



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0039698
 (43) 공개일자 2013년04월22일

(51) 국제특허분류(Int. Cl.)
 H04N 7/32 (2006.01)

(21) 출원번호 10-2012-0113086

(22) 출원일자 2012년10월11일

심사청구일자 없음

(30) 우선권주장

11306322.6 2011년10월12일
 유럽특허청(EPO)(EP)

(71) 출원인

툼슨 라이센싱

프랑스 92130 이씨레물리노 잔 다르크 튀 1-5

(72) 발명자

위르방, 파브리쓰

프랑스 35510 쉐송-셰비네 아브뉴 드 벨르 폰텐느
 1 페끄니 끌로르 에르 에 데 프랑스

샤마레, 크리스텔

프랑스 35510 쉐송-셰비네 아브뉴 드 벨르 폰텐느
 1 페끄니 끌로르 에르 에 데 프랑스

슈방스, 크리스토프

프랑스 35510 쉐송-셰비네 아브뉴 드 벨르 폰텐느
 1 페끄니 끌로르 에르 에 데 프랑스

(74) 대리인

백만기, 양영준, 전경석

전체 청구항 수 : 총 12 항

(54) 발명의 명칭 데이터 스트림에서 블록-단위로 예측 인코딩된 비디오 프레임의 블록의 돌출 값을 결정하기 위한 방법 및 디바이스

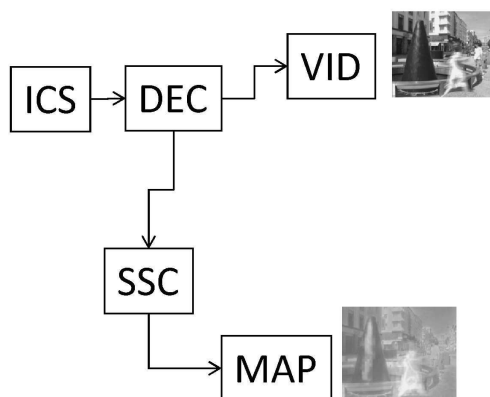
(57) 요약

본 발명은 데이터 스트림에서 블록-단위로(block-wise) 예측 인코딩된 비디오들에 대한 돌출(saliency) 결정의 분야에서 이루어진다.

본 방법은, 처리 수단을 사용하여, 블록들의 변환된 레시듀얼들의 코딩 비용을 결정하는 단계 및 결정된 코딩 비용을 사용하여 돌출 맵을 결정하는 단계를 포함하도록 제안된다.

블록들의 변환된 레시듀얼들의 코딩 비용은, 블록들이 얼마나 잘 예측되는지는 물론 블록들에 묘사된 내용의 선명함에 의존하므로, 돌출에 대한 좋은 지표이다.

대표도 - 도2



특허청구의 범위

청구항 1

데이터 스트림에서 블록-단위로(block-wise) 예측 인코딩된 비디오 프레임의 블록의 돌출 값(saliency value)을 결정하기 위한 방법으로서,

처리 수단을 사용하여,

상기 블록의 변환된 레시듀얼(transformed residual)의 코딩 비용을 결정하는 단계; 및

상기 결정된 코딩 비용을 사용하여 상기 돌출 값을 결정하는 단계

를 포함하는 돌출 값 결정 방법.

청구항 2

데이터 스트림에서 블록-단위로 예측 인코딩된 비디오 프레임의 블록의 돌출 값을 결정하기 위한 디바이스로서,

상기 블록의 변환된 레시듀얼의 코딩 비용을 결정하고, 상기 결정된 코딩 비용을 사용하여 상기 돌출 값을 결정하도록 구성된 처리 수단

을 포함하는 돌출 값 결정 디바이스.

청구항 3

제1항 또는 제2항에 있어서, 상기 블록은 인트라 예측(intra-predictive) 인코딩되고, 상기 코딩 비용의 결정은 로-도메인 모델(rho-domain model)을 사용하여 결정하는 것을 포함하는 돌출 값 결정 방법 또는 디바이스.

