

특허청구의 범위

청구항 1

오리지날 비디오 신호(11, 21)를 리프레이밍(16, 23)한 후 상기 리프레이밍된 비디오 신호를 인코딩(18)하는 방법으로서,

상기 오리지날 비디오 신호의 현재 픽처에 대한 돌출 맵(saliency map)을 계산하는 단계(12, 22) - 상기 돌출 맵은 사용자 주의 맵(user attention map)을 나타냄 -;

상기 오리지날 비디오 신호의 상기 현재 픽처에 대한 매크로블록 코딩 효율 코스트 맵을 제공하는 단계(14);

상기 오리지날 비디오 신호의 상기 현재 픽처 내의 후보 리프레이밍 윈도우 크기 및 위치를 선택하는 단계(231);

상기 돌출 맵 또는 상기 결합된 돌출(saliency) 및 매크로블록 코딩 효율 코스트 맵을 사용함으로써 상기 오리지날 비디오 신호에서 상기 리프레이밍 윈도우의 상이한 위치 및/또는 크기를 판정하는 단계(16, 17) - 상기 상이한 위치 및/또는 크기의 경우 상기 리프레이밍된 비디오 신호에 대한 코딩 코스트는 상기 후보 리프레이밍 윈도우 크기 및 위치에 대한 코딩 코스트보다 작음 -; 및

상기 대응하는 리프레이밍된 비디오 신호를 인코딩하는 단계(18)

를 포함하는 방법.

청구항 2

오리지날 비디오 신호(11, 21)를 리프레이밍(16, 23)한 후 상기 리프레이밍된 비디오 신호를 인코딩(18)하는 장치로서,

상기 오리지날 비디오 신호의 현재 픽처에 대한 돌출 맵을 계산하고 상기 현재 픽처에 대한 매크로블록 코딩 효율 코스트 맵을 제공하도록 구성된 수단(12, 14, 15, 22) - 상기 돌출 맵은 사용자 주의 맵을 나타냄 -;

상기 오리지날 비디오 신호의 상기 현재 픽처 내의 후보 리프레이밍 윈도우 크기 및 위치를 선택하고, 상기 돌출 맵 또는 상기 결합된 돌출 및 매크로블록 코딩 효율 코스트 맵을 사용함으로써 상기 오리지날 비디오 신호에서 상기 리프레이밍 윈도우의 상이한 위치 및/또는 크기를 판정하도록 구성된 수단(231, 16, 17) - 상기 상이한 위치 및/또는 크기의 경우 상기 리프레이밍된 비디오 신호에 대한 코딩 코스트는 상기 후보 리프레이밍 윈도우 크기 및 위치에 대한 코딩 코스트보다 작음 -; 및

상기 대응하는 리프레이밍된 비디오 신호를 위한 인코딩 수단(18)

을 포함하는 장치.

청구항 3

제1항에 따른 방법 또는 제2항에 따른 장치에 있어서,

시간에 따른 상기 돌출 맵의 콘텐츠의 변이(variation)로 인한 상기 리프레이밍 윈도우의 위치 및 크기의 변이를 제어하기 위해 칼만 필터(Kalman filter)가 사용되는데(232, 233), 방법 또는 장치.

청구항 4

제3항에 따른 방법 또는 제3항에 따른 장치에 있어서,

상기 리프레이밍 윈도우 파라미터들을 안정화하기 위해 이전 상태 x_{k-1} 및 상태 전이 모델에 대해 현재 상태 x_k 를 예측하고, 잡음 측정을 사용하여 상기 예측을 보정하기 위해 각각의 픽처에 대한 2개의 단계에서 칼만 필터가 사용되고, 상태 x_k 는 상기 리프레이밍 윈도우의 중심의 위치 및 그 크기를 정의하고,

$$x_k = \begin{bmatrix} centerColumn \\ centerLine \\ windowWidth \\ windowHeight \end{bmatrix}$$

인, 방법 또는 장치.

청구항 5

제1항, 제3항 또는 제4항 중 어느 한 항의 방법 또는 제2항 내지 제4항 중 어느 한 항의 장치에 있어서,

상기 매크로블록 코딩 효율 코스트 맵(14)이 제공되고, 상기 돌출 맵의 상기 현재 컨텐츠로부터 기인한 상기 후보 리프레이밍 윈도우의 종횡비(aspect ratio)가 상기 오리지널 비디오 신호(11, 21)의 종횡비와 상이한 경우, 상기 오리지널 비디오 신호의 상기 종횡비를 획득하기 위해 상기 후보 리프레이밍 윈도우는 수평 또는 수직 방향으로 대응적으로 확장되며, 상기 확장의 방향은 상기 현재 후보 리프레이밍 윈도우의 경우 상기 현재 매크로블록 코딩 효율 코스트 맵으로부터 도출된 매크로블록 코딩 코스트가 감소되도록 수행되는, 방법 또는 장치.

청구항 6

제1항 내지 제5항 중 어느 한 항의 방법 또는 제2항 내지 제5항 중 어느 한 항의 장치에 있어서,

상기 돌출 맵(13)은 상기 매크로블록 코스트 맵(14)과 통합되어(15), 적용된 최종 돌출 맵에서 고비용 코딩 코스트 매크로블록이 대응하는 돌출 값을 감소시키는, 방법 또는 장치.

청구항 7

제1항 내지 제6항 중 어느 한 항의 방법에 있어서,

상기 인코딩하는 단계(18)는 H.264/AVC 인코딩하는 것이고, 제2항 내지 제6항 중 어느 한 항의 장치에 있어서, 상기 인코딩 수단(18)은 H.264/AVC 인코딩 수단인, 방법 또는 장치.

청구항 8

제7항의 방법 또는 제7항의 장치에 있어서,

상기 리프레이밍은 상기 인코딩 루프 내에 포함되지만 전체 픽처가 인코딩되며, 사용될 리프레이밍 윈도우를 상기 디코더에 시그널링하는데 SEI 메시지들이 사용되는, 방법 또는 장치.

