

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。 Int. Cl. H04N 7/36 (2006.01)		(45) 공고일자	2006년05월24일
		(11) 등록번호	10-0582856
		(24) 등록일자	2006년05월17일
(21) 출원번호	10-1999-7004575	(65) 공개번호	10-2000-0069107
(22) 출원일자	1999년05월24일	(43) 공개일자	2000년11월25일
번역문 제출일자	1999년05월24일		
(86) 국제출원번호	PCT/IB1998/001241	(87) 국제공개번호	WO 1999/16251
국제출원일자	1998년08월12일	국제공개일자	1999년04월01일
(81) 지정국	국내특허 : 일본, 대한민국, EP 유럽특허 : 오스트리아, 벨기에, 스위스, 독일, 덴마크, 스페인, 프랑스, 영국, 그리스, 아일랜드, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투갈, 핀란드, 사이프러스,		
(30) 우선권주장	97202912.8	1997년09월23일	유럽특허청(EPO)(EP)
(73) 특허권자	코닌클리케 필립스 일렉트로닉스 엔.브이. 네덜란드왕국, 아인드호펜, 그로네보르스베그 1		
(72) 발명자	드하안게라드 네덜란드,아아아인드호펜5656,홀스트란6 슈트텐로버트안 네덜란드,아아아인드호펜5656,홀스트란6 펠라고티안나 네덜란드,아아아인드호펜5656,홀스트란6		
(74) 대리인	이병호 정상구 신현문 이범래		

심사관 : 최훈

(54) 움직임 추정 및 움직임 보상된 보간법

요약

본 발명의 움직임 추정 방법에서, 적어도 2개의 움직임 파라미터 세트들은 입력 비디오 데이터(n, n-1)로부터 생성되며 (PE1-PEn), 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터가 계산될 수 있다. 하나의 움직임 파라미터 세트는 화상 내의 모든 화상 부분들에 대한 제로 속도를 가리

키고, 움직임 파라미터 세트 각각은 대응하는 국부적 정합 에러를 갖는다. 출력 움직임 데이터는, 적어도 2개의 움직임 파라미터 세트들에 의존하여 입력 비디오 데이터(n, n-1)로부터 결정되고, 출력 움직임을 계산하는데 있어서 움직임 파라미터 세트 각각의 중요도는 움직임 파라미터 세트들의 국부적 정합 에러에 의존한다.

대표도

도 3

색인어

움직임 파라미터 세트, 움직임 추정, 움직임 보상, 정합 에러, 보간

명세서

기술분야

본 발명은 움직임 추정 방법 및 장치와 움직임-보상된 보간 장치를 포함하는 비디오 디스플레이 장치에 관한 것이다.

배경기술

움직임 벡터들은 코딩, 잡음 감소 및 스캔 속도 변환 등의 응용 영역에 사용된다. 이들 응용들 중 일부, 특히 프레임 속도 변환은 물체의 참-움직임을 추정할 필요가 있다[참고 10,11]. 다른 응용들, 예를 들면 비월-대-순차(interlaced-to-sequential) 스캔 변환은, 잔존 엘리어스(alias)의 진폭을 낮추기 위해 움직임 벡터들의 높은 정확성을 요한다[참고 12, 13]. 마지막으로, 움직임 추정기의 단가가 결정적으로 중요한 범주의 응용들, 예를 들면 움직임 추정의 소비자제품 응용 분야가 있다[참고 14, 15]. 참-움직임을 추정하기 위한 여러 가지 알고리즘들이 제안되어 왔다[참고 3,10,11,15-17]. 복잡도가 낮은 움직임 추정을 실현하기 위한 알고리즘 역시 제안되어 왔으며[참고 3,14,15,18-20], 대개 서브-화소 정확도를 허용하는 펠-재귀적(pel-recursive) 알고리즘 외에(예, [참고 21,22] 참조), 고도로 정확한 움직임 벡터들을 산출하는 많은 블록-매칭 알고리즘들이 보고되고 있다[참고 10,23,24].

수년 전에, 프레임 속도 변환에 필요한 참-움직임을 소비자 제품 응용에 요구된 낮은 복잡성 조건과 결합한 재귀적 검색 블록-정합기(recursive search block-matcher)가 제안되었다[참고 3]. 이러한 설계는, 필름 자료를 텔레비전 상에 표시할 때 그 움직임 묘사를 개선하고, 전송 속도가 상이한 화상 리프레시 속도로 시퀀스를 디스플레이할 때 발생하는 움직임의 경우에 화상 디테일이 흐릿해지는 것을 제거하기 위해 움직임 추정 및 보상 기술을 적용한 필립스 IC(MELZONIC, SAA4991)[참고 6,25]로 상용화되어 있다. 이러한 처리의 주된 해결 과제는, 스크린의 주어진 위치에서 물체가 이동하는지의 여부와, 이동하는 경우, 얼마나 빠르게 어느 방향으로 이동하는지를 가리키는 움직임 벡터를 추정하는 것이다. 공지된 IC에서 이러한 과제는, 화상을 블록들로 분할하고, 정합 기준을 최소화함으로써 모든 화소들의 블록에 대해 움직임 벡터를 계산하는 이른바 블록-정합기에 의해 수행된다. 이러한 처리의 위험은, 움직임 벡터 필드가 원치 않는 불균일성을 겪는 경우에 인접한 화상들로부터 보간되고 움직임 벡터들을 이용하는 움직임-보상된 화상은 블록 왜곡을 보일 수 있다는 점이다. 이러한 위험을 허용 가능한 수준으로 낮추기 위해, 참고문헌 [참고 6]에서의 IC는 후보 벡터[참고 3]의 공간적 및 시간적 예측에 기초한 개선된 일관성을 갖는 블록-정합기를 적용한다. 움직임 추정에 대한 이러한 접근법의 유익한 부수적 효과는 기능에 필요한 처리 전력이 상당히 감소한다는 것으로서, 이것은 특히 후보 벡터 개수가 대단히 제한되기 때문이다.

컴퓨터 비전 및 패턴 인식에 대한 1993 IEEE 컴퓨터 소사이어티 컨퍼런스 회보 361-366페이지, J.Y.A. Wang 및 E.H. Adelson의 논문 "움직임 분석을 위한 층별 표시(Layered representation for motion analysis)" [참고 29]는, 동종 움직임 분석(affine motion analysis) 및 클러스터링 기술(clustering technique)을 사용하여, 화상을 동일한 방향으로 움직이는 영역들로 분할하는 기술을 개시하고 있다. 화상은 어클루전(occlusion) 및 심도 정렬(depth ordering)에 관한 정보와 함께 한 세트의 층들로 분해된다. 장면은 4개의 층으로 분석되고, 그 다음 연관된 움직임 파라미터들과 함께 각각의 층의 단일 화상으로 시퀀스가 표시된다.

발명의 상세한 설명

특히, 본 발명의 목적은 복잡도가 한층 더 감소된 움직임 추정기를 제공하는 것이다. 이를 위해, 본 발명의 제1 특징은 특허 청구 범위 제1항 및 제9항에 정의된 방법 및 장치를 제공한다. 본 발명의 제2 특징은 특허 청구 범위 제7항 및 제10항에 정의된 방법 및 장치를 제공한다. 본 발명의 제3 특징은 특허 청구 범위 제11항에 정의된 비디오 디스플레이 장치를 제공한다. 유익한 실시예들은 종속항들에 정의된다.

본 발명의 주된 특징에 따라 움직임을 추정하는 방법에서, 적어도 2개의 움직임 파라미터 세트들이 입력 비디오 데이터로부터 생성되며, 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터들이 계산될 수 있다. 하나의 움직임 파라미터 세트는 화상 내의 모든 화상 부분들에 대한 제로 속도를 나타내고, 움직임 파라미터 세트 각각은, 화소 블럭마다 결정되는 정합 에러와 같은 대응하는 국부적인 정합 에러들을 가진다. 출력 움직임 데이터는 적어도 2개의 움직임 파라미터 세트들에 의존하여 입력 비디오 데이터로부터 결정되고, 여기서, 출력 움직임 데이터를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요성(가중치 인자들(W)에 의해 결정됨, 수학적 17,18 및 수학적 20, 21 참조)은 움직임 파라미터 세트들의 국부적 정합 에러들에 의존한다. 국부적 정합 에러들은, 전체 화상에 대해 계산된 정합 에러들과 같은 전역적 정합 에러와 대조되는 것으로 이해되어야 한다.

본 발명의 또 다른 특징에 따른 비디오 데이터의 움직임-보상 방법에서, 적어도 2개의 움직임 파라미터 세트들이 입력 비디오 데이터로부터 생성되며, 하나의 움직임 파라미터 세트는 제로 속도를 가리키며, 움직임 파라미터 각각은 대응하는 정합 에러들을 가지며, 출력 비디오 데이터는 적어도 2개의 움직임 파라미터 세트들에 의존하여 입력 비디오 데이터로부터 보간되며, 출력 비디오 데이터를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요성은 움직임 파라미터 세트들의 정합 에러들에 의존한다.

한 실시예에서, 그 감소는 상당하므로, 필립스 트리메디아(TriMedia) 프로세서와 같은 완전히 프로그램가능한 장치 상에서 처리가 수행될 수 있다.

본 발명의 이들 목적 및 다른 특징들은, 이하에 기술된 실시예들을 참조하면 명백해질 것이다.

도면의 간단한 설명

도 1은 본 발명에 따른 움직임 보상에 대한 2가지 가능성을 도시하는 도면.

도 2는 본 발명에 따른 움직임-보상된 보간기(interpolator)의 제1 실시예를 도시하는 도면.

도 3은 본 발명에 따른 움직임-보상된 보간기의 제2 실시예를 도시하는 도면.

도 4는 본 발명에 따른 움직임 파라미터 추정기의 블록도.

도 5는 본 발명에 따른 양호한 파라미터 추정기를 도시한 도면.

실시예

[참고 5]에서, 화상 시퀀스로부터 전역적 움직임을 추정하는 방법이 개시되었다. 화상 내의 움직임은, 인덱스 n 을 갖는 화상 내의 위치 \vec{x} 에서의 변위 벡터(displacement vector)로서 $\vec{D}(\vec{x}, n)$ 을 사용하여, 2차원 1차 선형 방정식으로 기술할 수 있다고 가정한다:

수학적 식 1

$$\vec{D}(\vec{x}, n) = \begin{bmatrix} p_1(n)+p_3(n)x \\ p_2(n)+p_3(n)y \end{bmatrix}$$

여기서, P_1 은 카메라의 패닝(panning)이며, P_2 는 틸팅(tilting)이며, P_3 은 주밍(zooming)이고, 전역적 움직임 벡터들의 추정만을 목표로 한다면, 파라미터 계산 수단으로의 입력은 [참고 5]에 기술된 것보다 더욱 간단해질 수 있음을 알 수 있다.