청구항 4

제1항 또는 제2항에 있어서, 상기 블록은 인터 예측(inter-predictive) 인코딩되고, 상기 코딩 비용의 결정은 상기 블록의 인터 예측에 사용되는 참조 블록의 변환된 레시듀얼의 코딩 비용을 결정하는 것을 포함하는 돌출 값 결정 방법 또는 디바이스.

청구항 5

제4항에 있어서, 상기 처리 수단을 사용하여 상기 참조 블록의 상기 결정된 코딩 비용을 상기 블록의 크기로 가중하는 것을 더 포함하는 돌출 값 결정 방법 또는 디바이스.

청구항 6

제3항 또는 제4항에 있어서, 상기 블록의 움직임 벡터의 코딩 비용을 사용하여 상기 돌출 값을 결정하는 것을 더 포함하는 돌출 값 결정 방법 또는 디바이스.

청구항 7

제1항, 제3항 내지 제6항 중 어느 한 항에 있어서, 상기 처리 수단을 사용하여 상기 결정된 코딩 비용을 정규화하고, 상기 정규화된 코딩 비용을 사용하여 상기 돌출 값을 결정하는 것을 더 포함하는 돌출 값 결정 방법 또는 디바이스.

청구항 8

제4항에 있어서, 상기 처리 수단은 상기 참조 블록의 상기 결정된 코딩 비용을 상기 블록의 크기로 가중하도록 또한 구성되는 돌출 값 결정 방법 또는 디바이스.

청구항 9

제3항 또는 제4항에 있어서, 상기 처리 수단은 상기 블록의 움직임 벡터의 코딩 비용을 사용하여 상기 돌출 값을 결정하도록 구성되는 돌출 값 결정 방법 또는 디바이스.

청구항 10

제2항 내지 제5항, 제8항 및 제9항 중 어느 한 항에 있어서, 상기 처리 수단은 상기 결정된 코딩 비용을 정규화하고, 상기 정규화된 코딩 비용을 사용하여 상기 돌출 값을 결정하도록 구성되는 돌출 값 결정 방법 또는 디바이스.

청구항 11

제4항 내지 제6항 중 어느 한 항에 있어서, 상기 처리 수단을 사용하여 상기 블록이 다이렉트/스킵(Direct/Skip) 모드에서 인코딩되는지 여부를 결정하고, 상기 블록이 다이렉트/스킵 모드에서 인코딩되는 경우에 감쇠 값(attenuation value)이 상기 돌출 값을 결정하는데 또한 사용되는 돌출 값 결정 방법 또는 디바이스.

청구항 12

제4항, 제5항 또는 제8항 중 어느 한 항에 있어서, 상기 처리 수단은 상기 블록이 다이렉트/스킵 모드에서 인코딩되는지 여부를 결정하도록 구성되고, 상기 블록이 다이렉트/스킵 모드에서 인코딩되는 경우에 감쇠 값이 상기 돌출 값을 결정하는데 또한 사용되는 돌출 값 결정 방법 또는 디바이스.

명세서

기술분야

[0001] 본 발명은 비디오들에 대한 돌출 결정(saliency determination)의 분야에서 실시된다.

배경기술

[0002] 비디오 이미지 프레임에서, 돌출 특징들(salient features)로도 불리는, 현저성(remarkability)의 특징들 또는 관심이 증가된 위치들을 검출하는 것은 많은 실-세계(real-world) 애플리케이션들을 갖는다. 예를 들어, 이것은, 탐색 지원(navigational assistance), 로봇 제어, 감시(surveillance) 시스템, 물체 탐지와 인식, 및 장면 이해와 같은 컴퓨터 비전 작업들에 적용될 수 있다. 이러한 예측들은 또한 광고 디자인, 이미지와 비디오 압축, 이미지와 비디오 리퍼포징(repurposing), 그림 데이터베이스 문의(pictorial database querying), 및 응시 애니메이션(gaze animation)을 포함하는 다른 영역에서 애플리케이션들을 찾는다.