청구항 9

제1항 내지 제8항 중 어느 한 항의 방법 또는 제2항 내지 제8항 중 어느 한 항의 장치에 있어서,

상기 리프레이밍 윈도우는 매크로블록 경계에 따라 정렬되는, 방법 또는 장치.

청구항 10

인코딩 전에 제1항 내지 제9항 중 어느 한 항의 방법에 따라 리프레이밍되는 인코딩된 디지털 비디오 신호.

명세서

기술분야

[0001] 본 발명은 오리지널 비디오 신호를 리프레이밍 및 인코딩하는 방법 및 장치에 관한 것으로서, 여기서 리프레이밍 윈도우 위치(position) 및/또는 크기는 리프레이밍된 비디오 신호의 인코딩 코스트(encoding cost)를 줄이기 위해 적응되어 있다.

배경기술

[0002] 리프레이밍(reframing)은 예를 들어, 주어진 종횡비(aspect ratio)를 갖는 비디오 신호를 상이한 종횡비를 갖는

디스플레이에 디스플레이하기 위해 이미지 또는 비디오 콘텐츠의 크기를 재조정하는데 사용된다. 예를 들어, 고해상도(HD:High Definition) 비디오 콘텐츠는 소형 휴대용 디바이스에 디스플레이하는데 적합하지 않을 수 있다.

- [0003] EP 1748385 A2호는 인간의 시각 주의 모델(visual attention model)에 기반한 동적 리프레이밍을 개시하는데, 여기서 소스 비디오 콘텐츠는 관심 영역을 유지하기 위해 적절히 크롭핑(cropping)된다. 출력 신호는 인코딩되어 네트워크를 통해 송신될 수 있다.
- [0004] C. Chamaret, O. LeMeur가 발표한 논문, "Attention-based video reframing: validation using eye-tracking", 19th International Conference on Pattern Recognition ICPR'08, 8-11 Dec.2008, Tampa, FL, USA 또한 리프레이밍 적용예를 설명한다.
- [0005] O. LeMeur, P. LeCallet 및 D. Barba가 발표한 논문, "Predicting visual fixations on video based on low-level visual features", Vision Research, vol.47, no.19, pp. 2483-2498, Sept.2007은 시각 주의 모델에 기반하여 동적 돌출 맵(saliency map)의 계산을 설명한다.

발명의 내용

- [0006] 공지되지 않은 리프레이밍 프로세싱은 출력된 크롭핑 및 인코딩된 비디오 신호의 비트 레이트 및 왜곡을 다룰 것이다. 예를 들어, 크롭핑 윈도우(cropping window)는 인코딩된 비디오 신호의 코딩 복잡성을 고려하지 않고 관심 영역을 추적할 수 있다. 이는 다수의 주밍(zooming) 및 디주밍(de-zooming), 패닝(panning)을 초래할 수 있으며 이로 인해 영역들을 나타내기 위한 높은 코딩 코스트를 초래할 수 있다. 예를 들어, 최종 리프레이밍된 비디오 신호가 H.264/AVC 인코더를 사용함으로써 인코딩되는 경우, 이는 비트 레이트의 증가 및/또는 비디오 품질의 감소를 초래할 수 있다.
- [0007] 본 발명에 의해 해결하고자 하는 문제점은 비디오 소스 신호 리프레이밍을 제공하는 것으로서, 여기서 크롭핑 윈도우 위치 및 크기는 그 크롭핑된 윈도우를 포함하는 인코딩된 출력 비디오 신호의 레이트/왜곡을 고려한다. 이러한 문제점은 청구항 제1항에 개시된 방법에 의해 해결된다. 이 방법을 활용한 장치가 청구항 제2항에 개시된다.
- [0008] 본 발명에 따르면, 인코딩된 출력 비디오 신호의 레이트/왜곡을 최적화하기 위해 크롭핑 윈도우 파라미터들(시간에 따른 위치 및 크기)이 제약된다. 비디오 코딩 콘텍스트를 고려하고, 리프레이밍된 비디오 시퀀스가 인코딩되면 유발되는 코딩 효율 코스트(coding efficiency)를 고려함으로써 초기 리프레이밍은 개선된다.
- [0009] 원칙적으로, 본 발명의 방법은 오리지날 비디오 신호를 리프레이밍한 후 리프레이밍된 비디오 신호를 인코딩하는데 적합하며, 상기 방법은,
- [0010] 상기 오리지날 비디오 신호의 현재 픽처에 대한 돌출 맵을 계산하는 단계 - 상기 돌출 맵은 사용자 주의 맵을 나타냄 -;
- [0011] 상기 오리지날 비디오 신호의 상기 현재 픽처에 대한 매크로블록 코딩 효율 코스트 맵을 제공하고, 상기 돌출 맵과 상기 매크로블록 코딩 효율 코스트 맵을 결합하는 단계;
- [0012] 상기 오리지날 비디오 신호의 상기 현재 픽처 내의 후보 리프레이밍 윈도우 크기 및 위치를 선택하는 단계;
- [0013] 상기 돌출 맵 또는 상기 결합된 돌출(saliency) 및 매크로블록 코딩 효율 코스트 맵을 사용함으로써 상기 오리지날 비디오 신호에서 리프레이밍 윈도우의 상이한 위치 및/또는 크기를 판정하는 단계 - 상기 상이한 위치 및/또는 크기의 경우 상기 리프레이밍된 비디오 신호에 대한 코딩 코스트는 상기 후보 리프레이밍 윈도우 크기 및 위치에 대한 코딩 코스트보다 더 작음 -; 및
- [0014] 상기 대응하는 리프레이밍된 비디오 신호를 인코딩하는 단계를 포함한다.
- [0015] 원칙적으로, 본 발명의 장치는 오리지날 비디오 신호를 리프레이밍한 후 리프레이밍된 비디오 신호를 인코딩하는데 적합하며, 상기 장치는,
- [0016] 상기 오리지날 비디오 신호의 현재 픽처에 대한 돌출 맵을 계산하고 상기 현재 픽처에 대한 매크로블록 코딩 효율 코스트 맵을 제공하고, 상기 돌출 맵과 상기 매크로블록 코딩 효율 코스트 맵을 결합하도록 구성된 수단 - 상기 돌출 맵은 사용자 주의 맵을 나타냄 -;
- [0017] 상기 오리지날 비디오 신호의 상기 현재 픽처 내의 후보 리프레이밍 윈도우 크기 및 위치를 선택하고, 상기 돌

