예측될 것이다.
- [0020] 본 명세서에서 인용된 모든 예 및 조건어는 기술을 발전시키도록 본 발명 및 본 발명자(들)에 의해 기여된 개념들을 독자가 이해하는 것을 돋는 교육 목적을 위해 의도되고, 이러한 특별하게 인용된 예 및 조건에 대해 제한이 없는 것으로서 해석되어야 한다.
- [0021] 더욱이, 본 발명의 원리들, 양태들, 및 실시예들을 인용하는 본 명세서의 모든 설명뿐만 아니라 그 특정 예는 그 구조적 및 기능적 등가물들 모두를 포함하도록 의도된다. 부가적으로, 이러한 등가물들이 현재 공지된 등가물뿐만 아니라 장래에 개발되는 등가물들, 즉 구조에 관계없이 동일한 기능을 수행하는 개발되는 임의의 요소 모두를 포함하도록 의도된다.
- [0022] 그러므로, 예를 들어, 당업자는, 본 명세서에 제공된 블록도들이 본 발명을 구현하는 예시적인 회로의 개념적 관점을 나타낸다는 것을 예측할 것이다. 유사하게, 임의의 플로우 차트, 흐름도, 상태 천이도, 의사코드 등이 컴퓨터 판독 가능한 매체에서 실질적으로 표현될 수 있고, 그래서, 컴퓨터 또는 프로세서가 명시적으로 도시되어 있는지 아닌지 간에 이러한 컴퓨터 또는 프로세서에 의해 실행될 수 있는 다양한 프로세스들을 나타낸다는 것을 예측할 것이다.
- [0023] 도면에 도시된 다양한 요소들의 기능들은 전용 하드웨어뿐만 아니라 적절한 소프트웨어와 관련하여 소프트웨어를 실행할 수 있는 하드웨어의 사용을 통해 제공될 수 있다. 프로세서에 의해 제공될 때, 기능은 단일의 전용 프로세서에 의해, 단일의 공유된 프로세서에 의해, 또는 그 일부가 공유될 수 있는 다수의 개별적인 프로세서들에 의해 제공될 수 있다. 더욱이, 용어 "프로세서" 또는 "컨트롤러"의 명시적인 사용은 소프트웨어를 실행할 수 있는 하드웨어를 배타적으로 지칭하는 것으로 해석되지 않아야 하며, 디지털 신호 프로세서("DSP") 하드웨어, 소프트웨어를 저장하기 위한 판독전용 메모리("ROM"), 랜덤 액세스 메모리("RAM") 및 비휘발성 기억 장치를 제한없이 내재적으로 포함할 수 있다.
- [0024] 종래 및/또는 통상적인 다른 하드웨어가 또한 포함될 수 있다. 유사하게, 도면에 도시된 임의의 스위치들은 단지 개념적이다. 스위치의 기능은 프로그램 로직의 동작을 통해, 전용 로직을 통해, 프로그램 컨트롤 및 전용 로직의 상호 작용을 통해, 또는 심지어 수동으로 수행될 수 있고, 특정 기술은 문맥으로부터 더욱 구체적으로 이해되는 바와 같이 실시자에 의해 선택 가능하다.
- [0025] 본 발명의 청구항에서, 특정 기능을 수행하기 위한 수단으로서 표현된 임의의 요소는 예를 들어, a) 그 기능을 수행하는 회로 소자들의 조합, 또는 b) 그 기능을 수행하는 소프트웨어를 실행하기 위한 적절한 회로와 조합된, 펌웨어, 마이크로코드 등을 포함하는 임의의 형태로 된 소프트웨어를 포함하는 기능을 수행하는 임의의 방식을 포함하도록 의도된다. 이러한 청구항에 의해 한정된 바와 같은 본 발명은, 다양한 인용 수단에 의해 제공된 기능들이 청구항들이 요구하는 방식으로 조합되고 합쳐진다는 사실에 있다. 그러므로, 이들 기능들을 제공할 수 있는 임의의 수단은 도면에 도시된 것들과 등가인 것으로 간주된다.
- [0026] 본 발명의 "하나의 실시예" 또는 "실시예" 뿐만 아니라 그 변형에 대한 명세서에서의 기준은, 그 실시예와 관련하여 기술된 특정의 특징, 구조, 특성 등이 본 발명의 적어도 하나의 실시예에 포함된다는 것을 의미한다. 그러므로, 본 명세서 전체에 걸쳐 다양한 곳에서 보이는 구문 "하나의 실시예에서", 또는 "실시예에서" 뿐만 아니라 그 임의의 다른 변형의 출현은 반드시 동일한 실시예를 모두 지칭하는 것은 아니다.
- [0027] 예를 들면 "A/B", "A 및/또는 B", 및 "A 및 B 중 적어도 하나"의 경우에서, 다음의 "/", " 및/또는", 및 " 중 적어도 하나"의 임의의 하나의 사용은 첫번째 열거된 옵션(A)만의 선택, 두번째 열거된 옵션(B)만의 선택, 또는 양쪽 옵션(A 및 B)의 선택을 포용하도록 의도되는 것은 자명하다. 하나의 추가 예로서, "A, B 및/또는 C" 및 "A, B 및 C 중 적어도 하나"의 경우에서, 이러한 구문들은 첫번째 열거된 선택(A)만의 선택, 또는 두번째 열거된 옵션(B)만의 선택, 또는 세번째 열거된 옵션(C)만의 선택, 또는 첫번째와 두번째 열거된 옵션(A 및 B)만의 선택, 또는 첫번째와 세번째 열거된 옵션(A 및 C)만의 선택, 또는 두번째와 세번째 열거된 옵션(B 및 C)만의 선택, 또는 3가지 옵션(A 및 B 및 C) 모두의 선택을 포용하도록 의도된다. 이러한 것은 당업자 및 관련 기술분야의 통상의 지식을 가진 자에게 자명한 것으로서, 열거된 다수의 물품에 대해 확장될 수 있다.
- [0028] 또한, 여기에 사용되는 바와 같이, 단어들 "화상" 및 "이미지"는 호환 가능하게 사용되고, 일반적으로는 비디오 시퀀스로부터의 정지 이미지 또는 화상을 지칭한다. 공지된 바와 같이, 화상은 프레임 또는 펠드일 수 있다.

- [0029] 도 2를 참조하여, 본 발명이 적용될 수 있는 예시적인 비디오 인코더는 대체로 도면부호 (200)으로 지시된다. 도 2가 단지 본 출원인의 제안된 비디오 인코딩 접근의 근간을 포함하고 디블록킹 필터(de-blocking filter), 예측(prediction) 등과 같은 블록에 의해 더욱 개선될 수 있다는 것을 유념하는 것은 중요하다. 비디오 인코더 (200)는 모션 매트릭스 발생기(220)의 제1입력과 신호 통신으로 연결되는 출력을 가지는 딕셔너리 발생기(210)를 포함한다. 모션 매트릭스 발생기(220)의 출력은 매트릭스 임베더(matrix embedder)(230)의 입력과 신호 통신으로 연결된다. 매트릭스 임베더(230)의 출력은 샘플러(240)의 입력과 신호 통신으로 연결된다. 샘플러(240)의 출력은 양자화기(250)의 입력과 신호 통신으로 연결된다. 양자화기(250)의 출력은 엔트로피 인코더(260)의 입력과 신호 통신으로 연결된다. 딕셔너리 발생기(210)의 입력과 모션 매트릭스 발생기(220)의 제2입력은 기준 프레임을 수신하기 위한 인코더(200)의 입력으로서 이용할 수 있다. 모션 매트릭스 발생기(220)의 제3입력은 인코딩될 프레임을 수신하기 위한 인코더(200)의 입력으로서 이용할 수 있다. 엔트로피 인코더(260)의 출력은 인코딩된 비트스트림을 출력하기 위한 인코더(200)의 출력으로서 이용할 수 있다.
- [0030] 도 3을 참조하여, 본 발명이 적용될 수 있는 예시적인 비디오 디코더는 대체로 도면부호 (300)으로 지시된다. 도 3이 단지 본 출원인의 제안된 비디오 디코딩 접근의 근간을 포함하고 디블록킹 필터, 예측 등과 같은 블록에 의해 더욱 개선될 수 있다는 것을 유념하는 것은 중요하다. 비디오 디코더(300)는 역양자화기(320)의 입력과 신호 통신으로 연결된 출력을 가지는 엔트로피 디코더(310)를 포함한다. 역양자화기(320)의 출력은 매트릭스 완료기(matrix completer)(330)의 입력과 신호 통신으로 연결된다. 매트릭스 완료기(330)의 출력은 모션 매트릭스 회수기(motion matrix recoverer)(340)의 입력과 신호 통신으로 연결된다. 모션 매트릭스 회수기(340)의 출력은 이미지 재구성기(350)의 제1입력과 신호 통신으로 연결된다. 딕셔너리 발생기(360)의 출력은 이미지 재구성기 (350)의 제2입력과 신호 통신으로 연결된다. 엔트로피 디코더(310)의 입력은 입력 비트스트림을 수신하기 위한 디코더(300)의 입력으로서 이용할 수 있다. 딕셔너리 발생기(360)의 입력은 기준 프레임을 수신하기 위한 디코더(300)의 입력으로서 이용할 수 있다. 이미지 재구성기(350)의 출력은 하나 이상의 재구성된 이미지를 출력하기 위한 디코더(300)의 출력으로서 이용할 수 있다. 이미지 재구성기(350)는 기준 프레임을 수신하기 위한 하나 이상의 추가적인 입력을 포함할 수 있다.
- [0031] 상기된 바와 같이, 본 발명은 모션 매트릭스를 사용하여 비디오 인코딩 및 디코딩하기 위한 방법 및 장치에 관한 것이다.
- [0032] 높은 압축 효율을 달성하도록, 사실상 모든 비디오 인코딩 체계는 다음과 같은 것을 활용한다:
- [0033] (a) 시간적 상관관계 또는 연이은 프레임들 중에서의 상관관계, 및
- [0034] (b) 공간적 상관관계 또는 동등하게 프레임 및/또는 화상 내에서 인접한 픽셀(또는 그렇지 않으면 인근의 픽셀) 중에서의 상관관계.
- [0035] 높은 공간적 상관관계는, 본 발명자들이 매트릭스 형태에서 $N \times M$ 픽셀 디지털 이미지(위치(i, j)에 서의 픽셀 값이 매트릭스의 (i, j)th 엔트리에 저장되는)를 고려하면, 그 이미지에 대한 픽셀 값은 강한 상관관계를 가지며 독립적이지 않다는 것을 의미한다. 수학적인 관점으로부터, 이러한 것은 대용 매트릭스가 저등급 또는 낮은 독립성을 가진다는 것을 의미한다. 한편, 높은 시간적 상관관계는 특정 시간에서 프레임의 주요 부분이 이전의 프레임에서(아마 다른 위치에 있는) 정확하게 또는 대략적으로 발견될 수 있다는 것을 의미한다. 본 발명에 따라서, 본 발명자들은 압축 션싱의 수학적 툴 및 저등급 매트릭스 완료를 이용하는 새로운 시스템을 개발하였다. 본 발명자들은 이러한 새로운 접근을 개념 "모션 매트릭스"로서 지칭하고, 이는 비디오 시퀀스를 인코딩하도록 높은 시간적 및 공간적 상관관계를 이용하기 위한 대안적인 접근으로서 표현한다.
- [0036] 본 발명자들은 모션 매트릭스의 원리와, 모션 매트릭스가 비디오 인코더 및 디코더에서 어떻게 실행되는지를 개시하고 기술한다. 또한, 제안된 모션 매트릭스의 예/실시예가 제공되며, 대용 인코더/디코더의 다른 구성요소가 기술된다. 상기 체계는 본 명세서에 기술된 몇 개의 새로운 코딩 블록을 포함한다.
- [0037] 또한, 보편성의 손실없이 설명의 용이성을 위하여, 본 발명은 단지 하나의 기준 프레임이 있는 P-프레임의 상황에서 기술된다. 명확하게, 아이디어는 다수의 기준 프레임을 가진 경우로 용이하게 확장될 수 있다.
- [0038] **개요**
- [0039] 본 발명에 따라서, 본 발명자들은 비디오 컨텐츠를 인코딩 및 디코딩하는 시스템을 기술한다. 기술의 명료성의 목적을 위하여 특정 상황 및 실시예에서 기술되었을지라도, 다른 실시예들은 본 명세서에 기술된 하나 이상의 동일 원리를 이용하고, 그러므로 본 발명의 주석 아래 놓이는 것으로부터 배제되지 않는다는 것이 당업자에게는