[0003] 일부 선행 기술에 따른 시각적 주목 산출 모델들(visual attention computational models)은, 색상, 강도, 콘트라스트, 방향들, 움직임 및 입력 이미지 또는 비디오 신호의 다른 통계적 분석과 같은 소스 데이터의 하위-수준(low-level) 특징들로부터 돌출 맵(saliency map)을 산출한다.

[0004] 예를 들어, 신경 정보 처리 시스템들의 발전들(Advances in neural information processing systems)(2006)의 155-162 페이지에서: Bruce, NDB, 및 Tsotsos, JK의: "정보 극대화에 기초한 돌출(Saliency based on information maximization)"은, 장면으로부터 샘플링된 정보를 극대화하는 상향식 명시적 주목(bottom-up overt attention)의 모델을 제안한다.

[0005] TPAMI(IEEE Trans Pattern Anal Mach Intell.)(1998)의 20(11):1254-9에서: Itti L., Koch C., 및 Niebur E.의: "빠른 장면 분석을 위한 돌출-기반 시각적 주목의 모델(Model of saliency-based visual attention for rapid scene analysis)"은, 초기 영장류 시각 시스템의 동작 및 뉴런 아키텍처에 의해 영감을 얻은, 시각적 주목 시스템(visual attention system)을 제시한다. 이 시스템은, 상세하게 분석될 눈에 잘 띄는 위치들을, 계산적으로 효율적인 방식으로, 빠르게 선택함으로써, 장면 이해의 복잡한 문제를 타파할 수 있다.

[0006] 인지 계산(Cognitive Computation)(2011), 3권, 1번, 37-47에서: Fabrice U. 등의: "중간 공간 주파수들, 돌출의 강한 예측자(Medium Spatial Frequencies, a Strong Predictor of Saliency)"는, 중간 주파수들이 전세계적으로 주목(attention)의 최고 예측을 허용한다는 것을 발견했고, 고정 위치들에 대해서는, 인공 거리 장면들(man-made street scenes)에서 중간 내지 높은(medium to high) 주파수들을 사용하고, 자연 경관 장면들(natural landscape scenes)에서 낮은 내지 중간(low to medium) 주파수들을 사용하여 더 잘 예측할 수 있다는 것을 발견했다.

발명의 내용

해결하려는 과제

[0007] 자료는 보통 - 공간적 변환들, 공간적 및 시간적 예측들, 및 움직임 정보에 기초하여 - 주목할만한 특징들 및 관심이 증가된 위치에서의 정보를 보존하는 방식으로 압축되지만, 압축-인코딩된 비디오 자료에 대한 종래 기술의 돌출 결정 방법들 및 디바이스들은 자료의 디코딩을 요구하고, 이에 따라 디코딩시에 손실하는 일부 돌출 정보를 이미 포함한다는 것을 본 발명자들이 알게 되었다.

과제의 해결 수단

[0008] 따라서, 본 발명자들은 낮은 계산 비용의 돌출 모델(low-computational cost saliency model)을 창출하기 위해 압축된 비디오로부터 돌출 정보를 추출할 것을 제안한다. 계산 비용 절감은 인코딩으로 인해 사용가능한 데이터의 재사용을 기반으로 한다.

[0009] 즉, 본 발명자들은, 데이터 스트림에서 블록-단위로(block-wise) 예측 인코딩된 비디오 프레임의 블록의 돌출 값을 결정하기 위한, 청구항 1에 따른 방법 및 청구항 2에 따른 디바이스를 제안한다. 상기 방법은, 처리 수단을 사용하여, 블록의 변환된 레시듀얼(transformed residual)의 코딩 비용을 결정하는 단계, 및 결정된 코딩 비용을 사용하여 돌출 값을 결정하는 단계를 포함한다.

[0010] 변환된 블록 레시듀얼의 코딩 비용은, 블록이 얼마나 잘 예측되는지는 물론 블록에 묘사된 내용의 선명함(vividness)에 의존한다. 코딩 비용은 따라서 돌출에 대한 좋은 지표(good indication)이다.