이와 같은 전역적 움직임 벡터만을 이용시에, 상향 변환(upconversion) 문제는 프로세싱의 가장 주된 해결 부분이 된다. [참고 1,4]는 화상 데이터의 강인한 움직임-보상된 시간적 보간 방법을 기술한다. 기본적인 개념은, 이전 필드(n-1)로부터 움직임 보상된 화소 mcl(left), 다음 필드(n)로부터 움직임-보상된 화소 mcr(right), 또는 양 인접 필드(n-1, n)로부터 움직임-보상되지 않은 평균(av) 중 어느 하나로 선택된 출력 화소를 생성하는 3-탭 메디안 필터로 구성되어 있다:

수학식 2

$$F_i(\bar{x}, n-\frac{1}{2}) = med(mcl, av, mcr)$$

이며, 여기서,

수학식 3

$$mcl = F(\bar{x} - \alpha \bar{D}(\bar{x}, n), n-1)$$

수학식 4

$$av = \frac{1}{2} (F(\bar{x}, n) + F(\bar{x}, n-1))$$

수학식 5

$$mcr = F(\bar{x} + (1-\alpha)\bar{D}(\bar{x}, n), n)$$

수학식 6

$$med(a, b, c) = \begin{cases} a, & (b \leq a \leq c \vee c \leq a \leq b) \\ b, & (a \leq b \leq c \vee c \leq b \leq a) \\ c, & (\text{그 외}) \end{cases}$$

움직임 보상에 사용된 화소들이 도 1에 개략적으로 도시된다. 상당히 강인(robust)하지만, 이전 필드(n-1)에서 대응하는 화소 l(left), 다음 필드(n)에서 대응하는 화소 r(right), 또는 양 인접 필드(n-1, n)로부터 움직임 보상된 평균 mcav를 선택하는 출력 화소를 생성하는 3-탭 메디안 필터를 포함하는 우리의 새로운 매우 제한된 움직임 추정기를 위한, 훨씬 더 강인한 알고리즘을 고려해 볼 수 있다.

수학식 7

$$F_i(\bar{x}, n-\frac{1}{2}) = med(l, mcav, r)$$

이며, 여기서,

수학식 8

$$l = F(\bar{x}, n-1)$$

수학식 9

$$mcav = \frac{1}{2} (F(\bar{x} - \alpha \bar{D}(\bar{x}, n), n-1) + F(\bar{x} + (1-\alpha)\bar{D}(\bar{x}, n), n))$$

수학식 10

$$r = F(\bar{x}, n)$$

그러나, 실제로 매우 강인한 이러한 상황-변환기는 움직임 보상의 이점을 심각하게 제한한다(움직임 보상은 보다 낮은 주파수들만으로 제한된다). 따라서, 양호한 실시예에 따라, 상황-변환기는 움직임 벡터의 기대되는 품질에 따라 제1 옵션과 제2 옵션 사이에서 적용된다. 제안된 보간기의 유리한 특징은 2개의 강인한 옵션들 간의 스위칭이 대단히 결정적이지는 않다는 것이다. 이는 대충의 결정이 허용될 수 있음을 의미하고, 이는 입력된 시퀀스의 (공간적으로) 감소된 크기 버전에 대한 적은 처리 전력으로 실현될 수 있다. 이와 같이 감소된 크기 입력 시퀀스는 파라미터 모델로부터 발생되거나 또는 제로 벡터로부터 발생되는, 위치당 (적어도) 2개의 움직임 벡터로 얻어진 정합 에러를 계산하는데 사용된다.

그 결과는 화상을 제로 벡터 모델 또는 계산된 파라미터 모델에 보다 적절한 층들로 분할하는 분할이다. 분할 마스크 (Segmentation Mask; SM)은 상황-변환기 UC의 추가 입력으로서 현재 사용되고 있다. 상황-변환기는 앞서 기술된 두 상황-변환기들 사이에서 스위칭/페이드하기 위해 분할 마스크(SM)를 사용한다(도 2 참조). 유효한 파라미터 모델의 경우에, 상황-변환기는 수학적 2의 보간쪽으로 향하는 경향이 있고, 그렇지 않으면, 수학적 7의 보간쪽으로 향하는 경향이 있다.

도 2에서, l 및 r 값(도 1 참조)은 av 값을 생성하기 위해 제1 평균기(AV1)에 적용된다. 제1 메디안 필터(MED1)는 av , mcl 및 mer 값 중 중앙값을 결정한다. mcl 및 mcr 값은 $mcav$ 값을 생성하기 위해 제2 평균기(AV2)에 적용된다. 제2 메디안 필터(MED2)는 $mcav$, l 및 r 중 중앙값을 결정한다. 상황-변환기(UC1)는 분할 마스크(SM)에 의존하여 메디안 필터(MED1, MED2)의 출력 신호로부터 보간된 값을 공급한다. 상황-변환기(UC1)의 출력 신호는, 입력 비디오 데이터(n , $n-1$) 사이에 출력 비디오 데이터($n-1/2$)를 디스플레이하기 위해 디스플레이 유닛(D)에 인가된다.

이 시점에서, 화상의 상이한 부분들(반드시 고정된 부분은 아님)에 대한 파라미터 모델 각각을 발생시키는 몇개의 파라미터 추정기 PE1 .. PEn들(충별 파라미터-기반의 추정기 및 상황-변환기를 보여주는 도 3 참조)가 병렬로 운영되는 다중 층들로의 확장이 고려될 수 있다. 이들 파라미터 추정기들(PEi)은 또 다시 각각의 모델이 유효한 경우에 대한 화상의 부분들을 발견하는 분할 회로(SC)의 입력이 된다. 즉, 화상의 각 부분에 대한 최상의 보간법을 가리키는 분할 마스크(SM)(파라미터 세트)를 결정한다. 상황-변환기(UC2)는 분할 마스크(SM)에 의존하여 화상 내의 각각의 분리된 층에 대한 최상의 보간법을 다시 선택하여야 한다.

도 3에서, 입력 필드(n)로부터 현재의 화상 데이터 및 입력 필드($n-1$)로부터 이전 화상 데이터가 움직임 파라미터($p21-p2m$.. $pn1-pnm$)를 결정하기 위해 파라미터 추정기(PE2 .. PEn)에 인가된다. 제1 파라미터 추정기(PE1)는 제로 파라미터를 제공한다. 입력 필드들(n 및 $n-1$)도 역시 다운-샘플러(D1, D2)를 통해 분할 회로(SC)에 인가된다. 상황-변환기(UC2)는 입력 필드들(n 및 $n-1$)로부터 출력 필드($n-1/2$)를 보간하기 위해, 분할 마스크(SM)가 가리키는 파라미터 세트로부터 수학적 1로 나타낸 방식으로 움직임 벡터를 계산한다. 가중치 인자(W)는 수학적 17 및 18과 중간 수학적 20과 21을 참조하여 이하에서 설명한다. 각각의 파라미터 추정기(PE2 .. PEn)는 움직임 파라미터들을 조절하기 위한 에러 계산을 포함한다. 이 계산은 대상 파라미터 추정기(subject parameter estimator, PE)에 의해 다루어지는 파라미터 세트에 대응하는 화상 부분들로 제한되는 것이 바람직하다. 이 목적을 위해, 대응하는 양에 가중치가 적합하게 되도록 가중치 인자(W)가 할당된다(일종의 퍼지 로직). 그 때문에, 보간 시에 각각의 화소에 대해 최저 추정 에러를 산출하는 파라미터 세트가 사용된다.

이하에서, 양호한 파라미터 추정기들이 논의될 것이다.

[참고 2]에서, 화상 시퀀스로부터 전역적 움직임 파라미터들의 추정 방법이 기술된다. 이 문헌은, 구배-기반(gradient-based)의 방법, 시뮬레이션된 어닐링(simulated annealing) 등의 다-차원 최적화 문제를 해결하기 위한 다양한 옵션들에 초점을 맞추고 있다. 본 발명의 또 다른 특징에 따르면, 전용 실리콘의 단가를 감소시키거나 또는 심지어 프로그램가능한 아키텍처(특히, 필립스 트리메디아 프로세서)상에서의 처리를 가능케하기 위해, 현저하게 감소된 개수의 동작들로 이들 움직임 파라미터들이 추정된다.

[참고 5]에서, 화상 시퀀스로부터 전역적 움직임 파라미터를 추정하는 한 방법이 개시되어 있다. 화상 내의 움직임은 2차원 1계 선형 방정식으로 기술될 수 있다고 가정된다. 보다 복잡한 파라메트릭 움직임 모델들이 제안되었으며[참고 2], 사실상 본 발명과 조합하여 적용될 수 있지만, 본 명세서에서는 논의되지 않을 것이다. [참고 5]에서, 파라미터 모델은 블록-기반의 움직임 추정기에 대한 매력적인 후보 벡터를 발생시키기 위해 사용되었다. 파라미터 추정기로의 입력은 이 블록-기반의 추정기로부터 얻어진 이전의 출력 벡터 필드였다. 만일 전역적 움직임 벡터의 추정만을 목표로 한다면, 파라미터 계산 수단으로의 입력은 더 단순화될 수 있음을 알 수 있다.

명확성을 위해, 수학적 1의 4 파라미터 모델만으로 제한하고, 먼저 파라미터 벡터 \vec{P} 를 아래와 같이 정의한다:

수학식 11

$$\bar{P}(n) = \begin{bmatrix} p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \end{bmatrix}$$

다수의 후보 파라미터 벡터들 $\bar{C}_p(n)$ 로부터 하기 식에 따라 계산된 정합 기준의 최소값을 갖는 것으로서 $\bar{P}(n)$ 를 선택하는 것을 우리의 과제로서 정의한다.

수학식 12

$$\epsilon(\bar{C}_n, n) = \sum_x |F(\bar{x}, n) - F(\bar{x} - \bar{D}(\bar{x}, n), n-1)|$$

이 에러 함수의 계산은 강한 서브-샘플링을 적용함으로써 매우 단순화될 수 있다. 실험 결과, 필드당 약 300개의 화소, 즉, 1000! 정도의 서브샘플링 인자 상에서 계산된 정합 기준에서 양호한 결과가 달성될 수 있다는 것이 드러났다. 그러나, 더 효과적인 방법은 클러스터링된 서브-샘플링이며, 즉, 선택된 화소들은 필드에 걸쳐 드문드문 분산된 그룹들을 형성한다.