출 맵 또는 상기 결합된 돌출 및 매크로블록 코딩 효율 코스트 맵을 사용함으로써 상기 오리지널 비디오 신호에서 리프레이밍 윈도우의 상이한 위치 및/또는 크기를 판정하도록 구성된 수단 - 상기 상이한 위치 및/또는 크기의 경우 상기 리프레이밍된 비디오 신호에 대한 코딩 코스트는 상기 후보 리프레이밍 윈도우 크기 및 위치에 대한 코딩 코스트보다 더 작음 -; 및

[0018] 상기 대응하는 리프레이밍된 비디오 신호를 위한 인코딩 수단을 포함한다.

[0019] 본 발명의 추가적인 유익한 실시예들은 개별 종속항들에 개시되어 있다.

도면의 간단한 설명

[0020] 본 발명의 예시적인 실시예들은 다음의 첨부 도면을 참조하여 설명된다.

도 1은 본 발명의 프로세싱의 전체 블록도이다.

도 2는 본 발명에 따른 리프레이밍 프로세싱을 도시한다.

도 3은 크롭핑 윈도우의 이방성 확장(anisotropic extension)을 도시한다.

도 4는 매크로블록 코스트 맵에 기반한 이방성 확장을 도시한다.

도 5는 이들의 코딩 코스트 C_i 에 관한 여러 크롭핑 윈도우들 사이의 선택을 도시한다.

도 6은 비디오 코딩에 최적화된 크롭핑 윈도우 변위(displacement)를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0021] 도 1에서, 소스 비디오 신호가 돌출 계산 단계 또는 스테이지(12) 및 리프레이밍 단계/스테이지(16)에 공급된다. 돌출 계산 단계 또는 스테이지(12)는 현재 소스 비디오 신호 콘텐츠의 돌출 맵(13)을 계산하며, 이로써 돌출 맵은 사용자 주의 맵을 나타낼 수 있고, 예를 들어 픽셀마다 그 지각 관련성(perceptual relevancy)을 특징으로 하는 돌출 값을 제공한다. 입력은 현재 픽처 콘텐츠이지만, 출력은 그 픽처를 바라보거나, 비디오 시퀀스 내의 그 픽처를 바라보는 경우 픽처의 어느 섹션이 사람들에게 의해 어떠한 주의를 얻고 있는지를 보여주는 맵이다. 블록 또는 매크로블록 코딩 효율 코스트 맵이 소스 비디오 신호의 인코딩 버전으로부터 그 입력 신호를 수신하는 대응하는 단계/스테이지(14)에서 계산된다. 단계/스테이지(14)가 존재하면, 그 출력 신호는 돌출 맵(13)과 결합된다(15). 그 결과는 또한 다음에 후술되는 대로 동작하는 리프레이밍 단계/스테이지(16)에 입력되어, 코딩 코스트가 감소된 리프레이밍된 비디오 신호(17)를 출력한다. 리프레이밍된 비디오 신호(17)는 (예를 들어, H.264 프로세싱을 사용함으로써) 인코더(18)에서 인코딩되어, 그 결과 최적화된 레이트/왜곡 특성을 갖는 인코딩 및 리프레이밍된 비디오 신호가 생성된다.

[0022] 출력 비디오 신호 인코딩의 레이트/왜곡을 최적화하기 위해, 크롭핑 윈도우는 다음 몇몇 규칙을 따르고 있다.

[0023] - 윈도우 경계에 영역 또는 물체들을 나타내는 것을 피하는데, 그 이유는 이것이 인트라-코딩 모드 블록들을 유발하고 인터-코딩 모드 또는 스킵 모드보다 더 많은 인코딩용 비트들을 요구하기 때문이다. 적절한 경우, 전자 이미지 안정화기(stabiliser)가 사용되어야 한다.

[0024] - 물체의 크기 변경을 피한다. 그 목적은 줌 팩터가 비디오 모션 파라미터들을 따라야 한다는 양호한 인터 예측을 선호하기 위한 것이다.

[0025] - 일정한 주요 모션(dominant motion)을 선호한다. 즉, 인코딩은 콘텐츠 모션이 일정한 경우 더 효율적일 수 있다(양방향 예측, 스킵, 및 직접 모드의 양호한 사용).

[0026] - 인코딩하기 어려운 매크로블록들을 포함하려 하지 않는다. 압축된 비디오를 리프레이밍하는 경우, 코스트가 많이 드는 이미지 영역들이 관심 영역이 아니면 이러한 코스트가 많이 드는 이미지 영역들을 거절하기 위한 매크로-블록 코딩 효율 코스트 맵을 사용한다.

[0027] 도 2는 리프레이밍 프로세싱을 더 상세하게 도시한다. 원 소스 시퀀스(21)는 시각 주의 모델(22)을 사용하여 확인되며, 이로써 현재 픽처(도 1의 12 및 13에 대응함)에 대한 돌출 맵을 초래한다. 다음의 크롭핑 윈도우 추출(23)(도 1의 16에 대응함)은 리프레이밍된 시퀀스(24)를 출력하고, 후술되는 바와 같은, 초기 윈도우 추출(231), 임시 일관성 검사(temporary consistency check)(232), 및 중형비 프로세싱(233)을 포함한다.