[0011] 일 실시예에서, 블록은 인트라-예측(intra-predictive) 인코딩되고, 코딩 비용을 결정하는 단계는 로-도메인 모델(rho-domain model)을 사용하여 결정하는 단계를 포함한다.

[0012] 다른 실시예에서, 블록은 인터-예측(inter-predictive) 인코딩되고, 코딩 비용을 결정하는 단계는 상기 블록의 인터-예측을 위해 사용된 참조 블록(reference block)의 변환된 레시듀얼의 코딩 비용을 결정하는 단계를 포함한다.

[0013] 또 다른 실시예에서, 참조 블록의 코딩 비용을 결정하는 것은 블록의 크기에 의해 가중된다.

[0014] 또 다른 실시예에서, 블록의 움직임 벡터의 코딩 비용은 돌출 값을 결정하기 위해 아직 더 사용된다.

[0015] 또 다른 실시예에서, 결정된 코딩 비용은 정규화되고(normalized), 정규화된 코딩 비용은 돌출 값을 결정하기 위해 사용된다.

[0016] 블록이 다이렉트/스킵 모드(Direct/Skip mode)에서 인코딩되면, 돌출 값을 결정하기 위해 감쇠 값(attenuation value)이 추가로 사용될 수 있다.

[0017] 추가의 유리한 실시예들의 특징들은 종속 청구항들에 명시되어 있다.

도면의 간단한 설명

[0018] 본 발명의 예시적인 실시예들은 도면들에 도시되어 있고, 다음 설명에 보다 상세히 설명된다. 예시적인 실시예들은 단지 본 발명을 명확히 하기 위해 설명된 것으로, 청구항들에서 정의된 본 발명의 개시 또는 범위를 제한하기 위한 것이 아니다.

도 1은 종래 기술에 따른 돌출 맵(saliency map)의 도출에 대한 예시적인 순서도를 도시한다.

도 2는, 스트림으로부터 공간적 돌출 맵을 도출함으로써, 압축된 비디오 스트림으로부터 돌출 맵을 도출하는 것에 대한 제1 실시예의 예시적인 순서도를 도시한다.

도 3은, 스트림으로부터 시간적 돌출 맵을 도출함으로써, 압축된 비디오 스트림으로부터 돌출 맵을 도출하는 것에 대한 제2 실시예의 예시적인 순서도를 도시한다.

도 4는, 스트림으로부터 공간적 돌출 맵과 시간적 돌출 맵과 도출된 맵들의 융합(fusion)을 도출함으로써, 압축된 비디오 스트림으로부터 돌출 맵을 도출하는 것에 대한 제3 실시예의 예시적인 순서도를 도시한다.

도 5는 압축된 비디오 스트림으로부터 공간적 돌출 맵을 도출하는 것에 대한 예시적인 순서도를 도시한다.

도 6은 압축된 비디오 스트림으로부터 시간적 돌출 맵을 도출하는 것에 대한 예시적인 순서도를 도시한다.