자명하게 될 것이다. 예를 들어, 디코더에서의 압축 센싱 모듈은 "기본 추구"(basis pursuit) 알고리즘을 사용하여 달성될 수 있지만, 동일 목적을 달성할 수 있는 다른 공통의 알고리즘(뿐만 아니라 추후에 개발되는 알고리즘)이 문헌에 존재할 수 있으며, 이것들은 본 발명의 일반적인 원리와 범위 하에 포함되도록 의도된다. 본 발명자들은 기본 추구(BP) 알고리즘이 컨벡스 최적화(convex optimization) 문제를 해결하는 단계를 수반한다는 것에 주목했다. 특히, BP는 최소 L1 norm(절대값의 합)에 의한 선형 방정식 $y = Px$ 의 열결정계(under-determined system)에 대한 해결책을 찾아냈으며, 상기 식에서 y 의 값과 계수 P 의 매트릭스가 주어지고, x 는 미지의 벡터이다. 더욱이, 선형 방정식의 열결정계에 의해, 본 발명자들은 방정식의 수(여기에서, P 의 매트릭스의 열들의 수)는 미지의 수(미지의 벡터(x)의 길이)의 수보다 작다는 것을 의도한다. 수학적 형태에서, BP는 다음의 최적화 문제를 해결한다:

$$\text{Arg} \min \|x\|_1 = \sum |x_i| \text{ 이어서 } y = Px$$

이는 특정 조건하에서, BP에 의해 찾아낸 해결책이 가장 희박한 해결책, 즉 제로 이외의 최소수를 가진 해결책일 것으로 판명되었다.

비디오 시퀀스로부터 시간(t_1)에서의 이미지(f_1)를 고려한다. 이러한 것을 가정하면, 이미지는 예를 들어 도 1에 도시된 바와 같이 시간(t_2)에서 f_2 로 변한다. 개략적으로 이야기하면, 가장 전통적인 비디오 인코딩 체계에서, 이러한 움직임은 다음과 같이 분석된다:

$$f_2 = X(f_1) + R$$

여기에서, $X(\cdot)$ 은 기본 함수이고 R 은 임의의 나머지이다. 예를 들어, 도 1의 예에 관하여, $X(\cdot)$ 은 입력으로서 (f_1)을 취하고 그런 다음 출력을 위하여 하나의 픽셀만큼 이미지(f_1)를 우측으로 시프팅하는 함수이며, R 은 나머지이다. 압축은 $X(\cdot)$ 과 R 을 효율적인 방식으로 인코딩하는 것에 의해 얻어지게 된다. 본 예로 되돌아가, $X(\cdot)$ 는 모션 벡터로서 인코딩될 수 있으며, R 은 이산 코사인 변환(discrete cosine transform, DCT)과 같은 변환 영역(transform domain)에서 R 을 표현하는 것에 의해 효율적으로 압축될 수 있다. 인코딩의 이러한 포괄적 프레임, 즉 기준 프레임에서 임의의 움직임과 일부 나머지에 대한 이미지를 분석하는 것은 압축을 달성하도록 시간적 상관관계를 광범위하게 이용한다는 것을 유념한다. 공간적 상관관계는 나머지의 변환 코딩에 의해 또한 이용된다. 다른 한편으로, DCT, 웨이브릿(wavelet) 등과 같은 이용할 수 있는 변환은 나머지 대신, 매우 상관되는 이미지를 희박화하도록(희박을 만들도록) 디자인된다. 다시 말하면, 나머지를 인코딩하도록 변환을 이용하는 것이 과거에 성공적이었을지라도, 공간적 상관관계는 비디오 데이터를 압축하는 능력을 한층 더 향상시키도록 추후의 개발되는 체계에서 사용될 수 있다.

본 발명의 원리에 따라서, 그리고 선형 방정식의 열결정계와 또한 저등급 매트릭스 완료의 문제를 위해 최근에 개발된 개발품을 사용하는, 본 발명자들이 기술하는 방법 및 장치는 대안적인 인코딩 및 디코딩 체계를 디자인하도록 사용된다.

모션 매트릭스

인코딩될 프레임(예를 들어, 도 1에서 f_2)을 위한 모션 매트릭스는 매트릭스(Y)이다. 디코더는 모션 매트릭스와 기준 프레임(예를 들어, f_1)을 사용하여 인코딩된 프레임을 특유의 형태로 재구성할 수 있다. 또한, 모션 매트릭스는 다음의 2개의 부가적인 핵심적 특성을 가진다.

• 모션 매트릭스는 저등급을 가진다. 다시 말하면, Y 의 엔트리는 매우 상관된다.

• 모션 매트릭스는 딕셔너리(Ψ), 즉 인코딩될 프레임을 아는 것(knowledge)을 요구함이 없이 디코더에서 파생될 수 있는 아톰의 세트/기본 벡터(예를 들어, 이산 코사인 변환은 이미지를 표현하기 위한 딕셔너리이다)에 대한 희박 표현(α)을 가지며; 즉 $\text{Vec}(Y) = \Psi \alpha$ 이고, 여기에서, $\text{Vec}(\cdot)$ 는 매트릭스의 행을 적층하는 것에 의해 형성된 매트릭스의 벡터화된 형태(도 4 참조)를 나타내고, α 는 그 성분이 주로 0인 희박 벡터(sparse vector)이다.

여기에서, 본 발명자들은 본 명세서에서 도입된 모션 매트릭스가 압축을 달성하도록 선택된 개념 및 접근 모두에서 종래의 작업의 변위 매트릭스와 모션 매트릭스에 있는 것들로부터 출발한다는 것에 유의하여야 한다. 다음에, 본 발명자들은 일부 핵심적인 차이를 강조한다:

• 변위 및 종래의 모션 매트릭스는 주로 이진(binary) 매트릭스를 이용한다. 다른 한편으로, 제안된 모션 매트

릭스의 엔트리는 음 및 부동소수점수(floating point number)를 포함하는 임의의 수를 취할 수 있다.

[0052]

- 종래의 작업에서 변위 및 모션 매트릭스의 엔트리는 일부 블록 및/또는 픽셀에서 임의의 종류의 모션의 존재를 나타낸다. 다른 한편으로, 제안된 매트릭스의 엔트리는 일부 블록에서의 움직임을 나타내는 역할로 제한되지 않는다. 예를 들어, 제안된 모션 매트릭스의 엔트리는 전체 프레임에서 일부 글로벌한 움직임(프레임에서의 모든 움직임의 합수)을 알릴 수 있다.

[0053]

- 종래의 작업에서, 기준 프레임이 주어지면, 모션 및/또는 변위 매트릭스는 통상적으로 프레임을 재구성하는 데 충분하지 않다. 예를 들어, 디코더측에서 필요한 재구성을 위해 나머지 매트릭스 또는 다른 시그널링과 같은 오버헤드 데이터(overhead data)를 전송할 필요가 있을 수 있다. 본 발명자들의 방법은, 기준 프레임이 주어지면, 필요한 에러를 가진 이미 코딩된 프레임을 근사치로 계산하도록 모션 매트릭스를 가지는 것이 오직 충분한 독립적 모션 매트릭스에 관한 것이다.