최소화를 수행하는 제안은, 예측 벡터를 취하는 단계(적어도 3차원, 실시예에서는, 4차원), 적어도 하나의 갱신 벡터를 부가하는 단계, 및 수학식 13에 따라 최상의 벡터를 선택하는 단계를 포함한다. 하기 식에 따라 3개의 후보 파라미터 벡터들 $\bar{C}_p(n)$ 을 포함하는 후보 벡터 세트 $CS_p(n)$ 를 발생할 때, 양호한 결과가 실험적으로 얻어질 수 있었다.

수학식 13

$$CS_p(n) = \left\{ \bar{C}_p(n) \mid \bar{C}_p(n) = \bar{P}(n-1) + m\bar{U}_p(n), \bar{U}_p(n) \in US_p(n), m = -1, 0, 1 \right\}$$

단, $US_p(n)$ 는 이전 필드에 대한 파라미터 벡터와 다른, 0이 아닌 m에 대한 후보 벡터를 획득하기 위한 갱신 벡터를 정의하며, 하기 식에 따라 선택된다.

수학식 14

$$US_p(n) = \left\{ \begin{bmatrix} i \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ i \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \end{bmatrix} \right\}, (i=1,2,4,8,16)$$

예를 들어, 시간적 평활(smoothness)을 얻기 위해, 개개의 후보 벡터들(파라미터 세트들)의 정합 에러에 패널티가 부가될 수 있다. 또한, 예측 루프 내부나 외부의 파라미터 벡터들의 시간적 필터링은, 하나의 화상으로부터 다른 화상으로의 움직임 벡터의 갑작스런 변화를 방지하도록 고려된다.

지금까지 명세서에서 파라메트릭 움직임 모델은 전체 화상의 전역적 움직임을 기술하는 것으로 제안하였지만, 화상이 몇 개의 블록, 예를 들면 9개의 큰 블록으로 분할되고, 가능한 예측으로서 시간적 예측 뿐만 아니라 하나 이상의 공간적 예측도 역시 가능한, 그러한 대안들을 고려해 볼 수도 있다. 또한, 분할을 포함하고, 고정된 개수의 파라미터 추정기가 병렬로 운영되고, 각각의 파라미터 추정기는 분할 알고리즘이 가리키는 화상의 한 세그먼트에 집중하는, 그러한 또 다른 대안도 가능하다.

동작 개수(operations count)는 믿을 수 없을 정도로 낮다. 에러 기준의 계산은, 매 반복(iteration)마다 후보 벡터당 약

1000개의 연산들에 상당한다. 상기 기술된 구현의 경우, 이것은 화소당 $\frac{3.16.1000}{720.288} \approx \frac{48}{207} \approx 0.23$ 연산을 초래한다. 이는 [참고 6]의 추정기와 비교해 보면 그 크기가 한 차수 또는 두 차수 정도 감소된 값이다.

도 4는 본 발명에 따른 움직임 파라미터 추정기의 블록도를 도시한다. 제1 및 제2 후보 파라미터 세트들($Cp1$, $Cp2$)은, 2개의 후보 움직임 벡터들($Cv1$, $Cv2$)을 얻기 위해 멀티플렉서(MUX) 및 파라미터-대-벡터 변환기(PVC)에 적용된다. 제1 후보 파라미터 세트($Cp1$)는 멀티플렉서(MUX)의 이전 출력 파라미터 세트 $P(n)$ 이다. 제2 후보 파라미터 세트($Cp2$)는 갱신 파라미터 세트(Up)를 제1 후보 파라미터 세트($Cp1$)에 부가함으로써(가산기 AD) 얻어진다. 갱신 파라미터 세트(UP)는 $\text{mod}(n)$ 카운터(CNT)의 결과를 룩업 테이블(LUT)에 적용함으로써 얻어진다. 2개의 에러($E1$, $E2$)를 얻기 위해 후보 움직임 벡터($Cv1$, $Cv2$)는 에러 계산기(EC)에 적용되고, 현재 필드 및 이전 필드(n , $n-1$)도 역시 적용된다. 최소 회로(MIN)는, 출력 파라미터 세트 $P(n)$ 를 얻기 위해, 어떤 에러가 더 작은지를 판정하여 멀티플렉서(MUX)에 대한 선택 신호(s)를 얻는다.

이하의 부분은 비디오 데이터로부터 움직임 파라미터들을 추정하는 양호한 방법들을 기술한다. 움직임 추정은 비디오 데이터의 코딩 및 스캔 속도 변환에 적용된다. 움직임 추정기로의 입력시에 이러한 비디오 데이터의 화상 속도는 통상적으로 고정되지만, 이러한 데이터의 발생지인 비디오 소스의 화상 속도는 처리된 데이터의 화상 속도와 상이할 수 있다. 특히, 이것은 필름 자료가 비디오로 변환될 때, 또는 한 비디오 표준의 비디오 자료가, 움직임 추정기 이전의 비디오 체인내의 다른 표준으로 변환될 때 발생한다.

필요한 화상 속도 변환을 다루는 통상적인 방식은, 새로운 화상이 이용가능할때까지 가장 최근의 화상을 이용하는 것이다. 느린 화상 속도로부터 보다 빠른 화상 속도로 변환될 때, 이는 새로운 포맷의 소스 화상의 반복을 의미하는 반면, 빠른 속도로부터 느린 속도로의 변환은 소스 재료 화상의 간헐적인 스킵을 초래한다. 어느 경우든, 생성된 비디오는 불규칙적인 움직임 패턴(이상 진동)을 나타내고, 이것은, 움직임은 강한 시간적 일관성을 가진다는 움직임 추정기에서의 공통적 가정을 위반한다. 이러한 가정으로부터 이익을 얻고자 시도하는 움직임 추정기에서, 시간적 예측 벡터들을 사용함으로써, 불규칙한 움직임 작용으로 인해 시간적 예측 벡터들의 유용성이 없어진다는 문제점이 발생한다. 추정된 움직임 벡터들의 심각한 열화가 발생할 수 있다.

[참고 9]에서, 50Hz 텔레비전 표준으로 전송되는 영화 자료에 대해 이 문제에 대한 해결책이 개시되었다. 그 아이디어는 반복적인 화상이 발생할 때 벡터 예측 메모리를 재순환(recirculate)시키는 것이다. [참고 8]에서, 반복적이지 않은 화상이 발생될 때까지 '이전' 화상을 저장한 화상 메모리가 재순환되는 방법이 기술되어 있다. 두 선행 기술 방법의 공통된 특징은, 메모리 제어를 변경하기 위해 패턴이 알려져야만 한다는 것이다.

본 발명의 현 특징의 목적은, 신뢰성있게 움직임을 추정하기 위해 반복 패턴의 선행적 지식을 요구하지 않으면서 매우 강한 움직임 추정 방법을 제공하는 것이다. 이 목적을 위해, 움직임 추정기는 하나 이상의 이전 화상 쌍(반복 패턴의 최대 길이만큼)으로부터 시간적 예측 벡터들을 취하고, 이들중 최상의 것을 추정 과정에 대한 기초로서 선택하거나, 이들 모두를 정합 과정에서의 후보로서 사용한다.

이러한 해결책은 특히, 기억시킬 움직임 벡터의 개수가 매우 적은 객체-기반의 움직임 추정기에서, 경제적으로 정당화할 수 있다. 알고리즘의 소프트웨어 버전은 필립스 TM1000(TriMedia) 프로세서에 대해 실시간으로 수행하는 것으로 나타났다.

도 5는 본 발명에 따른 양호한 파라미터 추정기를 예시한다. 움직임 파라미터들 $P(n)$ 를 얻기 위해, 현재 필드(n)로부터 현재 화상 데이터 및 이전 필드($n-1$)로부터 이전 화상 데이터가 움직임 파라미터 추정기 유닛(MPE)에 적용된다. 화상 지연($D1$, $D2$, .. Dn)은 움직임 파라미터 추정기 유닛(MPE)에 대한 움직임 파라미터 $P(n)$ 의 지연된 버전들($TP1$, $TP2$, .. TPn)을 제공한다.

본 개시의 이하의 부분은, 화상이 복수 개의 층으로 분할되는 것, 즉 층별 움직임 추정에 관한 것이다.

영역-기반의 움직임 추정기는 블록-기반의 움직임 추정기에 대한 대안으로서 도입되어 왔다. 블록-기술 움직임 보상은 H.261/263(ISDN 라인을 통한 비디오-회의), MPEG-1(멀티미디어) 및 MPEG-2(모든-디지털 TV 애플리케이션) 등의 비디오 압축에 대한 국제 표준으로 채택되고 있다. 이들 표준은 특정 움직임 추정 방법을 명시하고 있지 않지만, 블록-기반의 움직임 추정 방법이 자연스럽게 선택되고 있다.

그러나, 움직임 추정을 위한 단위로서 블록의 사용은 아티팩트를 감소시킬수도 있는데, 이것은, 물체의 경계는 일반적으로 블록의 경계에 대응하지 않으며, 어떠한 공간적-시간적 일관성 조건도 존재하지 않는다면, 인접한 블록들은 실질적으로 상이한 움직임 벡터들을 할당받을 수 있기 때문이다.

블록 아티팩트들의 문제를 해결하고 이동 가장자리를 따라 보다 정확한 예측을 제공하려는 한 유망한 시도는, 움직임 필드를 분할한 것이다. 움직임 정보 및 패턴 정보(세기, 윤곽 텍스처)는 영역-기반의 (임의적 형상의) 움직임 추정을 달성하기 위해 사용되고, 다음 목표는 물체, 가능하면 MPEG-4 '시청각적(AV) 물체'를 다루는 것이다.

화상을 분할하고, 화상 시퀀스로부터 이들 화상 세그먼트에 대한 움직임 파라미터를 추정하는 여러 가지 방법들이 제안되고 있다[참고 2]. 분할을 수행하는데 있어서 이들 방법의 전략에 따라, 이들 방법은 상향식 방법, 하향식 방법 및 층별 표시법으로 분류될 수 있다. 개개의 카테고리의 특성을 간단히 요약하자.

* 상향식 방법

이 과정은, 패턴 정보에 기초하거나 미리 계산된 밀도 움직임 벡터 필드에 기초한 화상의 인트라-프레임 분할에서 시작한다. 분할은 일반적으로 많은 수의 작은 영역들을 초래한다. 이어서, 이들 영역들은 일반적으로 그들의 움직임에 대한 정보를 사용하여 병합된다, 즉, 유사한 움직임을 갖는 영역들이 하나의 영역으로 병합되고, 그 다음, 움직임 파라미터들이 재계산된다. 이러한 방법은 그 목표가 객체-지향적 코딩일 때 상당히 대중적인 것으로 입증되었다. 이 예들은 [참고 26-29]에 있다.