도 7은 시간적 돌출 맵과 공간적 돌출 맵의 융합에 대한 예시적인 순서도를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0019] 본 발명은 상대적으로 적응된 처리 장치들을 포함하는 임의의 전자 디바이스에서 실현될 수 있다. 본 발명은 특히 돌출 기반 애플리케이션(saliency-based application)이 요구되는 저전력 디바이스에 유용하지만, 이에 한정되지 않는다. 예를 들어, 본 발명은 셋톱 박스, 태블릿, 게이트웨이, 텔레비전, 모바일 비디오 전화, 개인용 컴퓨터, 디지털 비디오 카메라 또는 자동차 엔터테인먼트 시스템에서 실현될 수 있다.
- [0020] 현 발명은, 인코딩된 스트림들이 약간의 추가 산출 비용으로 돌출 맵을 도출하는데 사용될 수 있는 정보를 이미 포함한다는 사실을 개시 및 활용한다. 정보는 풀 디코딩(full decoding) 동안 비디오 디코더에 의해 추출될 수 있다. 또한, 일부 디코더는, 그것을 완전히 디코딩하지 않고 단지 비디오 스트림을 분석(parse)하도록 구현될 수 있다.
- [0021] 도 2에 도시된 제1 예시적인 실시예에서, 돌출 맵 MAP의 산출은 공간적 돌출 맵 산출(spatial saliency map computation; SSC)을 단지 포함한다.
- [0022] 도 3에 도시된 제2 예시적인 실시예에서, 돌출 맵 MAP의 산출은 시간적 돌출 맵 산출(temporal saliency map computation; TSC)만을 포함한다.
- [0023] 도 4에 도시된 제3 예시적인 실시예에서, 돌출 맵 MAP의 산출은 공간적 돌출 맵 산출 SSC, 시간적 돌출 맵 산출 TSC, 및 산출된 시간적 돌출 맵과 산출된 공간적 돌출 맵의 융합 FUS를 포함한다.
- [0024] 제1, 제2, 및 제3 예시적인 실시예에서 산출된 공간적 및/또는 시간적 돌출 맵은, 진입 압축 스트림(incoming compressed stream; ICS) 내의 인코딩된 비디오 VID를 전체적으로 디코딩 DEC하지 않고 진입 압축 스트림 ICS으로부터 이용가능한 정보로부터 산출된다.
- [0025] 본 발명은 특정 코딩 체계에 한정되지 않는다. 진입 압축 스트림 ICS는 임의의 예측 인코딩 체계, 예를 들어, H.264/MPEG-4 AVC, MPEG-2, 또는 그 외의 것을 사용하여 압축될 수 있다.
- [0026] 다른 예시적인 실시예들에서, 공간적 돌출 맵 산출 SCC는 코딩 비용 추정에 기초한 것이다. Z. He의: 산타 바바라, 캘리포니아 대학, 박사-논문 (2001), "비주얼 코딩 및 통신에 대한 레이트 제어 및 ρ -도메인 레이트-왜곡 분석(ρ -domain rate-distortion analysis and rate control for visual coding and communication)"은, 블록 변환의 논-제로(non-zero) 변환 계수의 개수는 블록의 코딩 비용에 비례한다는 것을 기술한다. 도 5에 예시적으로 도시된 공간적 돌출 맵 산출 SCC는 이러한 사실을 활용하고 이러한 블록들의 코딩 비용들을 사용하여 결정된 돌출 값들을 인트라-코딩된 블록들에 할당하는데, 코딩 비용은 He에 의해 기술된 바와 같이 로-도메인 모델을 사용하여 결정된다.
- [0027] 대부분의 시간동안 단지 상대적인 돌출이 중요하므로, 돌출 맵이 정규화될 수 있다.
- [0028] 코딩 비용 외에, 돌출 값을 결정하는데 블록 크기가 더 사용될 수 있다. 블록 크기가 작을수록 보통 물체들(objects)의 가장자리와 연관되므로 관심을 끈다. 매크로-블록 비용 맵은 더 작은 블록들로 분해한 수만큼 증가된다. 예를 들어, 각 블록에 대한 비용 값은 하위-블록 분해의 경우에 두 배가 된다.
- [0029] 인터-예측 또는 바이-예측(bi-prediction)을 사용하여 인코딩된 블록들에 대해, 움직임 정보는 스트림으로부터 추출될 수 있고, 이는 결과적으로 인터-예측 또는 바이-예측을 위해 사용된 하나 이상의 참조 이미지에 대해 결정된 공간적 돌출 맵의 움직임 보상을 위해 사용될 수 있다.
- [0030] 시간적 돌출 산출 TSC는 도 6에 예시적으로 도시된 바와 같이 움직임 정보에 기초한 것이다. 따라서, 인터-예측 또는 바이-예측 프레임들에 대해서만 결정된다. 인터- 또는 바이- 예측된 프레임들 내에서, 인트라-코딩된 매크로-블록들은, 인터- 또는 바이- 예측에 의해 잘 예측할 수 없는, 고속 움직임을 보이거나 커버되지 않는 영역들 나타낸다. 예시적인 실시예에서, 바이너리 인트라-코딩된 블록맵 ICM은 시간적 돌출 맵을 결정하는데 사용된다. 바이너리 인트라-코딩된 블록 맵에서, 각각의 인트라 블록은, 예를 들어, 값 1을 취한다.
- [0031] 두드러지고, 주목을 끄는 움직임을 나타내는 움직임 벡터들은 잘 예측될 수 없고 따라서 코딩을 위해 상당히 더 많은 비트를 요구하기 때문에, 움직임 벡터 코딩 비용 맵 MCM은 시간적 돌출 맵을 결정하는데 더 사용된다.
- [0032] 움직임 벡터 코딩 비용 맵 MCM과 인트라-코딩된 블록 맵 ICM이 정규화되고 추가된다. 결과적으로 얻은 맵에서 블록들에 할당된 시간적 돌출 값들은 스킵(SKIP) 또는 다이렉트(DIRECT) 모드에서 코딩되는 블록들에 대해 감쇠