[0028] 다음으로, 인코딩된 출력 리프레이밍된 비디오 시퀀스의 비트 레이트 감소를 다루는 여러 보충적인 실시예들을

설명한다.

[0029] A. 시간적 일관성(temporal consistency) 개선

[0030] 이 섹션에서는 시간적으로(temporally) 더 안정한 이미지가 원칙적으로 더 적은 인코딩용 비트들을 필요로 할 것이라는 사실이 사용된다.

[0031] A.1 간단한 시간적 제약

[0032] 크롭핑 윈도우의 위치 및 크기를 제어하기 위해 칼만 필터(Kalman filter)가 사용된다. 제1 구현예에서, 공분산 잡음 행렬(co-variance noise matrix) Q 가 제약된다. 칼만 필터 적용의 목적은 주의 모델(attention model)(22)에 의해 제공되는 원래 값(raw value)의 변이들(variations), 즉 시간에 따른 돌출 맵의 콘텐츠의 변이들을 평활화하는 것이다. 칼만 모델화에서, 주의 모델(22)에 의해 주어지는 원래 값들은 최적의 크롭핑 파라미터들(윈도우의 중심 및 크기)을 추정하는 것이 시도되는 동안 잡음 추정을 나타내는 것으로 간주된다. 크롭핑 윈도우의 파라미터들의 현재 상태 x_k 는 $x_k = A_k x_{k-1} + B_k u_k + w_k$ 로 정의되는데, 여기서, A_k 는 이전 상태에 적용된 상태 전이 모델(state transition model)이고, B_k 는 제어 입력 모델이고, u_k 는 제어 벡터이고, w_k 은 $w_k \approx N(0, Q_k)$ 인 상태 잡음이고, N 은 평균이 0이고 공분산이 Q_k 인 정규 분포이다.

[0033] 시간 k에서, 실제 상태 x_k 의 관측 또는 측정 z_k 이 수행되어, $z_k = H_k x_k + v_k$ 로 정의되는데, 여기서 H_k 는 실제 상태 공간을 관측된 공간에 맵핑하는 관측 모델이고, v_k 는 공분산 R_k 을 갖는 평균이 0인 가우시안 화이트 잡음(zero mean Gaussian white noise)이라고 가정되는 $v_k \approx N(0, R_k)$ 인 관측 잡음이다.

[0034] 리프레이밍 적용예에서, 칼만 필터는 다음과 같이 사용된다. 상태 x_k 는 크롭핑 윈도우 및 그 크기의 중심의 위치를 정의한다. 이는 다음과 같이 정의된다.

$$x_k = \begin{bmatrix} centerColumn \\ centerLine \\ windowWidth \\ windowHeight \end{bmatrix}$$

[0035]

[0036] 각각의 이미지에 대해, 칼만 필터가 2개의 단계, 즉 예측 페이즈(prediction phase) 및 업데이트 페이즈(update phase)로 적용된다.

[0037] 예측 페이즈에서, 예측 단계는 이전 상태 및 상태 전이 모델을 이용하여 현재 상태를 추정한다. 이러한 특별한 케이스에서, 칼만 필터는 크롭핑 윈도우 파라미터들을 안정화하기 위해 사용된다. 결과적으로, 예측 상태 \hat{x}_k 는 대개 이전 상태의 사본이다.

[0038] 즉, $\hat{x}_k = x_{k-1}$ 인데, A_k 는 항등 행렬(identity matrix)이고 $B_k = 0$ 이면, $\hat{P}_k = P_{k-1} + Q_{k-1}$ 이며, 여기서 P_k 는 시스템 잡음을 제어하는 공분산 행렬이다.

[0039] 업데이트 페이즈는 잡음 추정을 사용한 예측의 보정이다.

[0040] 즉, $\tilde{y}_k = z_k - \hat{x}_k$ 는 측정 잔차(measurement residual)이며, 여기서 z_k 는 주의 모델(attention model)에 의해 주어진 원래 윈도우 파라미터들이고, H_k 는 항등 행렬이다.

$$x_k = \hat{x}_k + K_k \tilde{y}_k, \quad P_k = (I - K_k) \hat{P}_{k-1}$$

[0041]

$$K_k = \frac{\hat{P}_k}{\hat{P}_k + R_k}$$

여기서, $\hat{P}_k + R_k$ 는 사후 오차 공분산(posteriori error covariance)을 최소화하는 행렬 이득(matrix gain)이다.

초기 추정치로서, 스크린의 중심이 선택될 수 있다. R 및 Q 는 각각 상태를 정의하고 공분산 잡음을 측정하는 상수 대각 행렬들(constant diagonal matrices)이다. 이러한 컨텍스트에서, 시간적 제약을 구현하기 위한 가장 간단한 방식은 상태 공분산 잡음 행렬(state covariance noise matrix) Q 의 값들을 낮추는 것이다.

이러한 해결 방안은 움직이는 배경을 갖는 비디오 콘텐츠(또는 0이 아닌 주요 모션)를 고려하지 않는다. 다음 섹션은 이러한 더 복잡한 케이스를 고려한다.

A.2 주요 모션 제약

0이 아닌 주요 모션, 즉 배경 변환 및/또는 줌을 갖는 비디오 콘텐츠의 경우, 코딩 효율을 개선하기 위해, 크롭핑 윈도우는 이미지의 경계에 블록들을 나타내는 것을 피하고, 더 우수한 인터-픽처 예측을 위한 물체들의 크기 변화를 피하기 위해 주요 모션을 따른다.

시각 주의 모델에서, 6 파라미터 어파인(affine) 모델을 이용한 주요 모션의 파라미터들을 계산하는 모션 추정기가 사용된다. 이전에 정의된 컨텍스트에서, 이들 파라미터는 백그라운드 변환 및 줌을 도출하는데 사용된다.