* 하향식 방법

이 과정은 큰 영역들에서 초기 화상 분할에서 시작한다. 이들 영역들은 계산된 움직임 모델이 정확도가 결여되는 경우에 세분되고, 움직임 파라미터들이 재계산된다. 초기 분할은 일반적으로 변경된/변경되지 않은 규칙에 일반적으로 기초한다. 즉, 현재 화상은 이전 화상과 비교되어: 동일한 위치에서 한 프레임에서 휘도 값이 다른 프레임 내의 휘도 값과 상당히 다르면, 화소는 '변경됨' 또는 그렇지 않으면 '변경되지 않음'으로 표시된다. 후속해서, '변경됨'으로 분류된 화상 부분은 그 영역에 대해 계산된 움직임 필드에 따라 움직임-보상될 수 있고, 상이한 움직임 영역들을 식별하기 위해 앞서 기술된 절차가 반복된다. 이 예들은 [참고 11, 30, 31, 36]에 있다.

2가지 기술은 조합될 수도 있다. 예를 들면, 초기 분할은 랜덤이거나 이전 추정[참고 11, 32]에 기초할 수 있다. 그리고, 연속적인 정밀화가 양 방향으로 일어난다. 추정 및 분할은 화상 시퀀스의 분석에 대한 통계학적 접근, 예를 들면 최대 가능성 추정 방법[참고 33]을 사용하여 동시에 수행될 수도 있다.

* 층별 표시법(layered representation methods)

이상적인 장면 분할은 분리된 객체들을 초래하며 3-D 정보를 포함하지만, 이것은 얻기 곤란하며 계산 집중적 방식이다. 따라서, 비디오 데이터는 분할되고 한 세트의 이동 층들, 즉, 단절된 경우에도 유사한 움직임을 보이는 화상 부분들로서 기술된다. 그 다음, 층들의 순서(깊이)가 결정된다. [참고 29,34,35]에 그 예가 있다. 완전한 3-D 모델보다 덜 복잡하고 시퀀스 내의 모든 물체를 다루는 모델보다 덜 복잡한 모델이 제안되었다. 이것은 양호한 실시예에서 채택된 모델이기 때문에, 이하에서 보다 상세히 기술할 것이다.

이들 층별 표시법에 따라, 비디오 데이터는 분할되고, 한 세트의 이동 층들, 즉, 단절된 경우조차 유사한 움직임을 보이는 영역들의 세트로서 기술된다. 그 다음, 층들의 순서(깊이)가 결정된다. 비디오 시퀀스의 층별 표시법은 스캔 속도 변환, 객체-트래킹, 비디오 압축, 코딩, 비디오 주해(annotation) 및 인텍싱 등의 여러 가지 용도에 대해 흥미를 유발한다. 층별 움직임 추정을 위한 많은 알고리즘이 이미 제시되었다[참고 29,34,37].

이들 알고리즘에서 중요한 점들중 하나는 움직임 추정/분할 문제를 해결하는 방식이다. 2가지 주요 접근법이 제안되었다.

* 순차 접근법

순차 접근법은 주된 움직임(dominant motion)을 순차적으로 추정함으로써 다중층들을 리졸브(resolve)하는데, 이것은 하향식 방법과 유사하다. 이러한 접근법의 주요 결점은, 한 층을 다루는 동안 최종 분할이 아직 알려지지 않았기 때문에, 상이한 움직임을 갖는 화상의 일부가 움직임 파라미터의 추정에 포함될 수 있고, 그 결과에 영향을 미칠 수 있다는 것이다.

* 동시 접근법

동시 접근법은 화상 내의 모든 층들을 동시에 추정하려 시도한다. 이는 미리-계산된 밀집한 움직임 벡터 필드를 사용함으로써 행해질 수 있다. 움직임 모델의 초기 세트는 주어진 움직임 벡터 필드[참고 29]에 대한 클러스터링 알고리즘을 사용함으로써 유도될 수 있다. 움직임 벡터 필드를 계산하는데 있어서, 약간의 평활 가정이 일반적으로 이루어진다. 이것은 그 경계가 객체/층 경계에 대응하지 않는 움직임 벡터 필드를 유도할 수 있으므로, 정확한 분할은 가능하지 않다. 대안으로서, 그 문제는 확률론적 문제로서 공식화될 수 있고, 다수의 모델들 및 이들의 지지층들의 최대 가능성 추정 방법이 기대치-최대화 알고리즘(expectation-maximisation algorithm)[참고 36]을 사용하여 수행될 수 있다. 마지막으로 언급한 2개 방법들의 주요 결점은 이들의 복잡도이다.

다른 중요한 점은 움직임 파라미터들이 추정되는 방식이다. 움직임 파라미터들의 추정이 휘도(luminance) 신호 자체에 대해 수행되는지의 여부에 의존하여, 직접 또는 간접으로서 분류될 수 있다. 직접적인 방법은 일반적으로 보다 강인한 것으로 간주된다. [참고 2]에서, 화상 시퀀스로부터 전역적 움직임 파라미터를 추정하는 여러 가지 방법이 기술된다. 구배-기반의 방법, 시물레이션된 어닐링 등의 다중-차원 최적화 문제를 해결하기 위한 여러 가지 옵션이 제안되고 있다. 제안된 알고리즘의 목적은 프로그램가능한 아키텍처 상의 움직임 추정이 가능하도록 현저하게 감소된 동작 개수로 이들 움직임 파라미터를 추정하는 것이다.

명세서의 현 부분의 논제는 전용 실리콘의 비용을 감소시키거나 심지어 프로그램가능한 아키텍처(특히, 필립스 트리메디아(TriMedia) 프로세서) 상에서의 처리를 가능케하기 위해 현저히 감소된 동작 수로 화상 시퀀스들을 추정/분할하는 것이다.

본 발명의 현재 특징은 층별 표시법을 목표로하는 움직임 추정/분할을 다룬다. 비록 간접적인 버전이 그럴듯하게 보이긴지만, 구현 비용을 가능한 한 낮게 유지하기 위해, 직접적인 방법으로서 구현에 초점을 맞추었다. 이는 조합된 움직임/분할의 선후 문제에 대한 격조 높은 해결책을 제공한다. 그 해답은 병렬로 동작하는 다른 파라미터 추정기들에 의해 다뤄진 정보에 의해, 주어진 층에 대한 파라미터 추정기의 최적 기준의 오염을 제한하는 가중치 과정으로 구성된다. 기존 프로그램가능한 아키텍처 상에서 실시간 동작하도록 움직임 추정기를 설계하는 것은 움직임 추정 문제에 대하여 심각한 제약을 부과하는데, 그 이유는 알고리즘의 복잡도가 극적으로 감소되어야 하기 때문이다. 층별 움직임 추정기는 이러한 목적으로 선택되었는데, 그 이유는, 블록들보다 더 적은 수의 층이 존재한다는 것을 알고 있을 때, 예를 들면 블록-기반의 움직임 추정보다 프로그램가능한 아키텍처 상의 구현이 잠재적으로 더 용이한 것으로 믿어지기 때문이다.

층별 표시법에서, 화상은, 다수의 층들, 즉, 단절된 경우조차 일관된 움직임을 보이는 화상 부분들로 분할된다. 어느 시퀀스의 외관적 움직임(광학적 흐름)은 파라메트릭 모델로 기술될 수 있다고 가정한다. 즉, 움직임은 단지 카메라 움직임과 불투명한 객체의 그리드 움직임(rigid motion)의 조합에 기인하는 것으로 가정한다. 따라서, 한 세트의 움직임 파라미터들은 움직임 필드 그 자체 대신에 각각의 층에 대해 추정될 수 있다.

유사한 움직임을 보이는 영역들 내의 화상들의 시퀀스를 분할하고 동시에 이들의 움직임을 추정하는 것은 그 자체로서 까다로운 문제인데, 이는 2가지 할당이 상호-의존적이기 때문이다. 하나의 영역에서 움직임을 정확히 추정하기 위해서는, 그 영역은 알려져야 한다. 그러나, 일관적으로 이동하는 화상의 영역들을 판별하기 위해서는 그들의 이동이 알려져야 한다. 유사-동시 움직임 추정 및 고정된 수의 층들까지의 분할을 위한 새로운 방법이 제공된다. 우리는 각각의 층에 대한 움직임 파라미터를 추정하는 문제를 해결하고, 동시에 각 층에 상이한 등급을 주는 계층구조를 도입하여 화상을 분할한다. 이러한 계층구조가 의미하는 2가지 목표는,

* 특정 층이 계층구조내의 보다 상위층에 의해 잘 커버되는 모든 화상 부분에 대해 추정되는 것을 방지하고,

* 특정 층이 계층구조내의 보다 하위층에 의해 잘 커버되는 화상 부분들에 의해 오염되는 것을 방지하는 것이다.

이어서, 파라미터 벡터들은 재귀적 접근법을 사용하여 병렬로 추정된다. 즉, 각각의 층에 대해 앞서 추정된 파라미터 벡터는 어느 갱신 벡터들이 부가되는지에 대한 예측으로서 사용된다. 선택된 파라미터 벡터는 최저 정합 에러를 초래하는 것이다. 이 후, 모든 층들의 파라미터 벡터들은, 화상을 상이한 층들로 분할하는데 있어서 함께 사용된다.

각각의 층(1)의 움직임은 단순 움직임 모델에 의해 기술된다. 층 내의 움직임은, 화상 내의 위치 $\vec{x} = (x, y)^T$ 에서 층(1)의 변위 벡터에 대하여 인덱스 n과 함께 $\vec{D}(\vec{x}, l, n)$ 을 사용한 2차원 1차 선형 모델에 의해 기술될 수 있다고 가정한다.

수학식 15

$$\bar{D}(\bar{x}, l, n) = \begin{bmatrix} s_x(l, n) + d_x(l, n).x \\ s_y(l, n) + d_y(l, n).y \end{bmatrix}$$

이러한 4-파라미터 모델에 의해, 줌과 마찬가지로 수평 및 수직적 변환(팬 및 틸트)이 기술될 수 있다. 보다 복잡한 파라메트릭 움직임 모델들이 제안되었으며[참고 2], 실제로 제안된 알고리즘과의 조합하여 적용될 수 있지만, 이하에서 논의되는 것은 아니다. 실험에서, 이러한 움직임 모델은 몇가지 자유도와 함께 사용되고 있다:

* 4개의 파라미터 모두 자유.

* 파라미터 s_x 및 s_y 는 자유이고, d_x 및 d_y 는 화상의 종횡비에 따라 고정된 비율로 결합됨(3-파라미터 모델).