[0054]

- 전통적인 모션 및/또는 변위 매트릭스는 이것들이 (통상) 본 발명자들의 정의에서 모든 필요한 조건을 위반하기 때문에 본 발명자들의 제안된 모션 매트릭스 프레임워크에 놓이지 못한다. 예를 들어, 대각선이 일부 픽셀 만큼 우측 또는 좌측으로 시프팅하는 프레임 또는 블록을 고려한다. 그런 다음, 대웅 변위 매트릭스가 완전 등급(최대 가능한 등급)인 대각 매트릭스일 것이라는 것을 찾는 것은 복잡하지 않다. 더욱이, 이러한 대각 매트릭스(또는 동등하게 이미지)는 DCT, 웨이브릿 등과 같은 전통적인 변환에서 희박하지 않다(이러한 종류의 매트릭스에 적절한 임의의 지향성 변환(directional transform)이 있을 수 있을지라도, 그러나, 본 발명자들은 여분의 오버헤드로 바뀌는 디코더로 이러한 변환/딕셔너리의 파라미터를 전송하는 것을 필요로 한다). 그러므로, 전통적인 변위/모션 매트릭스는 모션 매트릭스를 위한 2개의 필요한 조건(즉, 저등급의 존재 및 희박화 영역의 존재)을 쉽게 위반할 수 있다.

[0055]

- 또 다른 차이는 디코딩 스테이지에서 나타난다. 특히, 전통적인 인코딩 방법은 압축을 달성하도록 예측 및 나머지 코딩(residue coding)을 이용한다. 다른 한편으로, 본 발명자들은 매트릭스 완료와 같은 최근에 개발된/해결된 알고리즘/문제를 이용하고, 유사한 목적을 달성하도록 (압축된 센싱을 사용하여) 선형 방정식의 열결정계를 해결한다.

[0056]

마지막으로 그러나 역시 주요한 것이지만, 전통적인 변위 및 모션 매트릭스는 전형적으로 상이한 프레임 중에서 시간적 상관관계의 함수이며, 프레임 내에서 인접한 픽셀 중에서 공간적 상관관계(있으면)에 덜 민감하다. 그러나, 본 발명자들의 제안된 모션 매트릭스는 비디오 컨텐츠를 압축하도록 시간적 및 공간적 상관관계를 동시에 공동으로 고려한다.

[0057]

도 4를 참조하여, 모션 매트릭스를 벡터화하는 예시적인 작업 및 벡터화 작업의 반전은 대체로 도면부호 (400)으로 지시된다. 벡터화된 형태는 도면부호 (410)으로 지시되고, 벡터화된 형태의 반전은 도면부호 (420)으로 지시된다.

[0058]

모션 매트릭스는 높은 공간적 상관관계(저등급 매트릭스로 만드는) 및 높은 시간적 상관관계(희박화 프레임의 존재를 이끄는)를 공동으로 활용하는 것에 의해 형성될 수 있다. 모션 매트릭스의 형성이 이전의 접근으로 한정되지 않고, 그러므로 모션 매트릭스를 형성하기 위한 다른 접근이 본 명세서에 제공된 본 발명의 교시에 따라서 또한 사용될 수 있는 한편, 본 발명의 사상을 유지한다는 것이 예측될 것이다.

[0059]

이는 모션 매트릭스 기반 비디오 인코딩 체계를 위한 대웅 인코더와 디코더의 제시를 따른다.

[0060]

도 5를 참조하여, 모션 매트릭스 기반 비디오 코딩 체계를 사용하여 인코딩하기 위한 예시적인 방법은 대체로 도면부호 (500)으로 지시된다. 방법(500)은 기능 블록(510)과 기능 블록(515)에 제어를 패스하는 시작 블록(505)을 포함한다. 기능 블록(510)은 기준 $\{f_1, \dots, f_{q-1}\}$ 프레임을 입력하고, 기능 블록(520)으로 제어를 패스한다. 기능 블록(520)은 인코딩될 프레임(f_q)을 입력하고, 기능 블록(520)으로 제어를 패스한다. 기능 블록(520)은 인코딩될 프레임(f_q)을 입력하고, 기능 블록(525)으로 제어를 패스한다. 기능 블록(525)은 모션 매트릭스를 찾아내고, 기능 블록(530)으로 제어를 패스한다. 기능 블록(530)은 압축하는 것이 더욱 용이한 매트릭스(가능한 작은) 내로 모션 매트릭스를 임베딩하고, 기능 블록(535)으로 제어를 패스한다. 기능 블록(535)은 임베딩된 매트릭스로부터 샘플링하고(즉, 엔트리의 서브세트(subset)를 유지하고 잔여분(rest)을 버린다), 기능 블록(540)으로 제어를 패스한다. 기능 블록(540)은 임베딩된 매트릭스의 샘플을 양자화하고, 기능 블록(545)으로 제어를 패스한다. 기능 블록(545)은 임베딩된 매트릭스의 양자화된 샘플을 엔트로피 코딩하고, 기능 블록(550)으로 제어를 패스한다. 기능 블록(550)은 임베딩

된 매트릭스의 엔트로피 코딩된 샘플을 저장 및/또는 전송하고, 종료 블록(599)으로 제어를 패스한다.

[0061] 기술된 모션 매트릭스를 위한 인코딩 방법(500)은 5개의 주요 구성요소를 포함한다:

1. 인코딩될 프레임과 일련의 기준 프레임이 주어지면, 덕셔너리(Ψ)를 형성하고(인코딩될 프레임은 덕셔너리에 대해 희박하게 되며), 모션 매트릭스(Y)를 찾아낸다. 상기된 바와 같이, 모션 매트릭스(Y)와 기준 프레임(및, 그러므로 희박화한 덕셔너리(Ψ))을 가지면, 디코더는 인코딩된 프레임을 재구성할 수 있게 된다.

2. 임베딩된 매트릭스($Z = p_1Y p_2$) 내로 모션 매트릭스를 임베딩하고, 여기에서, p_1 및 p_2 는 2개의 매트릭스이다. 디자인에 의해, 임베딩된 매트릭스(Z)는 모션 매트릭스의 모든 필요한 특성들, 즉 추가의 압축을 달성하도록 이용될 수 있는 것을 위한 저등급 특성과 희박화한 기본의 존재를 물려받는다. 전개된 매트릭스(p_1 및 p_2)는 압축 시 본 발명자들을 또한 도울 수 있다. 예를 들어, p_1 및 p_2 가 각각 넓고 높은(tall) 매트릭스이면, 임베딩된 매트릭스(Z)는 압축의 다른 레벨로 바뀌는 모션 매트릭스보다 작게 된다. 이러한 단계의 목적은, Z 의 많은 엔트리들이 안전하게 버려지는 한편, 디코더가 버려진 엔트리를 회수할 수 있도록 하는 것이다. 또한, 이 단계는 추후에 기술된 바와 같은 디코더의 단계 3 및 4의 성공을 위해 매우 중대하다. (a) Z , p_1 및 p_2 를 가지면, 본 발명자가 모션 매트릭스(Y)를 다시 재구성할 수 있고, (b) Z 의 고유 벡터가 충분히 랜덤하도록, 매트릭스(p_1 및 p_2)가 선택되는 것을 유념하는 것은 중요하다. 예를 들어, p_1 및 p_2 가 각각 넓은 랜덤 매트릭스 및 높은 랜덤 매트릭스이면, 비록 임베딩된 매트릭스(Z)가 모션 매트릭스(Y)보다 작지만, 매트릭스(p_1 및 p_2)의 이러한 랜덤 선택이 압축의 측면에서 최적의 접근이 아닐지라도, 상기된 두 필요 조건이 충족되게 된다.

3. 샘플링: 임베딩된 매트릭스(Z)의 엔트리의 (가능한 랜덤한) 서브세트(S)를 유지하고, 엔트리의 잔여분을 버린다. 본 발명자들이 나머지(또는 이미지 자체)의 상부의 가장 큰 DCT 계수를 유지하고 이러한 큰 크기의 엔트리의 값과 위치를 신호로 알리는 전통적인 비디오 체계에서의 압축 단계를 이러한 단계가 연상시키는 것을 생각할 수 있다. 다른 한편으로, 본 발명자의 체계에서 이러한 단계는 완전히 랜덤하며, 다른 엔트리로 임베딩된 매트릭스의 임의의 엔트리를 구별하지 못한다. 결과적으로, 인코더와 디코더가 동일 랜덤 시드(random seed) 및 랜덤 수 발생기를 사용하면, 이것들은 Z 의 엔트리의 위치에 모두 일치한다. 그러므로, 이러한 엔트리의 위치를 신호로 알릴 필요가 없게 되고, 인코더는 단지 샘플 값만을 저장/전송할 수 있다.

4. 임베딩된 매트릭스로부터의 샘플(S)을 \hat{S} 로 양자화한다. 명확하게, 이것은 비가역 압축 단계이다. 그러나, 이 단계는 사실상 모든 압축 기술에 대해 필수적이다.

5. 임베딩된 매트릭스의 양자화된 샘플(\hat{S})을 압축된 심볼(C)로 엔트로피 인코딩한다. 심볼(C)은 저장되거나 또는 디코더로 전송되게 된다.