전술한 바와 같이, 칼만 필터 상태는 다음과 같이 정의된다.

$$x_k = \begin{bmatrix} centerColumn \\ centerLine \\ windowWidth \\ windowHeight \end{bmatrix}$$

이때, 커맨드는 널이 아니고(non-null), 주요 모션을 따르는데 사용된다.

$$u_k = \begin{bmatrix} translation_x \\ translation_y \\ zoom_dx \\ zoom_dy \end{bmatrix} \quad B_k = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix}$$

다시, 각각의 이미지에 대해, 칼만 필터가 2개의 단계, 즉 예측 페이즈 및 업데이트 페이즈로 적용된다.

예측 페이즈에서, 예측 단계는 이전 상태 및 상태 전이 모델을 이용하여 현재 상태를 추정한다. 크롭핑 윈도우 파라미터들은 단지 주요 모션에 따라서만 변하게 된다.

즉, $\hat{x}_k = x_{k-1} + B_k u_k$, 여기서 A_k 는 항등 행렬이며, $\hat{P}_k = P_{k-1} + Q_{k-1}$ 이다.

업데이트 페이즈는 잡음 측정을 사용한 예측의 보정이다.

즉, $\tilde{y}_k = z_k - \hat{x}_k$ 는 측정 잔차이며, 여기서 z_k 는 주의 모델에 의해 주어진 원래 윈도우 파라미터들이고, H_k 는 항등 행렬이다.

$$K_k = \frac{\hat{P}_k}{\hat{P}_k + R_k}$$

[0057] 는 사후 오차 공분산을 최소화하는 행렬 이득이다.

$$x_k = \hat{x}_k + K_k \tilde{y}_k, \quad P_k = (I - K_k) \hat{P}_{k-1}$$

[0058]

[0059] 초기 추정치로서, 스크린의 중심이 선택될 수 있다. Q 및 R 은 상태를 정의하고 공분산 잡음을 측정하는 상수 대각 행렬들이다. 여기서, 상태 공분산 잡음 행렬 Q 는 다시 모델과 실제 출력 사이의 관계를 정의한다. 또한, 주요 모션 파라미터들의 추정에 잡음을 고려한다. Q 가 낮은 값들을 가지면, 출력은 이러한 모델로 강하게 제약되지만, 그렇지 않으면(Q 가 높은 값들을 가지면), 잡음으로 인해 주의 모델 출력 및 가능하게는 높은 변이들을 더 빠르게 따를 것이다.

[0060] 주요 모션의 출력 u_k 은 시간적 일관성을 개선하기 위하여 칼만 필터에 통합될 수 있다.

[0061] B. 매크로블록 코스트에 의한 제약

[0062] 원 비디오 시퀀스(11 및 21)는 압축된 비디오 시퀀스인 경우 인코딩 코스트의 값들과 함께 들어온다고 가정한다. 공지된 바와 같이, 소정의 매크로블록의 코딩 코스트는 현재 양자화 파라미터 q 를 사용하여 그 매크로블록을 인코딩하는데 요구되는 비트들의 개수로 표현된다. 본 발명에 따르면, 이들 입력 시퀀스 코딩 코스트는 리프레이밍을 제약하는데 사용된다. 다음에 정의된 프로세싱은 섹션 A에 설명된 프로세싱에 추가적 또는 독립적으로 사용될 수 있다.

[0063] B.1 전체 이미지 코스트에 의한 제약

[0064] 칼만 필터의 상태 공분산 잡음 행렬 Q 는 픽처 시퀀스의 전체 코스트로부터 도출될 수 있다. 입력 시퀀스의 코스트가 낮으면, 크롭핑된 픽처의 코스트가 낮을 것이라고 예측될 수 있고, 그 결과 크롭핑된 시퀀스의 코스트를 감소시키기 위한 제약들이 줄어들 수 있다.

[0065] 일례로서, 행렬 Q 는 $Q = I \cdot (\sigma - \lambda \cdot \text{cost})$ 로서 정의될 수 있는데, 여기서 I 는 항등 행렬이고, σ 는 상수이고, λ 는 cost 에 다소의 가중치를 부여하는 가중 파라미터이고, cost 는 초당 메가바이트(MB/s)의 시퀀스의 코딩 코스트이다.

[0066] B.2 매크로블록 코스트 맵을 이용한 윈도우 확대 제약

[0067] 이 구현은 도 2에 설명된 중형비 단계/스테이지(233)를 다룬다. 일단 제1 크롭핑 윈도우가 돌출 맵으로부터 단계/스테이지(231)에서 외삽되면(extrapolated), 그 로케이션 및 크기가 시간적으로(temporally) 평활화된다. 그 후, 이방성 확장(anisotropic extension)이 수행되고, 그의 상이한 변수 및 기호들이 도 3에 도시되어 있다.

[0068] 중형비 AR은 오리지날 비디오 신호(11 및 22)의 폭과 높이 사이의 비이다. 이방성 확장은 현재의 중형비 R_{SM} 에 따른 방향으로 크롭핑 윈도우 $CW_i^{SM}(x_{SM}, y_{SM}, w_{SM}, h_{SM})$ 를 확장함으로써 크롭핑 윈도우 크기를 개량하는데, 여기서 SM 은 돌출 맵을 지칭한다. 확장은 타겟 중형비 R_{TG} 를 이루기 위한 폭 또는 높이에 대한 것이다.

$$R_{SM} = \frac{w_{SM}}{h_{SM}}$$

[0069] 는 돌출 맵으로부터의 추출로부터 유발되는 중형비이고, 이다.

$$R_{TG} = \frac{w_{TG}}{h_{TG}} = \frac{w_{AR}}{h_{AR}}$$

는 타겟 중형비

[0070] $R_{TG} > R_{SM}$ 이면, (현재의 직사각형의 폭에 대한) 수평 확장이 수행되고, 그렇지 않으면 (현재의 직사각형의 높이에 대한) 수직 확장이 수행된다.