* 파라미터 s_x 및 s_y 는 자유이고, d_x 및 d_y 는 0으로 고정됨(2-파라미터, 변환 모델).

* 파라미터 s_x 및 s_y 는 자유이고, d_x 및 d_y 는 0으로 고정됨(1-파라미터, 패닝 모델).

한 실시예에서, 제1층은 4개 또는 8개의 자유 파라미터를 갖는 반면, 후속하는 각각의 층은 계산 부담을 감소시키기 위해 선행층보다 적은 자유 파라미터를 갖는다.

본 발명은, 제로 벡터(움직임 없음)는 매우 흔하며 비디오 시퀀스에서 중요하며, 특히 스캔 속도 변환에서 의도된 용도에 대해 중요하다는 인식에 기초한다. 따라서, 제안된 알고리즘은 (명백히 추정되지 않는) 제로 파라미터 벡터에 의해 기술된 움직임과 함께, 층(0)에서 시작한다. 추가 층들(1, 1>0)의 파라미터 벡터들은 이들 각각의 파라미터 추정기(PE₁)에 의해 별도로 추정된다.

각각의 PE₁은 [참고 3]의 3D 재귀적 검색 블록 정합기와 동일한 기본 원리를 가진다. 앞서 추정된 파라미터 벡터는, 최상의 정합 파라미터 벡터가 선택된 후, 의사-랜덤 잡음 벡터에 따라 갱신된다.

수학식(15)의 파라미터 모델을 고려하면, 층(1, 1>0)의 파라미터들은 파라미터 벡터 \bar{P}_l 로 간주되며:

수학식 16

$$\bar{P}_l(n) = \begin{bmatrix} s_x(l, n) \\ s_y(l, n) \\ d_x(l, n) \\ d_y(l, n) \end{bmatrix}$$

다수의 후보 파라미터 벡터들 $\bar{CP}_l(n)$ 로부터, 최소값의 정합 기준을 갖는 것으로서 $\bar{P}_l(n)$ 을 선택하는 것을 과제로 정의한다. 에러 함수는 하기 식에 따라 계산되고:

수학식 17

$$\epsilon'(\bar{CP}_l(n)) = \epsilon(\bar{CP}_l(n)) + \sum_{x \in X_l} W_l(\bar{x}) \cdot \Pi(\bar{CP}_l(n))$$

여기서, 예를 들면 공간 평활도를 얻기 위해 개개의 후보 벡터(파라미터 세트)의 정합 에러에 패널티 $\Pi(\bar{CP}_l(n))$ 가 부가되고, ϵ 는 아래와 같고:

수학식 18

$$\epsilon(\bar{CP}_l(n)) = \sum_{x \in X_l} W_l(\bar{x}) \cdot |F_y(\bar{x}, n) - F_y(\bar{x} - \bar{D}(\bar{x}, l, n), n - 1)|$$

여기서, $W_i(\vec{x})$ 는 위치 \vec{x} 에 의존하는 가중치 인자이고, $F_i(\vec{x}, n)$ 은 서브-샘플링된 화상 내의 위치 \vec{x} 에 대한 인덱스 n을 갖는 휘도 값이며, X_1 은, 층(1)의 움직임 추정하고자하는 한 세트의 위치들 \vec{x} 이다(위치 \vec{x} 의 선택 모드를 아래 설명할 것이다).

화상들은 필드 베이스상에서 수평으로 4의 인자 및 수직으로 2의 인자에 의해 서브-샘플링되고, 각각의 원시 필드 F(n)로부터 서브-샘플링된 화상 $F_s(n)$ 을 발생한다. 이는 동작 개수의 원하는 감소에 강력하게 기여한다. 서브-샘플링은 움직임이 추정되는 객체들이 충분히 크기 때문에 허용된다. 화소 또는 심지어 F의 원시 화소 그리드에 대한 부분-화소 정확도를 달성하기 위해 부분 샘플링 그리드[참고 7]에 대한 보간이 요구된다.

제안된 최소화는 [참고 3,7]에 이용된 전략과 약간의 유사성을 보여준다. 즉, 예측 벡터(이 경우, 4-차원)를 취하고, 적어도 하나의 갱신 벡터를 가산하고, 수학식(18)에 따라 최상의 후보 벡터를 선택한다. 하기 식에 따라 3개의 후보 $\vec{CP}_i(n)$ 를 포함하는 후보 파라미터 세트 $S_{\vec{CP}_i}(n)$ 를 생성할 때, 양호한 결과가 실험적으로 얻어질 수 있다.

수학식 19

$$S_{\vec{CP}_i}(n) = \{ \vec{CP}_i(n) | \vec{CP}_i(n) = \vec{P}_i(n-1) + m\vec{UP}_i(n), \\ \vec{UP}_i(n) \in S_{\vec{UP}_i}(n), m = -1, 0, 1 \}$$

단, 갱신 파라미터 $\vec{UP}_i(n)$ 는 아래와 같이 갱신 파라미터 세트 $S_{\vec{UP}_i}(n)$ 로부터 선택된다.

수학식 20

$$S_{\vec{UP}_i}(n) = \left\{ \begin{bmatrix} i \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ i \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \end{bmatrix} \right\} \\ (i = 1, 2, 4, 8, 16)$$

예측 루프 내부 및 외부 모두에서 파라미터 벡터들의 시간적 필터링은, 하나의 화상으로부터 다른 화상으로 움직임 벡터의 갑작스런 변화를 방지하도록 적용된다.

지금까지 기술된 알고리즘은 한 쌍의 입력 화상에 대하여 하나의 반복을 수행한다. 이 알고리즘의 보다 빠른 수렴은 동일한 쌍의 입력 화상에 대하여 파라미터 추정기의 다중 반복에 의해 달성되고, 이 경우, 수학식(19)에서 $\vec{P}_i(n-1)$ 은 한쌍의 화상에 대한 초기 반복 후에 이전의 반복 $\vec{P}_i(n)$ 의 출력으로 대체된다.

층들의 계층적 구조가 제안된다. 이는 아래와 같이,

* 보다 상위층들로 잘 커버되는 화상 부분들을 배제한 X_1 에서 위치 \vec{x} 의 선택

* X_1 내에서, 계층구조의 보다 하위층들에 의해 잠재적으로 잘 커버되는 화상 부분들의 효과를 감소시키고, 이전의 분할에서 층(1)에 할당된 화소에 보다 큰 가중치 $W_i(\vec{x})$ 의 할당에 의해 달성된다.

계층구조의 최상위층을 제외한 각각의 추정기(제로 추정기)는, 보다 상위층의 모든 추정기들이 이전의 화상에서 성공적이지 못했던 영역들에서 계산된 정합 에러를 최소화한다. 위치들 X_1 의 세트는, 모든 상위 랭크된 층들의 정합 에러가 고정된 인자로 평균 블록 정합 에러를 초과하는 위치 \vec{x} 로 채워진다.

실험 결과, X_1 에서 위치들의 개수가 화상 내의 모든 화소의 약 2-5%만으로 제한될 때 양호한 결과가 잘 얻어질 수 있음을 나타낸다. 가장 효과적인 결과는 화상에서 클러스터된 서브-샘플링이다. 즉, 선택된 화소들은 전체 화상에 걸쳐 드문드문 분포된 그룹들을 형성한다. 현재의 응용에서, 16개 화소의 최대 50개의 클러스터가 선택된다(F_s 에서 모든 화소의 3%).

현재 추정기가 이전 층들에 의해 이미 커버된 움직임 추정하는 것을 방지하기 위해, X_1 의 정확한 선택이 필요하다.

위치-의존형 가중치 인자 $w_l(\bar{x})$ 는 이전 화상에서 발견된 분할 마스크 $SM(n-1)$ 에 의해 결정된다. 분할 마스크에 따라 현재 층(1)에 속하는 위치 \bar{x} 는 1보다 큰 가중치 인자를 가질 것이고, 상이한 층에 속하는 위치는 1의 가중치 인자를 갖는다. 현재 추정기가 계층구조에서 후속층들로 커버될 수 있는 움직임을 추정하는 것을 방지하기 위해 $w_l(\bar{x})$ 의 정확한 선택이 필요하다.

분할 단계는 알고리즘에서 가장 중요한 단계이다. 그 과정은 화상 내의 층들 중 하나, 즉 움직임의 한 모델을 각각의 화소 그룹에 할당하는 것이다. 이것은 기본적으로 각각의 화소 그룹에 가장 잘 정합하는 모델을 할당함으로써 달성된다(프레임 베이스 상에서 기본적으로 8X8 화소들만큼 큰 블록 \bar{B}).

각 층에 대해, 정합 에러는 하기 식에 따라 계산된다:

수학식 21

$$\sum_{\bar{x} \in \bar{B}} \epsilon(\bar{B}, l, n) = |F_s(\bar{x} + (1-\alpha)\bar{D}(\bar{x}, l, n), n) - F_s(\bar{x} - \alpha\bar{D}(\bar{x}, l, n), n-1)|$$

분할 마스크 $SM(\bar{B}, n)$ 는 최저 ϵ 를 갖는 층(1)을 블록 \bar{B} 에 할당한다. 분할의 시간적 위치는 α 에 의해 정의되고, 이 실험에서는 1/2로 설정하였다.

처리 전력을 절감하기 위해, 모든 블록 \bar{B} 에 대해 분할 마스크 SM 가 계산되어야 하는 것은 아니다. 대신에, 계산된 블록들은 오엽 배열(quincunx) 패턴으로 서브-샘플링될 수 있고, 그 후 분할 마스크에서 누락된 위치들은 보간된다(예를 들어, 인접한 것들로부터 가장 잘 발생하는 층 번호를 선택함으로써)[참고 7].

분할은 많은 층들이 존재할 때 보다 어려워지는데, 그 이유는 분할 작업은 완전 검색 블록 정합기(full search block matcher)의 분할 작업과 더욱 더 유사해질 것이기 때문이다. 완전 검색 블록 정합기와 유사한 비밀관성을 갖는 움직임 추정기의 출력을 방지하기 위해, 여분의 (평활화) 제약요건이 알고리즘에 추가되었다.

현재의 평활화 제약 요건은 아래와 같이 구성된다:

- * 공간적 평활화: ϵ 의 계산시에, 층을 할당받은 블록 \bar{B} 의 크기보다 더 큰 창을 취함으로써.
- * 시간적 평활화: 한 층의 계산된 ϵ 를, 상기 층이 이전 화상의 분할에서 선택되었다면, 보너스 값으로 감소시킴으로써.
- * 공간적 평활화: 분할에서 특이점(singular spot)을 제거하기 위해 머조리티 필터(majority filter)를 사용함으로써.