될 수 있다. 예를 들어, 스킵 또는 다이렉트 모드에서 인코딩된 블록들의 코딩 비용은 팩터 0.5씩 가중되는 한편 다른 모드들에서 인코딩된 블록들의 코딩 비용은 달라지지 않는다.

[0033] 공간적 돌출 산출 SSC과 시간적 돌출 산출 TSC로부터 얻어지는 돌출 맵들의 융합 FUS은 간단한 추가일 수 있다. 또한, 도 7에 예시적으로 도시된 바와 같이, 공간적 돌출 맵과 시간적 돌출 맵은 추가되기 전에 가중치 a, b로 가중된다. 가중치 a는 프레임에서 인트라-코딩된 블록들의 상대적인 양에 의존하고, 가중치 b는 프레임에서 인터- 또는 바이- 예측 블록들(P 또는 B)의 상대적인 양에 의존한다. 융합 FUS는 또한 비트-레이트 변화 및 코딩 유형에 의존하는 가중치 c로 가중된 이전 프레임의 이전 돌출 맵을 사용할 수 있다.

[0034] 본 발명자들의 실험들은 a, b 및 c에 대한 다음의 예시적인 값들이 양호한 결과를 생성한다는 것을 보여 주었다:

[0035]
$$a = \frac{1}{12} + \frac{\text{number_of_I_MB}}{4 \times \text{number_of_MB}}, \quad b = \frac{1}{12} + \frac{\text{number_of_P_MB} + \text{number_of_B_MB}}{4 \times \text{number_of_MB}}$$

[0036]
$$c = \frac{1}{12} + \frac{f(\text{bitRate}, \text{type})}{4}$$

[0037] 여기서,

[0038] 바이-예측된 프레임들(B-프레임들)에 대해:
$$f(\text{bitRate}, \text{type}) = \frac{1}{2} + \Delta\text{bitRate}$$

[0039] 인터-예측된 프레임들(P-프레임들)에 대해:
$$f(\text{bitRate}, \text{type}) = \frac{1}{4} + \Delta\text{bitRate}$$

[0040] 인트라-예측된 프레임들(I-프레임들)에 대해:
$$f(\text{bitRate}, \text{type}) = \frac{1}{8} + \Delta\text{bitRate}$$