도 6을 참조하여, 모션 매트릭스 기반 비디오 코딩 체계를 사용하여 디코딩하기 위한 예시적인 방법은 대체로 도면부호 (600)으로 지시된다. 방법(600)은 기능 블록(610)과 기능 블록(645)으로 제어를 패스하는 시작 블록(605)을 포함한다. 기능 블록(610)은 인코딩된 심볼(C)을 입력하고, 기능 블록(615)으로 제어를 패스한다. 기능 블록(615)은 수신된 심볼(C)을 엔트로피 디코딩하고, 기능 블록(620)으로 제어를 패스한다. 기능 블록(620)은 디코딩된 심볼(C)을 역양자화하고, 기능 블록(625)으로 제어를 패스한다. 기능 블록(625)은 매트릭스 완료를 실행하고, 기능 블록(630)으로 제어를 패스한다. 기능 블록(630)은 임베딩된 매트릭스로부터 모션 매트릭스를 회수하고, 기능 블록(635)으로 제어를 패스한다. 기능 블록(645)은 기준 프레임(f_1, \dots, f_{q-1})을 입력하고, 기능 블록(650)으로 제어를 패스한다. 기능 블록(650)은 덕셔너리를 발생시키고, 기능 블록(635)으로 제어를 패스한다. 기능 블록(635)은 출력(f_q)으로 모션 매트릭스를 디코딩하고, 종료 블록(699)으로 제어를 패스한다.

[0068] 디코딩 방법(600)은 5개의 주요 구성요소를 가진다:

1. 임베딩된 매트릭스(Z)의 양자화된 샘플(\hat{S})을 구하도록 압축된 심볼(C)을 엔트로피 디코딩한다.

2. 임베딩된 매트릭스의 샘플의 근사(\bar{S})를 구하도록 샘플(\hat{S})을 역양자화한다.

[0071] 3. 임베딩된 매트릭스(Z)의 일부 엔트리의 근사(\bar{S})를 가지면, 매트릭스 완료를 수행하는 것에 의해 임베딩된 매트릭스(Z)의 버려진 엔트리를 회수한다. 이 단계가 인코더에서 수행되는 샘플링(단계 2)의 반전이라는 것을 유념한다. 버려진 엔트리가 양자화의 비가역 단계로 인하여 정확하게 회수되지 않을 가능성이 상당하다. 그러므로, 본 발명자들은 \hat{Z} 로 이 스테이지의 출력을 인용한다.

[0072] 4. \hat{Z} (이전 단계에서 회수된(recorverd) 임베딩된 매트릭스)로부터 \hat{Y} (모션 매트릭스의 근사)를 재구성한다. 모션 매트릭스가 선형 방정식의 시스템(가능하게 열결정계)을 형성하도록 덕셔너리(ψ , 즉, $\text{Vec}(\hat{Y}) \square \text{Vec}(Y) = \psi \alpha$)에 대하여 회박 표현(α)을 가진다는 사실을 이용할 수 있다. 특히, \hat{Z} 를 가지면, 본 발명자들은 다음에 대한 해결책($\hat{\alpha}$)을 찾아낸다:

$$\text{Vec}(\hat{Z}) \approx \text{Vec}(Z) = \text{Vec}(p_1 Y p_2) = (p_2^T \otimes p_1) \text{Vec}(Y) = (p_2^T \otimes p_1) \psi \alpha$$

[0073] 이러한 해결책은 압축된 센싱 디코더(기본 추구 또는 직교 매칭 추구(Orthogonal Matching Pursuit)와 같은)를 사용하는 것에 의해 구해질 수 있다. 결과적으로, 모션 매트릭스(에 대한 근사)는 $\hat{Y} = \text{Vec}^{-1}(\psi \hat{\alpha})$ 에 의해 추정될 수 있으며, 여기에서 $\text{Vec}^{-1}(.)$ 는 매트릭스 형태로 다시 벡터를 재형상화한다(도 4 참조). 덕셔너리(ψ)는 인코딩된 프레임의 함수가 아니며, 그러므로 매우 작은 오버헤드가 없거나 또는 오버헤드를 가진 디코더측에서 복제될 수 있다는 것을 유념하는 것은 중요하다.

[0075] 5. 인코딩된 프레임(f_2)를 재구성하도록, 모션 매트릭스(Y)와 기준 프레임을 가지는 것이 충분하다는 것이 가정된다. 그러므로, 본 발명자들은 본래의 모션 매트릭스(이전의 단계로부터)의 근사($\hat{Y} \approx Y$)와 재구성하는 기준 프레임(\hat{f}_2 , 인코딩된 프레임(f_2)의 근사)를 이용할 수 있다.

인코더

[0077] 이 섹션에서, 모션 매트릭스의 실시예의 예가 제시된다. 또한, 이러한 모션 매트릭스를 인코딩 및 디코딩하기 위해 요구되는 단계들이 더욱 상세하게 주어진다.

단순 모션 매트릭스

[0079] 본 발명자들이 기준 프레임(f_1)을 이용하는 것에 의해 프레임(f_2)을 인코딩하기를 원한다고 가정한다. 이전에 기술된 바와 같이, 제안된 인코더의 제1스테이지는 모션 매트릭스(Y)를 구하는 것이다. 모션 매트릭스의 간단한 형태는 f_2 (즉, 인코딩될 프레임)의 변형된 버전이어서, $Y = \hat{f}_2$ 로 인용되는 이러한 변형된 버전은 다음과 같은 3개의 핵심 특성을 가진다:

[0080] (a) Y 는 저등급의 것이다.

[0081] (b) Y 는 기준 프레임과 DCT 또는 웨이브릿과 같은 아마도 다수의 다른 변형으로 만들어진 오버 캠플리트(over-complete) 덕셔너리에 관한 회박 표현(α)을 가진다.

[0082] (c) Y 는 f_2 (인코딩될 프레임)의 양호한 근사치를 가진다; 즉:

$$Y = \hat{f}_2 \approx f_2.$$

[0084] 첫번째의 2개의 특성들은 모션 매트릭스의 필요조건이며, 세번째 특성은 모션 매트릭스와 기준 프레임으로부터 인코딩된 프레임을 추론할 필요가 없게되는 것을 기술한다(디코더의 단계 5). 다시 말하면, 모션 매트릭스 자체는 인코딩된 프레임에 매우 근접하고 직접 디스플레이될 수 있다.

[0085] 지금, 본 발명자는 이러한 모션 매트릭스를 찾는데 요구되는 단계를 더욱 상세하게 기술한다. 이러한 모션 매트릭스를 찾는 제1단계는 모션 매트릭스($Y = \hat{f}_2$)가 이에 대하여 회박하게 될 덕셔너리(ψ)를 발생시키는 것이다.

이전에 기술된 바와 같이, 이러한 딕셔너리는 오직 기준 프레임의 함수이게 된다(인코딩될 프레임의 함수가 아니다). 그러므로, 딕셔너리는 디코더측에서 정확하게 복제될 수 있다. 예를 들어, 이러한 딕셔너리는 상이한 움직임(상이한 방향으로의 상이한 시프팅, 회전, 줌 인 등) 또는 기준 프레임 상에서 일부 예측을 컴퓨터로 계산하는 것에 의해 형성될 수 있다. 이러한 딕셔너리를 향상시키고 또한 이러한 딕셔너리가 모든 가능한 이미지를 발생시킬 수 있는 것(즉, 이 딕셔너리는 완전 등급임)을 보장하도록, 일부 변환(예를 들어, DCT, 웨이브릿 등)은 딕셔너리에 부가될 수 있다. 적절하게 디자인되면, 높은 시간적 상관관계는 인코딩될 프레임이 이러한 딕셔너리에 관한 희박 표현을 가지는 것을 보장한다는 것을 유념한다.

[0086] 목표 희박화 딕셔너리(ψ)를 형성한 후에, 모션 매트릭스($Y = \hat{f}_2$)를 찾는 것이 남는다. 이러한 모션 매트릭스는 다음의 방법에 의해 컴퓨터로 계산될 수 있다:

[0087] 입력: 희박화 딕셔너리(ψ), 기준 프레임(f_1), 인코딩될 프레임(f_2), 목표 등급(r) 및 목표 희박도(k)

[0088] 출력: 모션 매트릭스(Y)

[0089] $Y = f_2$ 를 초기화

[0090] · ψ 에 관한 Y 의 표현(a)을 구함: $\text{Vec}(Y) = \psi a$

[0091] · 벡터($\hat{\alpha}$)로 두고, 나머지 지수에서 a 의 가장 큰 엔트리(k) 및 0을 유지하는 것에 의해 컴퓨터로 계산

[0092] · Y 를 $\text{Vec}(Y) = \psi \hat{\alpha}$ 로 업데이트

[0093] · Y 의 가장 큰 고유값(r)을 유지하고 나머지 고유값을 0로 설정

[0094] 수렴 또는 에러가 충분히 작을 때까지 반복

[0095] 근본적으로, 이러한 방법은 2개의 제약 사이에서 반복하고, 즉 (a) 출력 매트릭스(Y)가 등급(r)을 가지며, (b) 출력 매트릭스(Y)는 k (딕셔너리(ψ)에 관한 희박 표현)를 가진다. 다음과 같이 다수의 주석을 강조하는 것은 중요하다:

[0096] · k 및 r 을 조정하는 것은 압축의 양을 제어한다. 특히, r 과 k 가 증가함으로써, 디코더에서 재구성되는 모션 매트릭스의 품질은 압축을 달성하는 비용에서 보다 적게 증가시키고, 또한 그 반대도 같다.

[0097] · 높은 공간적 상관관계는 모션 매트릭스($Y = \hat{f}_2$)의 품질에서 급격한 저하없이 인코딩될 프레임(r)의 등급을 상당히 감소시키는 것을 허용한다.

[0098] · 높은 시간적 상관관계는 딕셔너리(ψ)에서 단지 소수의 아톰(k)과 함께 근사(f_2), 프레임이 인코딩되는 것을 허용한다.