실험 결과로서, TriMedia상의 제1 구현에서 3-층 구조가 선택되었고, 층 0은 움직임 없음에 대응하여 추정되지 않고, 즉, 모든 파라미터는 0과 동일하다. 층 1은 2개의 자유 파라미터를 갖고, 층 2는 단 하나의 자유 파라미터를 갖는다. 각각의 입력 화상 쌍에 대해, 층 1의 파라미터 추정기는 5회 반복되고, 층 2의 추정기는 3회 반복된다.

서브-샘플링의 간단한 예비-필터링은 4x2 화소 블록 내의 화소값들을 평균함으로써 달성된다. 이는 서브-샘플링된 출력

화소당 약 10개의 연산 또는 입력 그리드의 화소당(CCIR 601/625 라인/2:1) $\frac{180.144.10}{720.288} \approx 1.25$ 연산을 취한다.

파라미터 추정기에서 에러 기준의 계산은 매 반복마다 후보 벡터당 약 1000개의 연산을 취한다. 상기 기술된 구현을 위해, 이는 화소당 $\frac{3 \cdot (5+3) \cdot 1000}{720 \cdot 288} \approx 0.12$ 연산을 초래한다(이것은 파라미터 추정의 모든 기능을 담당하지 못함). 분할에서 에러

기준의 계산은 매 블록마다 층당 약 10개의 연산, 따라서, 화소당 $\frac{3 \cdot (72 \cdot 90 / 2) \cdot 10}{720 \cdot 288} \approx 0.47$ 연산을 취한다(이는 분할의 모든 기능을 담당하지 못함). 이것은 MELZONIC(SAA4991)의 추정기와 비교해 한 차수 감소된 크기이다. TriMedia에 대해 부분적으로 최적화된 측정치는, 서브-샘플링에 대해 약 1.25, 파라미터 추정기에 대해 1.0, 및 분할에 대해 화소당 6.1의 동작 개수가 달성됨을 보여준다.

최종적인 변위 벡터의 사용을 포함하여 제안된 층별 움직임 추정기가, 25Hz 필름에서 50 Hz 디스플레이로의 화상 변환 속도에 대하여 시뮬레이션되었다.

움직임 추정기로부터 발생하는 벡터 필드는 매우 일관성있는 것으로 입증되었고, 스캔 속도 변환에 적합하였다. 얻어진 품질은 매력적이고, 대부분의 장면에 대해 MELZONIC(SAA4991)에 의해 달성된 품질에 필적하는 것으로 생각된다.

제안된 움직임 추정 알고리즘은, 구현에 기인한 벡터 범위 제한이 없었고, 이는 MELZONIC(SAA4991)에 비해 유리하다.

층별 비디오 표시법에서 전역적 움직임 추정 알고리즘을 움직임 파라미터의 추정으로까지 확장하는 방법이 제공되었다. 고정된 개수의 파라미터 추정기들이 병렬로 운영되고, 그 각각은 하나의 화상층에 대한 파라미터들을 계산한다. 분할은 화상의 각각의 부분을 정확한 층에 할당한다.

추정기는 병렬로 동작하지만, 어느 정도 계층성이 존재한다. 계층구조에서 최상위를 제외한 각각의 추정기는, 계층구조에서 보다 상위의 추정기가 이전 화상에서 성공적이지 못했던 화상 부분들에 대해 동작한다. 두 번째로, 각각의 추정기는 계층구조에서 보다 하위의 추정기에 의해 잘 다루어질 화상의 부분들에 의한 오염이 방지된다.

실험은, 현재의 결과가 전용의 설계(MELZONIC(SAA4991)에 의한 Natural Motion)에 의해 얻어진 것과 그렇게 동떨어지지 않음을 보여준다. 그러나, 이 알고리즘은 TriMedia와 같은 프로세서 상에서의 소프트웨어적 구현에 훨씬 더 적절하다.

마지막으로, 본 발명에 따른 알고리즘은 비디오 압축 및 코딩, 비디오 주해 및 인텍싱, 객체 트래킹 및 잡음 제거와 같은 움직임 추정의 다른 응용 분야에 대해서도 흥미를 유발할 수 있다.

본 발명의 제1 특징은 다음과 같이 요약될 수 있다. 전역적 움직임-보상된 상향-변환을 위한 새로운 방법이 개시되며, 그 방식은 층별 비디오 표시로의 용도로 확장되는 것으로 나타났다. 본질적으로, 전역적 움직임을 기술하는 파라미터들은 양호하게는 재귀적 접근법을 사용하여 추정된다. 이들 파라미터에 의해 생성된 국부적인 움직임 벡터는 움직임-보상된 화상을 발생시키기 위해 사용된다. 동시에, 분할 마스크는 감소된 크기의 화상 상에서 계산되고, 그 출력은 상이한 파라미터 세트들과 보간법들 사이에서 전환하기 위해 사용된다. 매력적인 저비용 버전이, 현재 입수할 수 있는 완전히 프로그램가능한 장치(Natural Motion on a TriMedia) 상에서 구현하는데 적합하다.

양호한 실시예의 하기 두드러진 특징은 주목할만한 가치가 있다. 비디오 데이터의 움직임 보상을 위해 이러한 방법을 실현하는 방법 및 장치는,

입력 비디오 데이터로부터 전역적 움직임 파라미터를 계산하기 위한 적어도 2개의 수단,

상기 적어도 2개의 전역적 움직임 파라미터 세트에 의존하여, 1개 이상의 입력 필드로부터 출력 비디오 데이터를 계산하기 위한 보간 수단을 포함하고, 여기서, 전역적 움직임 파라미터를 계산하기 위한 상기 적어도 2개의 수단은, 화상의 내용과 무관하게, 전체 화상에 대해 제로 속도를 나타내는 파라미터들을 제공한다.

양호하게는, 보간 수단은, 예를 들면, 3-탭 메디안 필터와 같은 정렬 통계 필터(order statistical filter)이며, 3-탭 메디안 필터는,

이전 필드에 대응하는 화소, 다음 필드에 대응하는 화소, 및 양 인접 필드로부터의 움직임-보상된 평균(제1 옵션) 또는,

이전 필드로부터 움직임-보상된 화소, 다음 필드로부터 움직임-보상된 화소, 및 양 인접 필드로부터 움직임-보상되지 않은 평균(제2 옵션)으로부터 출력 화소를 생성한다.

양호하게는, 분할 신호는, 제2 파라미터 세트로부터 계산된 국부적인 움직임 벡터가 감소된 크기의 입력 화상에 대해 최상의 정합을 산출하는 경우에, 제1 결정을 활성화(activate)한다.

양호하게는, 분할 신호는 입력 신호의 감소된 크기 버전으로부터 유도된다.

비디오 데이터의 움직임 보상을 위해 이러한 방법을 실현하는 방법 및 장치는,

입력 비디오 데이터로부터 전역적 움직임 파라미터를 계산하기 위한 적어도 2개의 수단,

상기 적어도 2개의 전역적 움직임 파라미터 세트에 의존하여, 1개 이상의 입력 필드로부터 출력 비디오 데이터, 및 입력 신호의 감소된 크기 버전으로부터 유도되는 분할 신호를 계산하기 위한 보간 수단을 포함한다.

양호하게는, 전역적 움직임 파라미터 계산 수단들 중 하나는, 화상의 내용과 무관하게 전체 화상에 대하여 제로 속도를 나타내는 파라미터들을 제공한다.

본 발명의 제2 특징은 다음과 같이 요약될 수 있다. 전역적 움직임 파라미터 추정을 위한 신규한 방법이 기술된다. 본질적으로, 화상 내의 전역적 움직임을 기술하는 파라미터들은 재귀적 접근법을 사용하여 추정되는데, 즉, 이전의 n -차 (n 은 움직임 모델에서 파라미터들의 개수임) 추정치가 (n -차원) 갱신 벡터가 부가되는 예측값으로서 사용된다. 출력 파라미터 벡터는 가장 작은 정합 에러를 초래하는 것이다. 이 알고리즘의 복잡도가 극도로 낮고, 그 품질이 높음으로써 TV 및 멀티미디어 용도로 장치 사용할 수 있게 하고, 트리메디아(TriMedia)와 같은 완전히 프로그래밍가능한 장치 상에서 실행할 수 있게 한다.

양호한 실시예의 하기 두드러진 특징은 주목할만한 가치가 있다. 화상 시퀀스의 움직임 파라미터들(파라미터 벡터)을 추정하기 위해 이러한 방법을 실현하는 방법 및 장치는,

예측 파라미터 벡터, 즉, 이전에 계산된 움직임 파라미터 추정치를 제공하기 위한 수단,

갱신 벡터 세트로부터 적어도 하나의 갱신 파라미터 벡터를 선택하기 위한 수단,

적어도 하나의 갱신 벡터에 상기 예측 벡터를 부가하기 위한 수단,

적어도 2개 필드들로부터의 데이터를 사용하여, 생성된 적어도 2개의 파라미터 벡터들의 품질(단가 함수)을 계산하기 위한 수단,

그들의 품질에 기초하여 상기 언급한 적어도 2개의 파라미터 벡터들로부터 최상의 벡터를 선택하기 위한 수단, 및

상기 선택된 파라미터 벡터를 움직임 파라미터 추정치로서 출력하기 위한 수단을 포함한다.

양호하게는, 패넬티, 시간적 필터링, 및 시간적 및/또는 공간적 예측이 적용된다.

본 발명의 제3 특징은 다음과 같이 요약될 수 있다. 비디오 데이터로부터 움직임 파라미터들을 추정하는 방법이 개시된다. 본 발명은, 간단한 화상 속도 변환 기술(가장 최근 화상의 반복)로 인해 불규칙적인 움직임을 나타내는 비디오 데이터에 대한 시간적 예측 움직임 추정을 허용한다. 그 솔루션은 다양한 이전 화상 쌍들로부터 취해진 다수의 시간적 예측 벡터들을 사용하는 것으로 이루어져 있다. 이 솔루션은 기억될 움직임 벡터수가 매우 적은, 특히 객체-기반의 움직임 추정기의 경우에 경제적으로 정당화할만하다. 이 알고리즘의 소프트웨어 버전은 필립스 TM1000(TriMedia) 프로세서 상에서 실시간 실행되는 것을 보여준다.