도면

도면1

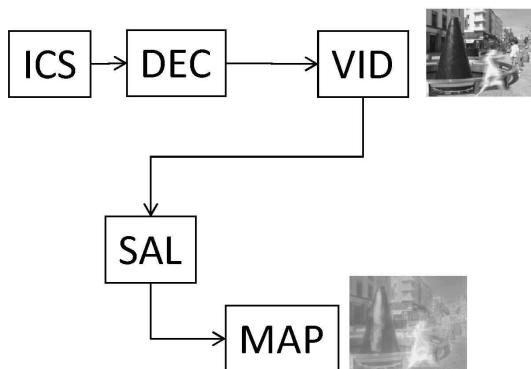
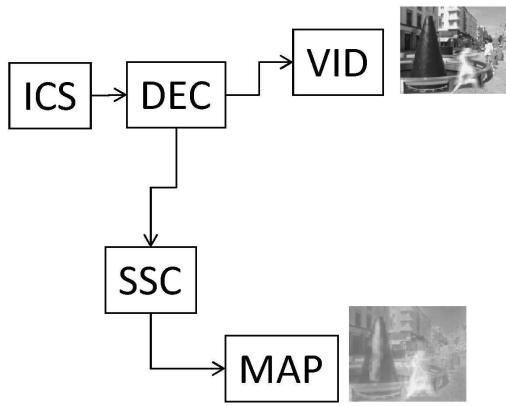