[0071] (괄호 안의 수직 확장에 대한) 수평 확장을 가정하면, 다음을 정의할 수 있다.

$$\begin{aligned} h_{AR} &= h_{SM} \quad (w_{AR} = w_{SM}) \\ w_{AR} &= w_{SM} + d_w \quad (h_{AR} = h_{SM} + d_h) \\ x_{AR} &= x_{SM} + d_x \quad (x_{AR} = x_{SM}) \\ y_{AR} &= y_{SM} \quad (y_{AR} = y_{SM} + d_y) \end{aligned}$$

[0072]

[0073] 여기서,

$$d_w = d_{left} + d_{right} = \frac{w_{AR}}{h_{AR}} \times h_{SM} - w_{SM} = R_{TG} \times h_{SM} - w_{SM} \quad d_x = x_{SM} + \frac{(d_{right} - d_{left})}{2}$$

및 $w_{AR} = w_{SM} + d_{left} + d_{right}$ 이다.

[0074] 일단 확장의 측(side)이 정의되면, 윈도우를 확장하는 여러 방식이 여전히 존재한다. 즉, d_{right} 및 d_{left} 이 상이한 방식으로 계산될 수 있다. 다음으로, 폭 w_{SM} 이 최종 종횡비에 도달하도록 확장될 수 있다고 가정한다. 확장은 전부 $d_{left} = d_w$ 및 $d_{right} = 0$ 와 같이 좌측으로 이동될 수 있고, $d_{left} = 0$ 및

$$d_{right} = d_w \text{와 같이 우측으로 이동될 수 있고, 또는 } d_{left} = \frac{d_w}{2} \text{ 및 } d_{right} = \frac{d_w}{2} \text{와 같이 동일한 비율로 양측으로 이동될 수 있다. 이러한 방안들은 콘텐츠 관점에서 최적이지 않다. 따라서, 종래 기술에서 하나의 측 또는 나머지 측을 선호하도록 돌출 맵의 더 미세한 분석이 실시되었다.}$$

[0075] 본 발명에 따르면, 새로운 기준, 즉, 도 4에 설명된 현재의 픽처에 대한 매크로블록 코딩 효율 코스트 맵에 기반한 코딩 코스트가 확장 방향의 선택에 사용된다. 일 방향의 영역이 나머지 방향에 비해 코딩될 비트가 더 소요되는 경우(매크로블록 코스트 맵에 의해 추론됨), 나머지 방향이 선호된다. 즉, 크롭핑 윈도우는 각각의 방향에서의 비트 코스트들에 비례하여 반대 방향으로 확장되며, 이로써 이러한 코스트는 최소화된다. 좌측 및/또는 우측으로의 확장들은 다음과 같이 계산된다.

[0076] [수학식 1]

$$d_{right} \text{ 및 } d_{left} \text{ 은 } \frac{d_{right}}{d_{left}} = \frac{C_{right}^{\max}}{C_{left}^{\max}} \text{와 같이 구해져야 한다.}$$

[0077]

[0078] 비트 코스트 C_{right}^{\max} 및 C_{left}^{\max} 는 좌측으로의 완전 확장($d_{left} = w_{AR} - w_{SM}$ 및 $d_{right} = 0$) 및 우측으로의 완전 확장($d_{right} = w_{AR} - w_{SM}$ 및 $d_{left} = 0$)을 고려함으로써 계산되는데, 여기서

$$C_{left}^{\max} = \sum_{y=y_{SM}-\frac{h_{SM}}{2}}^{y=y_{SM}+\frac{h_{SM}}{2}} \sum_{x=x_{inf}}^{x=x_{inf}+d_w} C(x, y) \quad \text{및} \quad C_{right}^{\max} = \sum_{y=y_{SM}-\frac{h_{SM}}{2}}^{y=y_{SM}+\frac{h_{SM}}{2}} \sum_{x=x_{sup}-d_w}^{x=x_{sup}} C(x, y)$$

[0079] 이다.

$$x_{inf} = x_{SM} - \frac{w_{SM}}{2} - (w_{AR} - w_{SM}) \quad \text{및} \quad x_{sup} = x_{SM} + \frac{w_{SM}}{2} - (w_{AR} - w_{SM})$$

[0080] 가 정의될 수 있다.

일단 각각의 측에 이용 가능한 돌출 양이 알려지면, 수학적 (1)을 사용하는 경우 각각의 방향의 확장 d_{right} 및 d_{left} 를 다음과 같이 추정할 수 있다.

$$d_{left} = \frac{d_w \times C_{left}^{max}}{C_{left}^{max} + C_{right}^{max}}, \quad d_{right} = \frac{d_w \times C_{right}^{max}}{C_{left}^{max} + C_{right}^{max}}$$

B.3 매크로블록 코스트 맵에 의해 선택된 미리 정의된 윈도우 로케이션

크롭핑 윈도우 박스의 로케이션을 제약하는 다른 방법은 도 5에 도시된 바와 같은 인접한 여러 도달 가능 크롭핑 윈도우에 대응하는 코스트만을 계산하여 가장 작은 코딩 코스트를 갖는 크롭핑 윈도우를 유지하는 것이다.

M개의 후보 크롭핑 윈도우 세트는 대응하는 코스트 $\{C_0 \dots C_M\}$ 를 갖는다. 크롭핑 윈도우의 최종 로케이션의 코딩 코스트 및 이로 인한 최종 로케이션 자체는 $C_{final} = \text{Min}\{C_0 \dots C_M\}$ 에 의해 판정된다.

C. 그 밖의 특징들

C.1 돌출 맵 레벨에서의 제약

다른 관심 실시예는 돌출 맵을 매크로블록 코스트 맵과 통합함으로써 고비용 코딩 코스트 매크로블록이 최종 돌출 맵에서의 그 대응하는 돌출을 감소시키는 것이다. 이로써 고비용 매크로블록의 잠재적 영향은 최종 크롭핑 윈도우 로케이션의 판정 시에 감소된다.