양호한 실시예의 하기 두드러진 특징은 주목할만한 가치가 있다. 비디오 데이터로부터 움직임 파라미터 벡터들을 추정하는 방법을 실현하는 방법 및 장치는 적어도 일부의 화상-부분들에 대해, 상이한 이전 화상 쌍들의 데이터로부터 추정된 적어도 2개의 (시간적) 예측 벡터를 제공한다. 양호하게는, 상기 적어도 2개의 예측 벡터들은 화상(부분)에 대한 출력 벡터를

결정하는 벡터 선택 과정에서의 후보들이다. 유익하게는, 기준 함수에 따라, 상기 적어도 2개의 예측 벡터들 중 최상의 것이 화상(부분)에 대한 출력 벡터를 결정하는 벡터 선택 과정의 입력이 되는 후보 벡터들을 계산하기 위한 기초로서 사용된다(예, 갱신 과정). 양호하게는, 다수의 연속된 화상들(화상 부분들)에 대한 결정 정보(적어도 2개의 예측 벡터들 중 어느 것이 기준 함수에 따른 최상의 벡터인지를 가리키는 정보)가 화상 반복 패턴을 검출하기 위해 사용된다(예, 영화 자료의 3-2 폴-다운 및 2-2 폴-다운, 또한 소스-목적지 화상 주파수의 불일치로 인한 다른 패턴도 가능).

본 발명의 제4 특징은 비디오 데이터의 결합 움직임 추정과 분할에 관한 것이며, 아래와 같이 요약될 수 있다. 화상을 고정된 개수의 층들로 분할하고, 개개의 층들에 대해 움직임 파라미터들을 추정하는 방법이 개시된다. 본 발명은 결합된 움직임 추정 및 분할의 선후 문제(the chicken and egg problem)에 대한 솔루션을 제공한다. 이 솔루션은, 주어진 층에 대한 파라미터 추정기의 최적 기준의 훼손을, 병렬로 운영되는 다른 파라미터 추정기들에 의해 다루어진 정보에 의해 제한하는 가중치 처리로 구성된다.

이 알고리즘의 지극히 낮은 복잡도와 높은 품질로 인해, 장래에 TV 및 멀티미디어에 대해 이용될 수 있도록 해준다. 이 알고리즘의 소프트웨어 버전은 필립스 TM1000(TriMedia) 프로세서 상에서 실시간 실행되는 것을 보여준다.

양호한 실시예에서, 유사-동시 움직임 추정/분할을 고정된 최대 개수의 층까지 허용하는 층별 움직임 추정 알고리즘이 제안되었다. 이 추정은, 층당 하나의 움직임 파라미터 세트와 이들 세트들을 화상의 상이한 부분들에 할당하는 분할 맵으로 귀결된다(움직임 층). 층 내에서의 움직임은 팬(pan), 틸트(tilt) 및 줌(zoom)을 기술할 수 있는 최대 4개의 파라미터들로 모델링된다. 이 개념은 어느 정도 계층성, 즉, 움직임 층들의 등급화(ranking)를 보여준다. 이러한 방식으로, 하나의 층에 관한 움직임 파라미터 추정은 계층구조의 상위층에 의해 기술된 화상 부분들을 배제하고, 계층구조의 하위층에 의해 잘 기술된 화상 부분들에 의해 오염되지 않는다. 이 개념은 매우 낮은 동작 개수를 초래한다. 이는 중요한 스캔 속도 변환 용도에서, 특히 화상 속도 상향-변환에서 잘 수행되는 것으로 보였다. 3개 층을 포함하는 한 변형이 필립스 트리메디아(TriMedia) 프로세서 상에서 실시간 수행되도록 스케줄되었다.

양호한 실시예의 하기 두드러진 특징은 주목할 가치가 있다. 화상을 등급화된 세트의 층들로 분할하고, 모든 층들에 대해 움직임 파라미터를 추정하는 방법을 실현하는 방법 및 장치는,

현재 화상 내의 모든 층들에 대하여, 적어도 2개의 화상들로부터의 화소들(화소들의 그룹)에 기초하여 기준 함수를 최적화하는 파라미터 추정(PE) 과정,

움직임 파라미터 세트를 화상 부분들에 할당하는 분할 과정(SP),

상이한 화상 부분들로부터의 정보가 움직임 파라미터 추정기의 기준 함수에 미치는 개별적인 효과를 정의하는 가중치 처리(WP)를 포함하고, 상기 WP는 제1 기준을 충족하는 화상 부분들로부터의 정보의 영향을 감소 또는 제거하고,

제2 기준을 충족하는 화상 부분들로부터의 정보의 효과를 증가시킨다.

양호하게는, 제1 기준은, 동일하거나 다른 화상 쌍에 대한 알고리즘의 선행된 반복에서, 화상 부분들이, 보다 상위층의 액티브 PE들에 의해 추정된 움직임 파라미터 세트들 중 임의의 것에 의해 적절히 기술되었던 영역에 속하는 경우에 만족된다. "적절히"라는 용어는, 계층구조의 상기 보다 상위층들상의 액티브 PE들의 파라미터 세트를 사용하는 에러 함수가 임계값 이하에 머무는 것을 의미한다(예를 들어, 고정되거나 평균 에러에 적응됨).

양호하게는, 제2 기준은, 동일하거나 다른 화상 쌍에 대한 알고리즘의 선행된 반복에서, 화상 부분들이, 이와 같이 주어진 PE에 의해 추정된 움직임 파라미터 세트들에 의해 최상으로 기술되었던 영역에 속하는 경우에 만족된다. "최상"이라는 용어는, 상기 주어진 PE의 파라미터 세트들을 사용하는 에러 함수가 임의의 다른 PE들의 경우보다 더 낮다는 것을 의미한다.

양호하게는, 이러한 에러 함수는, 평가될 파라미터 세트를 사용하여(직접적 방법), 현재 필드의 화소들과 이전 필드에서 대응하는 화소들간의 움직임-보상된 차이에 기초한다.

양호하게는, 이러한 에러 함수는 소정 방법에 의해 계산된 움직임 벡터들간의 차이에 기초한다. 움직임 벡터들은 평가될 움직임 파라미터 세트로부터 초래된다(간접적 방법)

양호하게는, 제1 기준을 만족시키는 화상 부분들은 주어진 PE의 에러 함수에서 제거되고, 이 제1 기준은 기준 함수가 계산되는 화상 영역이 주어진 범위 내에 머물도록 하는 방식으로 적용된다(최대로 이용할 수 있는 처리 전력을 효율적으로 사용하는 제어 루프).

양호하게는, PE 및/또는 SP, 및/또는 WP는 다운-스케일된 및/또는 서브-샘플링된 비디오 데이터 상에서 동작한다.

화상을 등급화된 세트의 층들로 분할하고, 각각의 층에 대한 움직임 파라미터들을 추정하는 방법을 실현하는 방법 및 장치는

현재 화상의 모든 층에 대해, 적어도 2개의 화상들로부터 선택된 화소(화소 그룹)에 기초하여 기준 함수를 최적화하는 반복적 파라미터 추정 과정,

움직임 파라미터 세트들중 하나를 화상의 모든 부분에 할당하는 분할 과정,

적어도 2개의 화상으로부터 화소들(화소 그룹) 중 어느 것이 그들의 기준 함수를 최적화하는지를 정의하는 선택 과정을 포함하고, 여기서, 파라미터 추정 과정은 다른 과정보다 더 자주 그 데이터에 대해 반복된다.

양호하게는, 선택 과정은 주어진 층에 대해 화소들(화소 그룹)을 선택하는 과정으로서, 상기 화소들은, 상기 화소들에 대해 이전 화상의 계층구조의 보다 상위층의 파라미터 세트들이 규칙에 따라 만족할만한 결과를 제공하지 못하였던, 그러한 화소들이다. 양호하게는, 이 규칙은 고정된 또는 적응된 임계값과 화소들(그 그룹)의 에러(합)를 비교하는 것을 포함한다.

양호하게는, 이 임계값은 기준 함수가 계산되는 화소의 개수가 주어진 범위 내에 머물도록 하는 방식으로 적용된다.

양호하게는, 기준 함수는, 후보 움직임 파라미터에 따라, 이전 화상으로부터 선택된 화소들(그의 그룹)과 움직임을 위해 보상된 현재 화상으로부터 대응하는 화소들간에 계산된 합계된 에러이다.

양호하게는, 기준 함수에 대한 상기 선택된 화소들의 기여분은, 이들이 할당받은 (이전 화상) 층에 의존하여 가중된다.

양호하게는, 기준 함수에 대한 상기 선택된 화소들의 기여분은, 이들이 앞서 동일한 층에 할당된 경우에는 증가된다.

화상을 등급화된 세트의 층들로 분할하고, 모든 층에 대한 움직임 파라미터를 추정하는 방법을 실현하는 방법 및 장치는,

현재 화상 내의 모든 층에 대하여, 적어도 2개의 화상으로부터 화소들(화소들의 그룹들)에 기초하여 기준 함수를 최적화하는 파라미터 추정(PE) 과정,

움직임 파라미터 세트중의 하나를 화상의 모든 부분들에 할당하는 분할 과정(SP),

적어도 2개의 화상으로부터의 화소들(화소들의 그룹들) 중 어느 것이 그들의 기준 함수를 최적화하는지를 정의하는 선택 과정을 포함하고, 움직임 파라미터 추정기(들)은 그들의 기준 함수를 최적화해야 하며, 여기서 선택 과정은 분할 과정에 의해 이들 파라미터들을 할당받은 층들의 크기와 무관하게, 적은 양의 화소들만이 PE들에 의해 최적화된 기준 함수에 기여할 수 있도록 허용한다.

원래는 필립스 트리메디아(TriMedia) 프로세서 상에서 실행되는 것으로 설계되었지만, 다른 응용도 가능하다. 특히, 그 개념은 차세대 VGA-제어기들 내에 설계될 수 있다. 이는 전용 실리콘이기 때문에, 전체 단가는 무시할 수 있다. 이러한 VGA-제어기는 트리메디아(TriMedia) 솔루션과 비교해 개선된 성능을 가질 수 있는데, 그 이유는 훨씬 더 많은 처리 전력이 전용 실리콘에 이용될 수 있기 때문이다. 더욱이, 2개 이상의 병렬 파라미터 추정기가 적용되는 경우, 그 성능은 저가의 비용으로 현재의 하이-엔드 솔루션들보다 더 나은 수준이 될 수 있을 것으로 기대된다.

상기 실시예들은 본 발명을 제한하고자 하는 것이 아니라 예시하고자 하는 것으로, 당업자는 첨부된 특허청구의 범위에서 벗어나지 않는 많은 대안적 실시예를 설계할 수 있음에 유의해야 한다. 특허청구의 범위에서, 괄호 안의 참조 부호는 특허청구의 범위를 제한하는 것으로서 해석되어서는 안된다. 본 발명은 수개의 개별적 요소를 포함하는 하드웨어에 의해서, 및

적절히 프로그램된 컴퓨터에 의해 구현될 수 있다. 수개의 수단을 열거한 장치에서, 이들 수개의 수단들은 하나의 동일한 하드웨어의 항목에 의해 구현될 수 있다. 특허청구의 범위에서 "포함하는"이라는 표현은 특허청구의 범위에 열거된 것 이외의 다른 요소들이나 단계들의 존재를 배제하지 않는다.