매트릭스 임베딩

[0100] 변형된 프레임($Y = \hat{f}_2$, 또는 동등하게 모션 매트릭스)를 가지고 매트릭스(p_1 및 p_2)를 이용하면, 본 발명자들은 $Z = p_1 \hat{f}_2 p_2$ 내로 f_2 를 임베딩했다. 여기에서, 제1매트릭스(p_1)는 넓은 매트릭스인 반면에, 제2매트릭스(p_2)는 높은 매트릭스이다. 또한, 매트릭스(p_1 및 p_2)는, 임베딩된 매트릭스(Z)를 가지면, 본 발명자들이 $Y = \hat{f}_2$ 를 회수할 수 있도록 선택된다. 예를 들어, 이러한 선택이 최적인 것이 아닐지라도, p_1 및 p_2 는 랜덤 매트릭스일 수 있다. 도 7을 참조하여, 저등급 조밀 매트릭스(Z) 내로 모션 매트릭스를 임베딩하기 위한 예시적인 방법은 대체로 도면부호 (700)으로 지시된다. 상기 방법(700)은 랜덤 및 조밀 매트릭스(p_1 , 705), 모션 매트릭스(저등급)(710), 및 랜덤 및 조밀 매트릭스(p_2 , 715)를 수반하고, 이것들은 다른 저등급이지만 보다 컴팩트한 임베딩된 매트릭스(Z , 720) 내로 저등급 모션 매트릭스를 압축하도록 사용된다. 매트릭스 임베딩의 역할 및 또한 매트릭스(p_1 및 p_2)의 역할은 디코더 설명에 대하여 더욱 상세하게 기술된다. 매트릭스(Z)가 저등급(기껏해야 모션 매트릭스의 등급과 동일)을 가지는 것을 유념한다. 더욱 중요하게, 매트릭스(Z)는 인코딩될 프레임(f_2) 또는 프

레임(f_2)의 변형된/재구성된 버전, 즉 $Y = \hat{f}_2$ 보다 적은 엔트리를 가진다.

[0101] 그러므로, 이러한 것은 제안된 방법에서 압축의 제1레벨이다. 명확하게, 이 단계에서 달성될 수 있는 압축 레벨에는 일부 제한이 있다. 특히, 매트릭스(Z)의 엔트리의 수는 이전의 단계에서의 k 및 r 의 1차함수이다. 이 실시 예에서, k 및 r 의 다른 값이 사용될 수 있거나, 또는 이러한 값이 프레임(f_2) 또는 매트릭스(Z)의 특징에 적응될 수 있다는 것을 유념한다.

[0102] 샘플링

[0103] 압축의 또 다른 레벨은 임베딩된 매트릭스(Z)의 엔트리(S)의 비교적 작은 랜덤 서브세트만을 유지하고 엔트리의 잔여분을 버리는 것에 의해 달성된다. 도 8을 참조하여, 모션 매트릭스와 매트릭스 완료로부터 샘플링하기 위한 예시적인 방법은 대체로 도면부호 (800)으로 지시된다. 상기 방법(800)은 저등급 및 조밀 임베딩된 매트릭스 (805)와 저등급 및 조밀 임베딩된 매트릭스(810)를 수반한다. 디코더와 인코더가 동일 랜덤 시드 및 랜덤 수 발생기를 사용하면, 단지 엔트리값(S)이 보내지고, 디코더와 인코더 모두는 이러한 엔트리의 위치를 알게 되고, 이는 (지수들이 보내지지 않게 되기 때문에) 추가의 압축을 제공한다.

[0104] 양자화 및 엔트로피 코딩

[0105] 이전의 스테이지로부터 임베딩된 매트릭스(Z)의 엔트리(S)들의 랜덤 서브세트는 양자화되고 최종 압축된 심볼(C)을 취하도록 엔트로피 코더를 통과한다.

[0106] 디코더

[0107] 디코더의 다른 구성요소는 이전에 도입된 초기 모션 매트릭스의 경우에 대해 기술되고 설명된다.

[0108] 역양자화 및 엔트로피 디코딩

[0109] 대부분의 전통적인 비디오 인코딩 체계와 유사하게, 디코더의 제1스테이지는 압축된 심볼(C)을 디코딩하고 임베딩된 매트릭스(Z)의 엔트리의 랜덤 서브세트의 근사(\bar{S})를 취하도록 디코딩된 심볼을 역양자화한다.

[0110] 매트릭스 완료

[0111] 최근에, 큰 관심사의 문제, 즉 "저등급 매트릭스 완료"는 일부 조건 하에서 일정 범위 해결되었다. 대체로, 이러한 문제는 다음의 질문을 다룬다: 본 발명자들은 저등급 매트릭스를 갖고, 그 엔트리의 서브세트만을 찾을 수 있다는 것을 가정하여, 본 발명자들이 누락 엔트리(missing entry)를 계산할 수 있을까? 실제로, 이러한 질문에 대한 대답은 매트릭스에 의존한다. 도 9를 참조하여, 대응 고유 벡터가 충분히 랜덤할 때 저등급 매트릭스 완료의 예는 대체로 도면부호 (900)으로 지시된다. 도 10을 참조하여, 충분히 랜덤하지 않은 대응 고유 벡터로 인하여 실패하는 저등급 매트릭스의 예는 대체로 도면부호 (1000)으로 지시된다. 그러므로, 예를 들어, 예(900)에 관하여, 본 발명자들이 밑에 놓인 매트릭스가 저등급이라는 것을 특정하면, 제2행이 제1행의 세배이기 때문에, 누락 엔트리는 6일 것이다. 다른 한편으로, 예(1000)에 관하여, 누락 엔트리는 임의의 수일 것이고 추정되지 않을 수 있다. 일부 조건들이 일치되면, 예를 들어, 저등급 매트릭스의 고유 벡터가 충분히 랜덤하면, 저등급 매트릭스 완료가 용이하게 해결될 수 있으며, SVT(Singular Value Thresholding), FPC(Fixed Point Continuation) 등과 같은 다수의 알고리즘이 이러한 문제를 해결하도록 제안되었다는 것을 보여준다.

[0112] 지금, 임베딩하는 단계 및 매트릭스(p_1 및 p_2)의 역할이 명확하여야 한다. 특히, 이러한 매트릭스들은 임베딩된 매트릭스(Z)의 고유 벡터들이 충분히 랜덤하다는 것을 보장한다. 그러므로, 이러한 매트릭스의 많은 엔트리는 압축을 달성하도록 버려질 수 있으며, 한편 일부 매트릭스 완료 알고리즘은 이러한 버려진 엔트리를 회수할 수 있다. 임베딩된 매트릭스(Z)가 저등급의 것이라는 사실을 활용하면, Z 의 누락 엔트리를 회수하도록, 샘플(\bar{S})의 근사가 입력으로서 입력으로서 매트릭스 완료 알고리즘에 주어진다(도 8 참조). 회수된 매트릭스(\hat{Z})가 임베딩된 매트릭스(Z)와 동일하지 않다는 것을 유념한다. 이러한 것은 양자화 단계($\bar{S} \neq S$)로부터 에러전파(error propagation)에 기인한다. 그러나, 추정 에러($\|Z - \hat{Z}\|_F$, 여기에서, $\|\cdot\|_F$ 는 프로베니우스 매트릭스 놈(Frobenius matrix norm)을 인용)는 양자화 단계에서 도입된 에러에 비례하게 된다.

[0113] 선형 방정식의 열결정계 해결

[0114] 모션 매트릭스($Y = \hat{f}_2$) 또는 인코딩될 프레임의 동등하게 변형된 버전은 기준 프레임 위에 세워진 딕셔너리(ψ)에 관한 (대략) 희박 표현(α)을 가진다는 것, 즉 $\text{Vec}(Y) = \psi \alpha$ 를 상기한다. 또한, 임베딩된 모션 매트릭스($Z = p_1 \hat{f}_2 p_2$)의 벡터화된 형태가 다음의 수학식 1에 의해 다시 표현될 수 있다는 것을 유념한다:

수학식 1

$$\text{Vec}(Z) = \text{Vec}(p_1 Y p_2) = (p_2^T \otimes p_1) \text{Vec}(Y) = (p_2^T \otimes p_1) \psi \alpha$$

[0115] 여기에서, \otimes 는 크로네커 텐서 곱(Kronecker tensor product)이고, $\text{Vec}(\cdot)$ 는 매트릭스의 벡터화된 형태를 표현 한다(도 11 참조). 도 11을 참조하여, 변형된 프레임(f_q')의 희박 표현의 압축 샘플로부터 형성되는 임베딩된

매트릭스(Z)는 대체로 도면부호 (1100)으로 지시된다. 임베딩된 매트릭스(Z)가 $Y = \hat{f}_2$ 에 비교하여 보다 적은 수의 엔트리를 가지기 때문에, 그러므로 수학식 1은 임베딩된 매트릭스(Z)의 엔트리가 방정식(equation)이고 α 의 엔트리가 알려지지 않은 선형 방정식의 열결정계이다. 압축된 션싱(Compressed Sensing, CS)의 최근에 개발된 이론은, 해결책이 희박할 때 선형 방정식의 열결정계에 대한 해결책이 특정 조건 하에서 용이하게 컴퓨터로 계산될 수 있다는 것을 보여주었다. 여기에서, 매트릭스(p_1 및 p_2)의 역할은 이러한 조건을 충족시키는 것을 돋는 것이다. 그러므로, \hat{Z} (Z 의 근사)를 가지면, 희박 벡터(α)는 예를 들어 "기본 추구"의 알고리즘을 사용하는 것에 의해 포괄적인 CS 디코더에 의해 회수될 수 있다. 다른 실시예에서, 희박 벡터를 회수하기 위한 다른 방법이 사용될 수 있는 한편, 본 발명의 사상을 유지한다.