2개의 맵을 통합하는 하나의 방식은 CNSP(Coherent Normalisation, Sum plus Product)라고 지칭되는 다음의 프로세싱을 적용하는 것이다.

$$SM_{final} = N_c(SM) + N_c(MB_{inv}) + (1 + N_c(SM)) \cdot (1 + N_c(MB_{inv}))$$

여기서, MB 는 범위가 0 내지 255인 매크로블록 코딩 효율 코스트 맵 값이고, $MB_{inv} = 255 - MB$ 이고, SM 은 돌출 맵 값이고, N_c 는 선형적(a priori) 지식에 의해 구동되는 정규화 오퍼레이터이다. 이 오퍼레이터는 각각의 맵의 전역적 최대값(global maximum)을 사용하는 것 대신 실증 값(empirical value)을 사용한다.

C.2 인코더 아키텍처에 의한 크롭핑 윈도우 변위 제약

크롭핑 윈도우를 인코더 아키텍처에 적응시킬 뿐 아니라 다음과 같은 일부 케이스에서 코딩 효율을 개선하도록 약간의 간략화가 수행될 수 있다.

- 예측 프레임과 현재 프레임 사이의 매크로블록 그리드(macroblock grid)에 대한 크롭핑 윈도우의 정렬을 보장한다. 이는 시간적 예측 품질을 개선할 것이고, 또한 적절한 블록 크기를 선택하는 것을 도울 수 있다. 단점은 크롭핑 윈도우의 평활화된 시각적 변이가 어쨌든 보증되어야 하고, 그렇지 않으면 일부 시각적 아티팩트(artefact)들이 보일 것이라는 점이다.

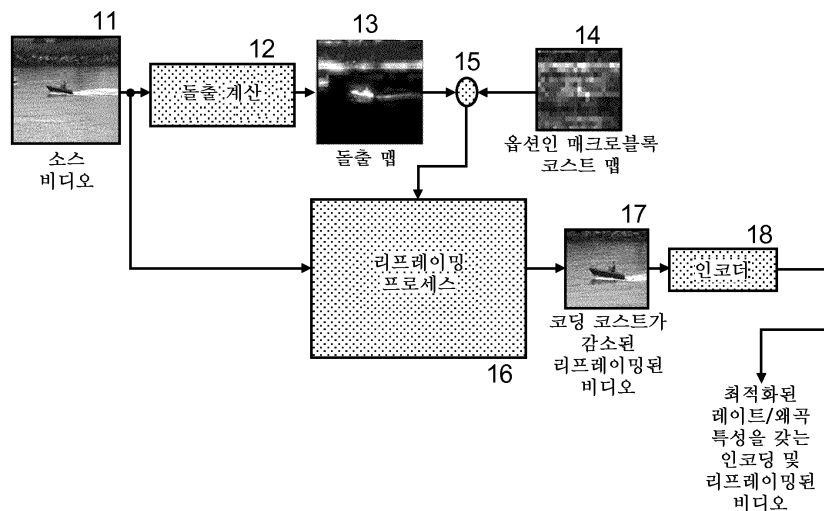
- 이와 유사하게, 크롭핑 윈도우 위치가 픽셀 샘플링 그리드에 매치되도록 설정될 수 있다. 그 경우에, 모션 벡터는 필수적으로 픽셀 개수의 배수(the motion vector will be necessarily a multiple of a number of pixels)일 것이며, 서브-픽셀(sub-pel) 모션 추정이 필요 없을 수 있다. 유리하게, 이는 인코더의 상당한 간략화를 초래한다.

- 리프레임 프로세스로부터의 임의의 정보를 재사용한다. 예를 들어, 돌출 영역들을 계산하기 위해, 프레임 사이의 모션을 계산하는 것이 필수적이다. 이러한 정보는 비디오 인코더로 전달되어 인코더 모션 추정을 돕거나 대체할 수 있다. 유사하게, 크롭핑 윈도우의 시간적 변이들은 카메라 맨의 촬영 방식과 유사한 평활한 거동을 보장하도록 모니터링되고 관리된다. 이들 가상 카메라 이동은 인코더가 모션 추정을 돕는데 사용될 수도 있다.