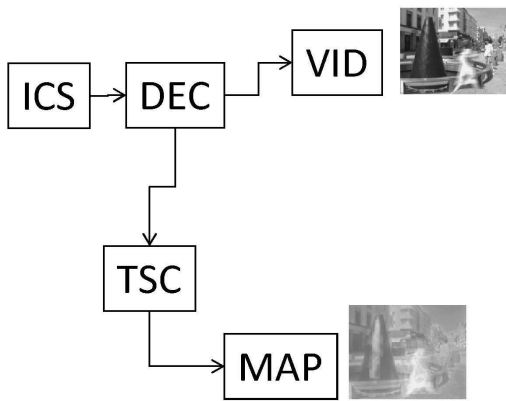
Fig. 1 - prior art

종래 기술

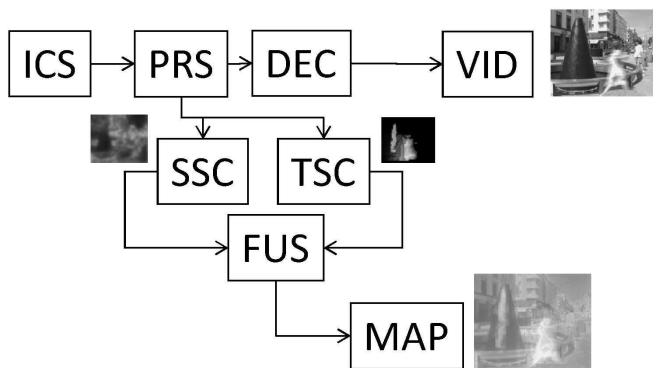
도면2



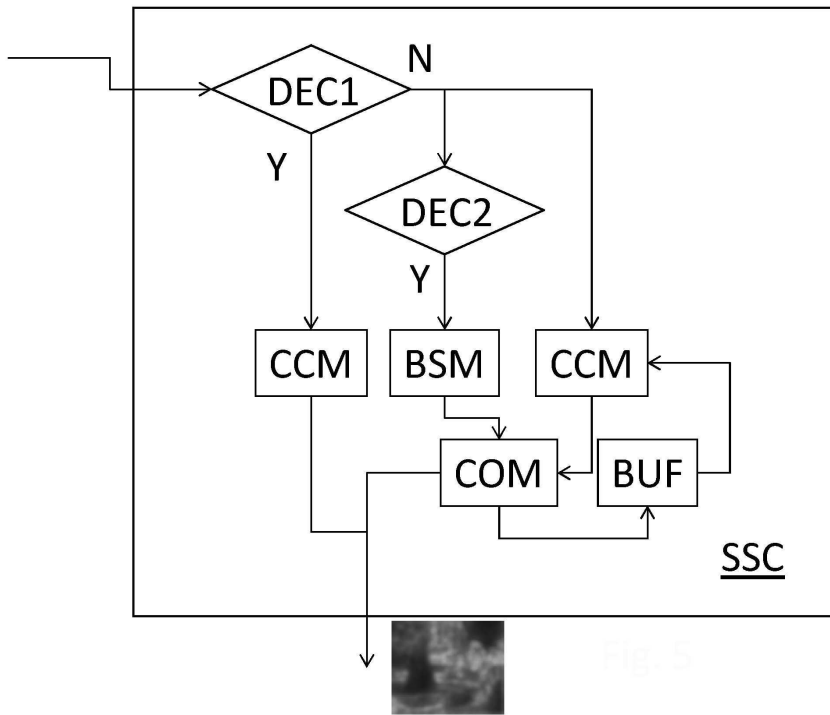
도면3



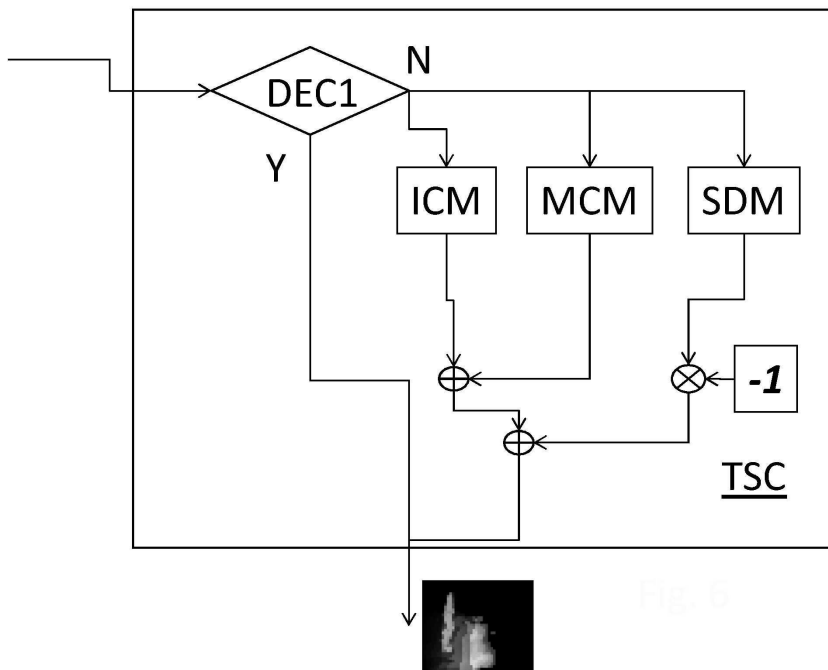
도면4



도면5



도면6



도면7

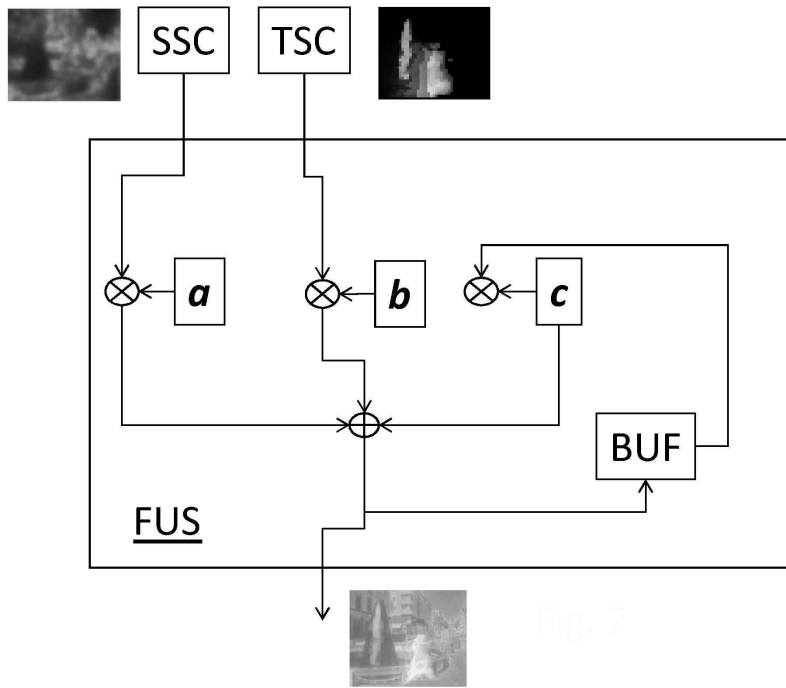


Fig. 7