[0116] 결과적으로, 모션 매트릭스의 벡터화된 형태는 다음과 같이 회수될 수 있다:

$$\text{Vec}(Y) = \psi \alpha$$

[0117] 모션 매트릭스의 벡터화된 형태를 재형상화하면, 본 발명자들은 모션 매트릭스를 얻는다. 그러나, 이 예에서; 본 발명자들은 모션 매트릭스가 인코딩된 프레임의 양호한 근사라는 것을 가정하였다. 결과적으로, 모션 매트릭스와 기준 프레임으로부터 인코딩된 프레임을 추론할 필요가 없으며, 본 발명자들은 파생된 모션 매트릭스를 직접 디스플레이할 수 있다.

[0118] 지금 본 발명의 다양한 수행의 장점/특징의 일부의 설명이 주어질 것이며, 장점/특징 중 일부는 위에서 기술되었다. 예를 들어, 하나의 장점/특징은 모션 매트릭스를 사용하여 비디오 시퀀스에서 화상을 인코딩하기 위한 비디오 인코더를 가지는 장치이다. 모션 매트릭스는 주어진 기준점 아래의 등급 및 딕셔너리에 관한 희박 표현을 가진다. 딕셔너리는 아톰의 세트와, 화상을 표현하고 단지 상기 세트만을 사용하여 대응 디코더에서 화상이 파생되도록 허용하기 위한 기본 벡터를 포함한다. 딕셔너리는 비디오 시퀀스에서 기준 화상의 세트로부터 형성된다.

[0119] 또 다른 장점/특징은 상기된 비디오 인코더를 가지는 장치이며, 딕셔너리는 화상과 상기 기준 화상의 세트 사이에서의 공간적 및 시간적 상관관계로부터 형성된다.

[0120] 여전히 또 다른 장점/특징은 상기된 바와 같은 비디오 인코더를 가지는 장치이며, 여기에서, 비디오 인코더는 임베딩된 매트릭스 내로 모션 매트릭스를 임베딩하기 위한 임베딩 디바이스, 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 임베딩된 매트릭스를 샘플링하기 위한 샘플러, 샘플의 양자화된 서브세트를 제공하도록 샘플의 서브세트를 양자화하기 위한 양자화기, 및 압축된 심볼 내로 샘플의 양자화된 서브세트를 엔트로피 코딩하기 위한 엔트로피 코더를 포함한다.

[0121] 여전히 또 다른 장점/특징은 비디오 인코더를 가지는 장치이며, 여기에서, 비디오 인코더는 임베딩된 매트릭스 내로 모션 매트릭스를 임베딩하기 위한 임베딩 디바이스, 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 임베딩된 매트릭스를 샘플링하기 위한 샘플러, 샘플의 양자화된 서브세트를 제공하도록 샘플의 서브세트를 양자화하기 위한 양자화기, 및 상기된 바와 같은 압축된 심볼 내로 샘플의 양자화된 서브세트를 엔트로피 코딩하기 위한 엔트로피 코더를 포함하며, 임베딩된 매트릭스는 모션 매트릭스가 곱해지는 2개의 매트릭스를 사용하여 결정되고, 상기 2개의 매트릭스는 임베딩된 매트릭스에서 기준량(threshold amount)의 랜덤 고유 벡터

(random Eigen-vector)를 제공하도록 선택된다.

[0124] 더욱이, 또 다른 장점/특징은 비디오 인코더를 가지는 장치이며, 여기에서, 비디오 인코더는 임베딩된 매트릭스 내로 모션 매트릭스를 임베딩하기 위한 임베딩 디바이스, 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 임베딩된 매트릭스를 샘플링하기 위한 샘플러, 샘플의 양자화된 서브세트를 제공하도록 샘플의 서브세트를 양자화하기 위한 양자화기, 및 상기된 바와 같은 압축된 심볼 내로 샘플의 양자화된 서브세트를 엔트로피 코딩하기 위해 엔트로피 코더를 포함하며, 샘플의 서브세트는 임베딩된 매트릭스로부터 랜덤하게 선택된다.

[0125] 추가적으로, 또 다른 장점/특징은 비디오 인코더를 가지는 장치이며, 여기에서, 샘플의 서브세트는 상기된 바와 같은 임베딩된 매트릭스로부터 랜덤하게 선택되고, 동일 랜덤 시드와 동일 랜덤 수 발생기는, 임베딩된 매트릭스로부터 샘플의 서브세트에 있는 샘플이 비디오 인코더와 대응 비디오 디코더에서 동일한 각각의 위치를 가지는 것을 보장하도록 상기 비디오 인코더와 상기 대응 비디오 디코더에 의해 사용된다.

[0126] 또한, 또 다른 장점/특징은 비디오 인코더를 가지는 장치이며, 여기에서, 비디오 인코더는 임베딩된 매트릭스 내로 모션 매트릭스를 임베딩하기 위한 임베딩 디바이스, 임베딩된 매트릭스로부터 샘플의 서브세트를 제공하도록 임베딩된 매트릭스를 샘플링하기 위한 샘플러, 샘플의 양자화된 서브세트를 제공하도록 샘플의 서브세트를 양자화하기 위한 양자화기, 및 상기된 바와 같은 압축된 심볼 내로 샘플의 양자화된 서브세트를 엔트로피 코딩하기 위해 엔트로피 코더를 포함하며, 상기 샘플러는 샘플의 서브세트에 있는 샘플과 다른, 임베딩된 매트릭스에서의 위치를 가진 임베딩된 매트릭스의 부분을 버린다.

[0127] 본 발명의 이들 및 다른 특징들 및 장점들은 본 명세서에 있는 교시에 기초하여 당업자에 의해 용이하게 확인될 수 있다. 본 발명의 교시는 하드웨어, 소프트웨어, 펌웨어, 특별 목적 프로세서들, 또는 그 조합들의 다양한 형태들로 구현될 수 있다.

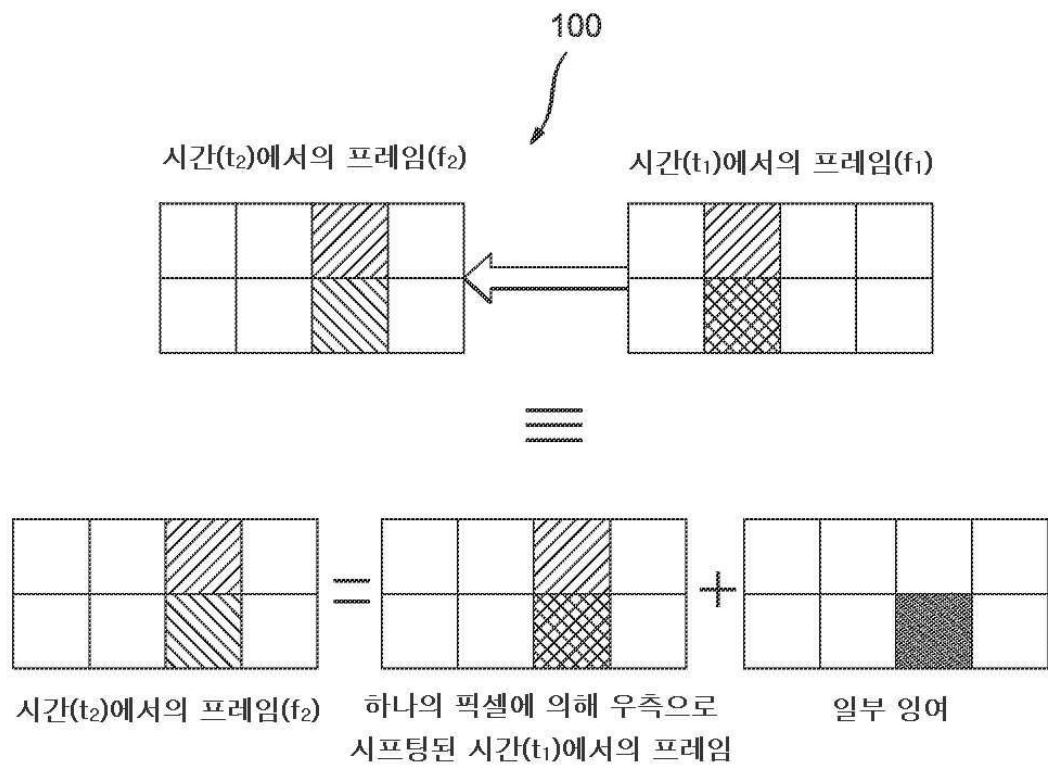
[0128] 가장 바람직하게, 본 발명의 교시는 하드웨어 및 소프트웨어의 조합으로서 실시된다. 더욱이, 소프트웨어는 프로그램 저장 유닛 상에서 현실적으로 구현되는 애플리케이션 프로그램으로서 실시될 수 있다. 애플리케이션 프로그램은 임의의 적합한 아키텍처를 포함하는 머신에 업로드되어 머신에 의해 실행될 수 있다. 바람직하게, 머신은 하나 이상의 중앙 처리 유닛("CPU"), 랜덤 액세스 메모리("RAM"), 및 입/출력("I/O") 인터페이스와 같은 하드웨어를 가지는 컴퓨터 플랫폼 상에서 실시된다. 컴퓨터 플랫폼은 또한 오퍼레이팅 시스템 및 마이크로명령 코드를 포함할 수 있다. 본 명세서에 기술된 다양한 프로세스 및 기능들은 마이크로명령 코드의 일부 또는 애플리케이션 프로그램의 일부, 또는 그 임의의 조합 중 하나일 수 있고, 이는 CPU에 의해 실행될 수 있다. 부가하여, 추가적인 데이터 저장 유닛 및 프린팅 유닛과 같은 다양한 다른 주변장치 유닛들이 컴퓨터 플랫폼에 연결될 수 있다.