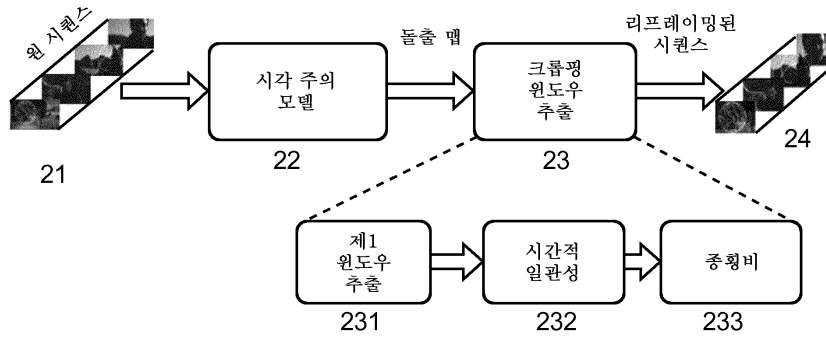
- [0095] - 소스 입력 비디오 신호가 인코딩된 신호인 경우(인코딩 포맷과 무관함), 비디오 신호 트랜스코딩 또는 변환으로부터 알 수 있듯이 파라미터와 같은 코딩 정보를 재사용하는 것이 가능할 것이다.
- [0096] 인코딩을 위한 리프레이밍 프로세스의 이러한 적응에 추가하여, 인코딩 루프 내에 리프레이밍 프로세스를 포함하는 것이 가능할 수도 있다. 예를 들어, 인코딩 및 리프레이밍이 사전 프로세싱으로서 수행되는 것 대신 함께 최적화되도록 프레임의 인코딩 중에 크롭핑 윈도우가 계산될 수 있다. 이들은 다음과 같은 여러 이점을 갖는다.
- [0097] - 첫째, 전체 픽처가 인코딩될 수 있고, 일부 인코딩 전략이 크롭핑 윈도우에 적응될 수 있는데, 즉 H.264/AVC SEI 메시지들이 옵션의 크롭핑 윈도우를 디코더로 시그널링하는데 사용되고, 더 높은 비트 레이트/인코딩 품질이 크롭핑 윈도우를 위해 설정될 수 있고, 상이한 슬라이스들이 2개의 영역을 위해 사용될 수 있고, 스킵 모드들이 비-크롭핑 윈도우를 위해 설정될 수 있다.
- [0098] - 둘째, 전체 프레임이 인코더에 사용될 수 있지만, 크롭핑 윈도우만이 인코딩된다. 이는 인코딩뿐 아니라 이전에 설명된 크롭핑 윈도우 크기 및 위치의 합동 최적화를 가능하게 하고, 모션 추정은 전체 프레임에 대해 수행되어 경계 효과를 방지할 수 있고 모션 필드의 인코딩을 용이하게 하고(우수한 균질성(homogeneity)), 크롭핑 윈도우로부터 전체 프레임으로 동적 스위칭할 가능성을 제공한다.
- [0099] - 셋째, 전술한 바와 같이, 하나의 프로세스로부터의 여러 계산들은 이들을 2번 수행하는 것 대신 후속 프로세스에 의해 재사용될 수 있다. 아키텍처 및 CPU 실행(또는 하드웨어 복잡성)은 전반적으로 감소될 것이고, 이러한 시스템은 더 낮은 기능성 및 유연성을 제공한다.

도면

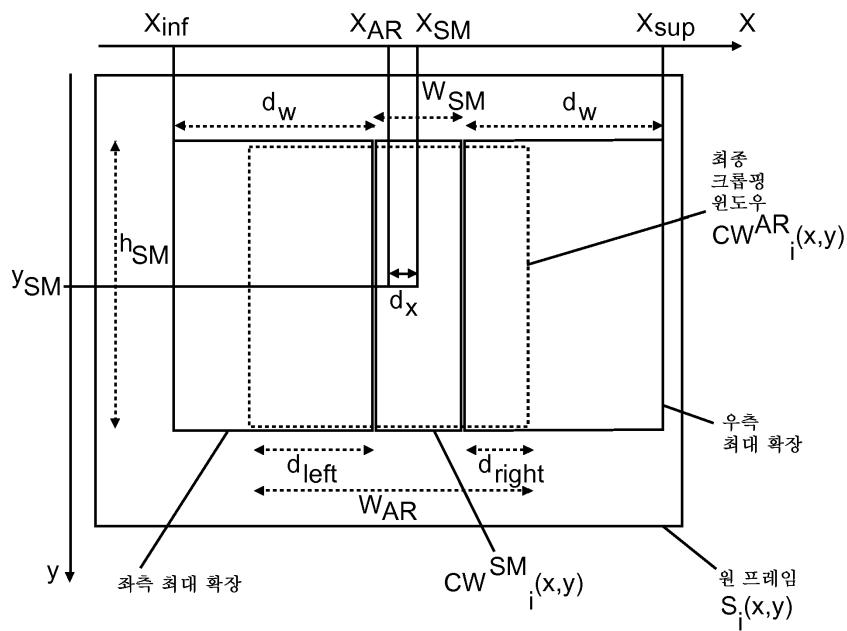
도면1



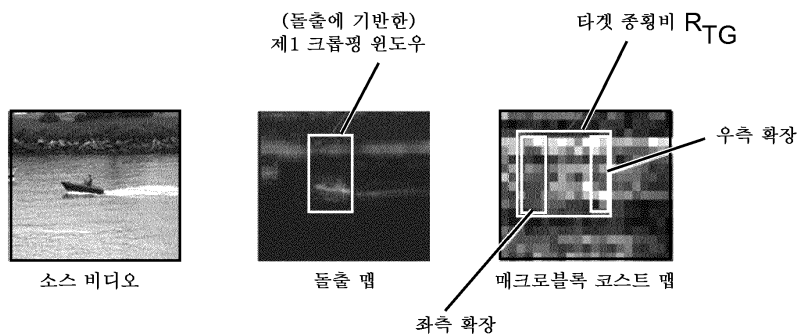
도면2



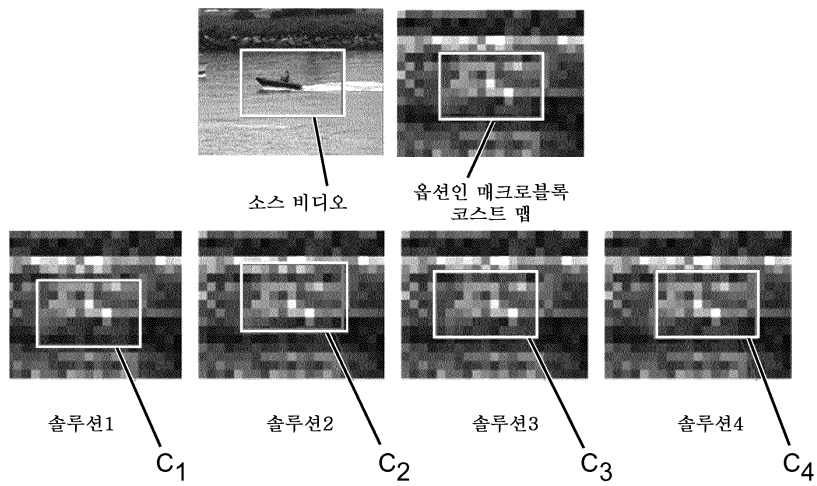
도면3



도면4



도면5



도면6