[0129] 첨부된 도면들에 도시된 구성 시스템 구성요소들 및 방법들의 일부가 바람직하게 소프트웨어에서 실시되기 때문에, 시스템 구성요소 또는 프로세스 기능 블록 사이의 실제 연결은 본 발명이 프로그래밍되는 방식에 따라 다를 수 있다는 것이 더욱 이해될 것이다. 본 교시들이 주어지면, 당업자는 본 발명의 이들 및 유사한 실행들 또는 구성들을 고려할 수 있을 것이다.

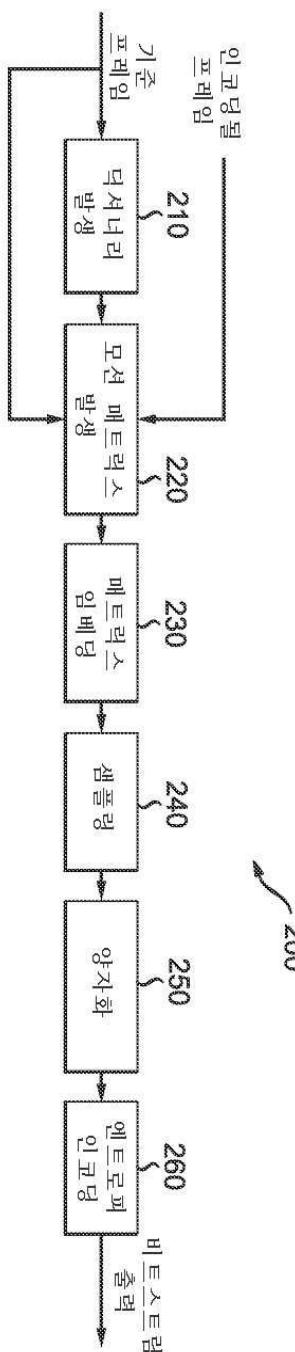
[0130] 여기에서 예시적 실시예들은 첨부된 도면들을 참조하여 설명되었을지라도, 본 발명은 이러한 정확한 실시예로 한정되지 않고, 본 발명의 범위 또는 사상으로부터 벗어남이 없이 당업자에 의해 다양한 변경 및 변형이 달성될 수 있다는 것이 이해될 것이다. 모든 이러한 변경 및 변형은 첨부된 청구항에 제시된 본 발명의 범위 내에 포함되도록 의도된다.

도면

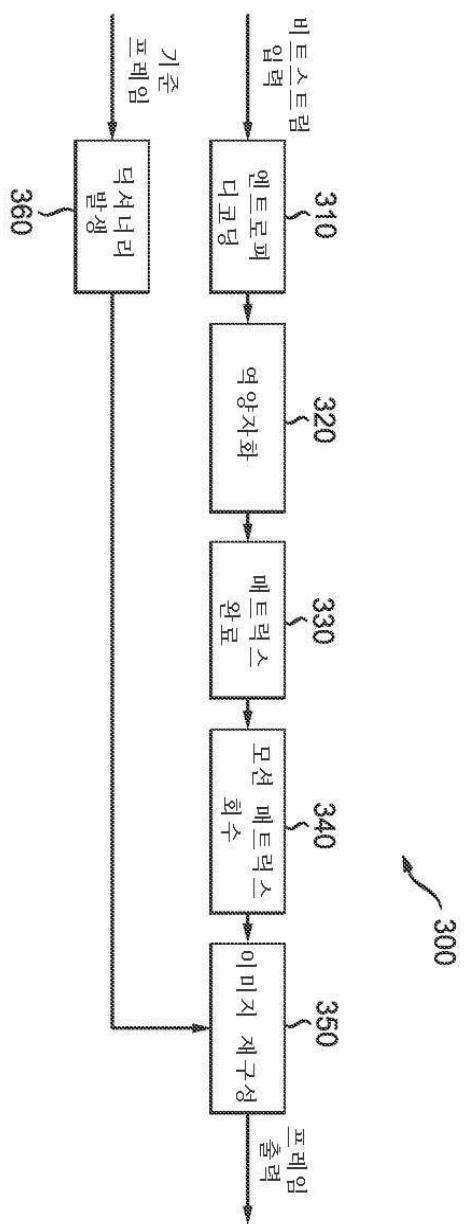
도면1



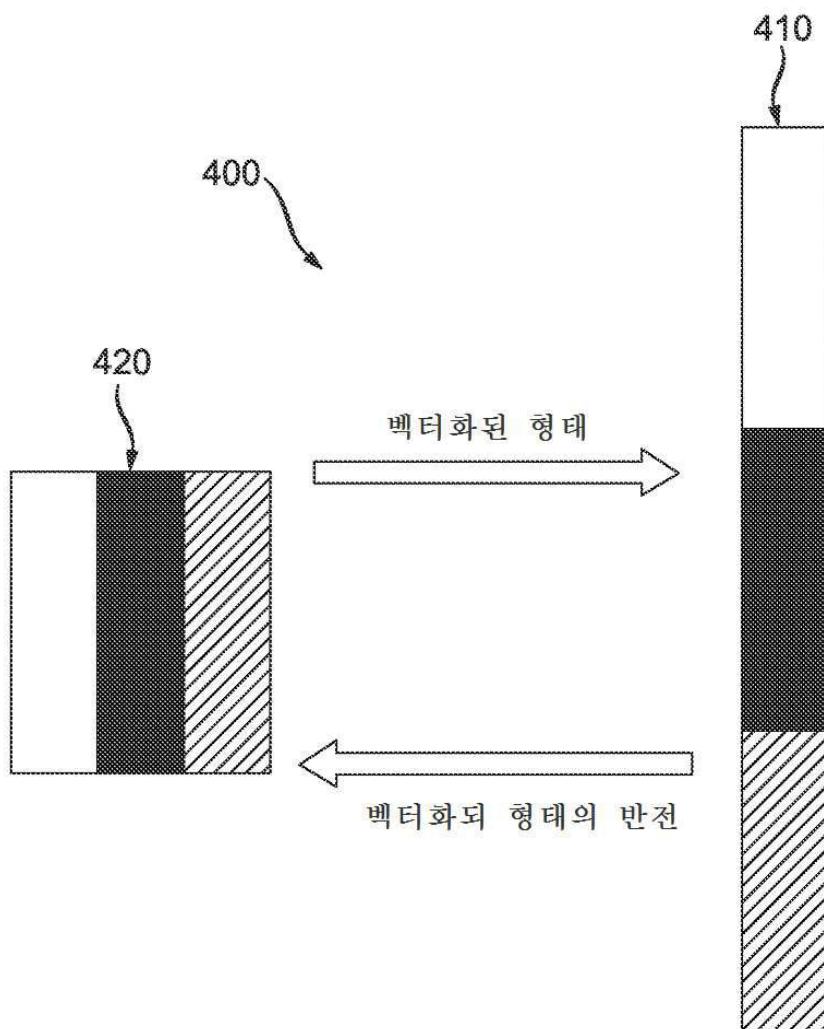
도면2



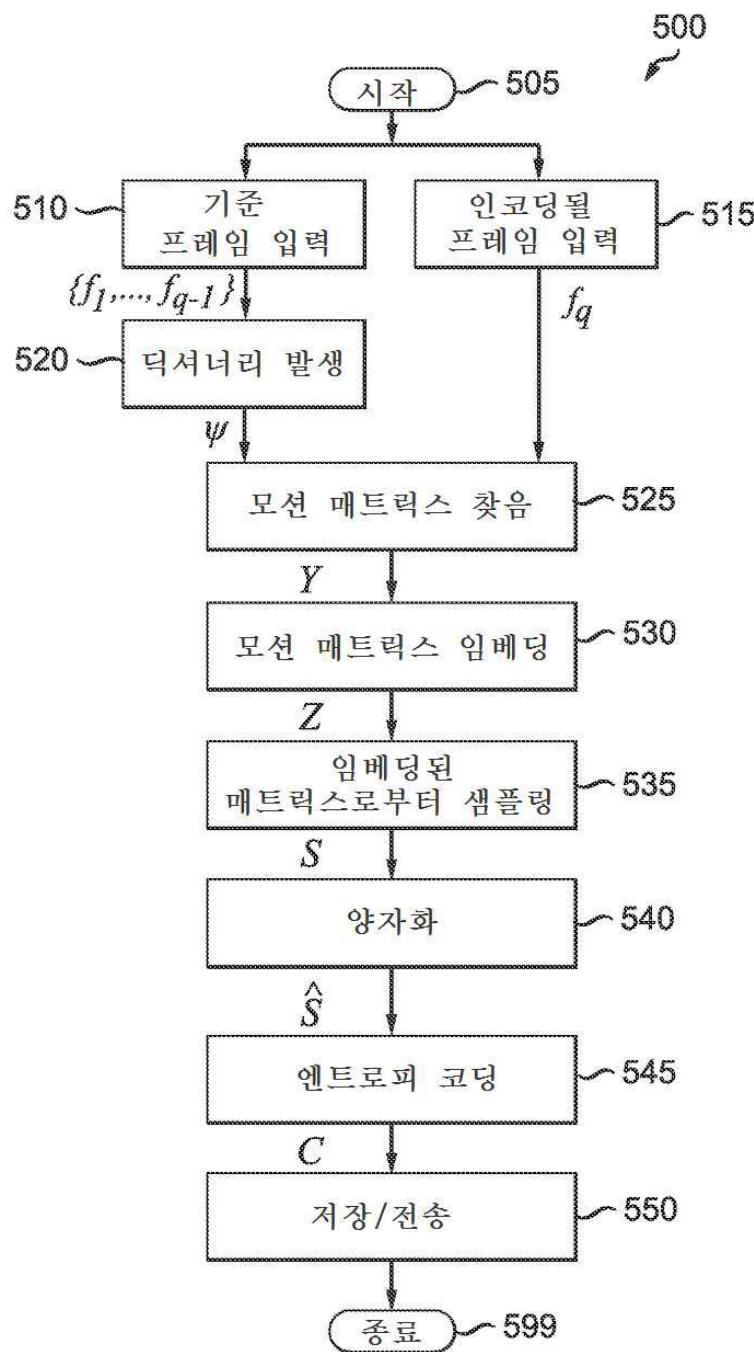
도면3



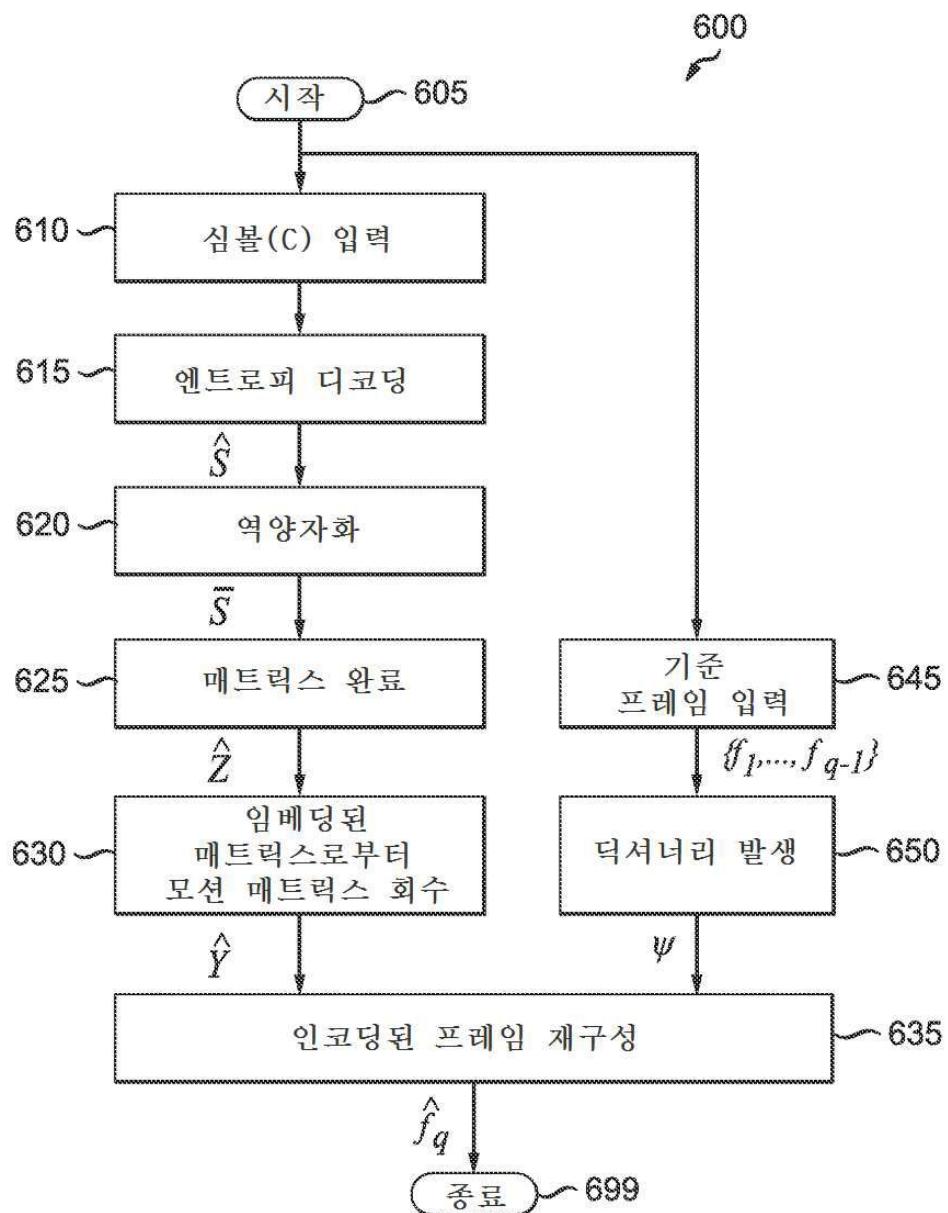
도면4



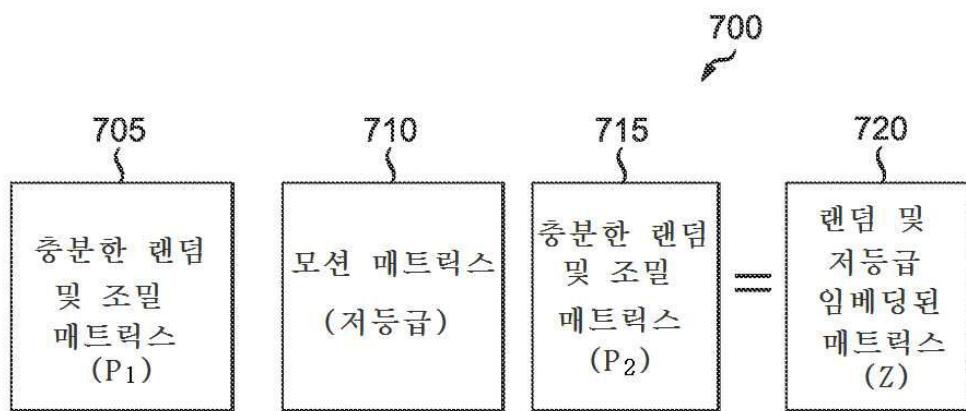
도면5



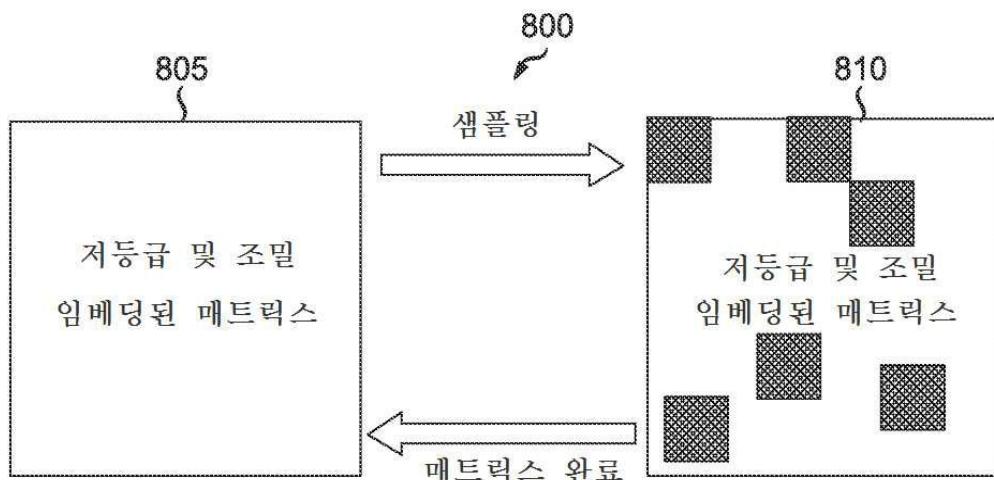
도면6



도면7



도면8



도면9

A 2x2 matrix with the following values:

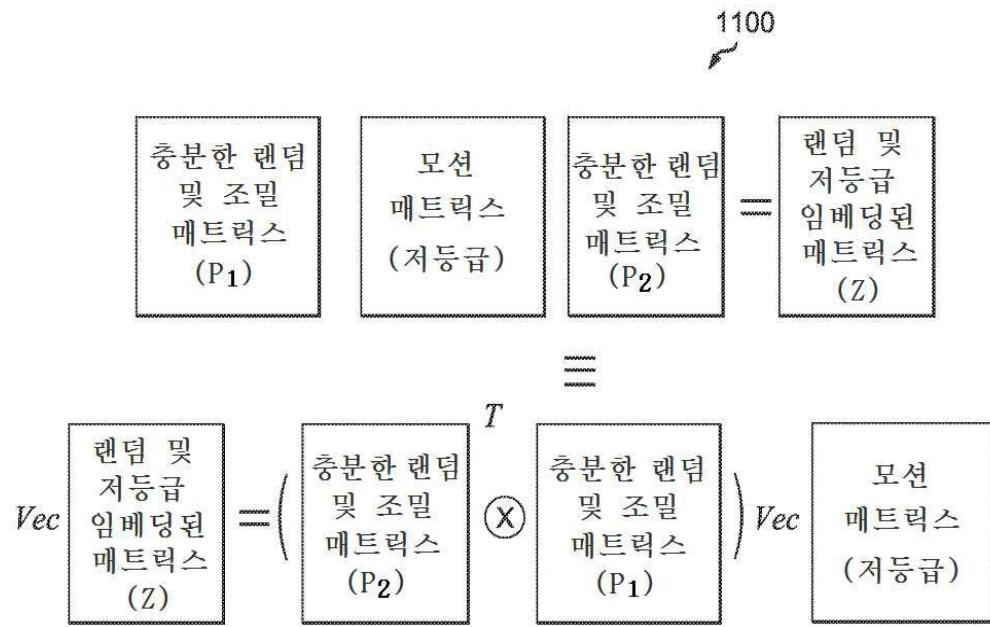
$$\begin{bmatrix} 1 & 3 \\ 2 & ? \end{bmatrix}$$

도면10

A 2x2 matrix with the following values:

$$\begin{bmatrix} ? & 0 \\ 0 & 0 \end{bmatrix}$$

도면11



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 제1, 8, 15, 22, 29항

【변경전】

상기 모션 모션 매트릭스를 개신하는 것

【변경후】

상기 모션 매트릭스를 개신하는 것