참고자료:

- [1] US-A-5,534,946 (Attorneys' docket PHN 14,066)
- [2] A.M. Tekalp, "Digital Video Processing", Prentice Hall Signal Processing Series, ISBN 0-13190075-7, pp. 200-203.
- [3] G. de Haan, P.W.A.C. Biezen, H. Huijgen and O.A. Ojo, "True Motion Estimation with 3-D Recursive Search Block-Matching", IEEE Transactions on Circuits and Systems for Video Technology, Vol.3, October 1993, pp. 368-388.
- [4] G. de Haan, P.W.A.C Biezen, H. Huijgen, and O.A. Ojo, "Graceful Degradation in Motion Compensated Field-Rate Conversion", in: Signal Processing of HDTV, V, L. Stenger, L. Chiariglione and M. Akgun (Eds.), Elsevier 1994, pp. 249-256.
- [5] PCT/IB97/00548 (Attorneys' docket PHN 16,112)
- [6] G. de Haan, J. Kettenis, and B. Déloore, 'IC for Motion Compensated 100 Hz TV, with a Smooth Motion Movie-Mode', International Conference on Consumer Electronics, ICCE 95, June 1995, Chicago.
- [7] G. de Haan, P.W.A.C Biezen, "Sub-pixel motion estimation with 3-D recursive search block-matching", Signal Processing: Image Communication 6 (1994), pp. 229-239.
- [8] PCT/IB97/00884 (Attorneys' docket PHN 15,943)
- [9] US-A-5,495,300 (Attorneys' docket PHN 14,079)
- [10] G. Thomas, "Television motion measurement for DATV and other applications," BBC Research Report, no. BBC RD 1987/11, 1987.
- [11] R. Thoma and M. Bierling, "Motion compensating interpolation considering covered and uncovered background," Signal Processing: Image Communications 1, pp. 191-212, 1989.
- [12] F. Wang, D. Anastassiou, and A. Netravali, "Time-recursive deinterlacing for IDTV and pyramid coding," Signal Processing: Image Communications 2, pp. 365-374, 1990.
- [13] Kwon, Seo, Kim, and Kim, "A motion adaptive deinterlacing method," IEEE Transactions on Consumer Electronics, vol. 38, pp. 145-150, August 1992.
- [14] G. de Haan and H. Huijgen, "New algorithm for motion estimation," in Chiariglione [38], pp. 109-116.
- [15] G. de Haan and H. Huijgen, "Motion estimation for TV picture enhancement,"

- in *Signal Processing of HDTV III* (H. Yasuda and L. Chiariglione, eds.), pp. 241-248, Elsevier Science Publishers B.V., 1992.
- [16] T. Reuter, "A modified block-matching algorithm with vector reliability checking and adaptive smoothing," in *Third International Conference on Image Processing and its Applications*, (England), University of Warwick, July 1989.
- [17] J. Konrad and E. Dubois, "A comparison of stochastic and deterministic solution methods in bayesian estimation of 2-d motion," *Image and Vision Computing*, vol. 8, pp. 304-317, November 1990.
- [18] J. Jain and A. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Transactions on Communications*, COM-29, no. 12, 1981.
- [19] T. Koga, K. Inuma, A. Hirano, Y. Iilima and T. Ishiguro, "Motion-compensated interframe coding for video conferencing," in *IEEE, Proceedings of the NTC 81, G5.3.1.*, (New Orleans LA), 1981.
- [20] R. Srinivasan and K. Rao, "Predictive coding based on efficient motion estimation," *IEEE Transactions on Communication*, no. 8, pp. 888-896, 1985.
- [21] H. Musmann, P. Pirsch, and J. Grallert, "Advances in picture coding," *Proceedings of the IEEE*, vol. 73, pp. 523-548, April 1985.
- [22] A. Netravali and J. Robbins, "Motion compensated television coding," *Bell Systems Technical Journal*, no. 3, pp. 629-668, 1979.
- [23] M. Ziegler, "Hierarchical motion estimation using the phase correlation method in 140 Mbit/s HDTV-coding," in Chiariglione [32], pp. 131-137.
- [24] DE-C 40 23 449
- [25] G. de Haan, J. Kettenis, and B. Deloore, "IC for motion compensated 100 Hz TV, with a smooth motion movie-mode," *IEEE Transactions on Consumer Electronics*, vol. 42, pp. 165-174, May 1996.
- [26] J. G. Choi and S.-D. Kim, "Multi-stage segmentation of optical flow field," *Signal Processing*, vol. 54, pp. 109-118, 1996.
- [27] D. Bagni, R. Lancini, P. Vicari, and S. Tubaro, "Motion estimation method using region-based segmentation methods," in *Proc. International Workshop on HDTV '96*, (Los Angeles), p. Sess. A2, October 1996.
- [28] D. LeQuang, Z. Zaccarin, and S. Caron, "Object-oriented coding using successive motion field segmentation and estimation," in *Proc. International Conference on Image Processing (ICIP'95)*, (Washington D.C.), pp. 207-210,

- October 1995.
- [29] J. Y. A. Wang and E. H. Adelson, "Layered representation for motion analysis," in Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, pp. 361-366, 1993.
 - [30] P. Csillag and L. Boroczky, "Frame rate conversion based on acceleration and motion-based segmentation," in SPIE, vol. 2952, pp. 438-448, 1996.
 - [31] F. Dufaux and F. Moscheni, "Motion estimation techniques for digital tv: a review and a new contribution," in Proceeding of the IEEE, vol. 83 n.6, pp. 858-876, 1995.
 - [32] S. Jeannin, "On the combination of a polynomial motion estimation with a hierarchical segmentation based video coding scheme," in Proc. International Conference on Image Processing (ICIP'96) (Lausanne, Switzerland), pp. 489-492, September 1996.
 - [33] K. E. Matthews and N. M. Namazi, "Simultaneous motion parameter estimation and image segmentation using the EM algorithm," in Proc. International Conference on Image Processing (ICIP'95), (Washington D.C.), pp. 542-545, October 1995.
 - [34] T. Darrel and D. Fleet, "Second-order method for occlusion relationships in motion layers," Tech. Rep. 314, MIT Media Laboratory Vision and Modelling Group, 1995.
 - [35] H. S. Sawhney, S. Ayer, and M. Gorkani, "Model-based 2D-3D dominant motion estimation for mosaicing video representation." On the net, 1995. A shorter version appeared in the IEEE Intl. Conf. on Computer Vision, Cambridge, MA, USA, June 1995.
 - [36] H. S. Sawhney and S. Ayer, "Layered representation of motion video using robust maximum-likelihood estimation of mixture models and MDL encoding." On the net, 1995. A shorter version appeared in the IEEE Intl. Conf. on Computer Vision, Cambridge, MA, USA, June 1995.
 - [37] J. Y. A. Wang and E. H. Adelson, "Spatio-temporal segmentation of video data," in Proceedings of the SPIE: Image and Video Processing II, vol. 2182, (San Jose), pp. 361-366, Feb 1994.
 - [38] L. Chiariglione, ed., Signal Processing of HDTV II, Elseviers Science Publishers B.V.. 1990.

(57) 청구의 범위

청구항 1.

비디오 데이터에서 움직임을 추정하는 방법으로서,

입력 비디오 데이터($n, n-1$)로부터 적어도 2개의 움직임 파라미터 세트들을 제공하는 단계(PE1-PE n)로서, 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터들이 계산될 수 있으며, 움직임 파라미터 세트 각각은 화상 내의 모든 화상 부분들에 대한 대응하는 국부적 정합 에러들을 가지며, 상기 적어도 2개의 움직임 파라미터 세트들은, 화상 내의 모든 화상 부분들에 대한 제로 속도를 나타내는 하나의 움직임 파라미터 세트를 포함하는, 상기 적어도 2개의 움직임 파라미터 세트들을 제공하는 단계; 및

상기 적어도 2개의 움직임 파라미터 세트들에 의존하여, 상기 입력 비디오 데이터($n, n-1$)로부터 출력 움직임 데이터(P(n))를 결정하는 단계(MPE)로서, 상기 출력 움직임 데이터(P(n))를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요도(가중치 인자들(W)에 의해 결정됨)는 상기 움직임 파라미터 세트들의 국부적인 정합 에러들에 의존하는, 상기 출력 움직임 데이터(P(n))를 결정하는 단계(MPE)를 포함하는, 움직임 추정 방법.

청구항 2.

제1항에 있어서, 상기 움직임 파라미터 세트들을 제공하는 단계는,

적어도 하나의 앞서 계산된 움직임 파라미터 세트를 적어도 하나의 제1 예측 파라미터 벡터(Cp1)로서 제공하는 단계;

추가 예측 파라미터 벡터(Cp2)를 얻기 위해, 상기 적어도 하나의 제1 예측 파라미터 벡터에 적어도 하나의 갱신 파라미터 벡터(U_p)를 추가하는 단계(AD); 및

상기 제1 예측 파라미터 벡터 및 상기 추가 예측 파라미터 벡터(Cp1, Cp2) 중 하나를 선택하는 단계(PVC, EC, MIN, MUX)를 포함하는, 움직임 추정 방법.

청구항 3.

제2항에 있어서, 상기 제공 단계는, 적어도 2개의 이전 필드들에 대해 계산된 움직임 파라미터 세트들(Tp1-Tpn)을 제공하는 단계를 포함하는, 움직임 추정 방법.

청구항 4.

제1항에 있어서, 상기 움직임 파라미터 세트들을 제공하는 단계는, 조절된 움직임 파라미터 세트를 제로 속도 세트 이외의 움직임 파라미터 세트 각각에 대해 전역적 정합 에러들(global match error)에 의존하여 결정하는 단계를 포함하고, 상기 전역적 정합 에러들은 화상 부분들에 할당된 가중치들에 따라 상기 화상 부분들에 대해 계산되는, 움직임 추정 방법.

청구항 5.

제4항에 있어서, 주어진 조절된 움직임 파라미터 세트 결정(PE_i)에서, 상기 주어진 움직임 파라미터 세트가 다른 조절된 움직임 파라미터 세트 결정들(PE1-PE_n)의 국부적 정합 에러들과 비교하여 최저 국부적 정합 에러를 갖는 화상 부분들에 대해서 상기 가중치들이 증가되는, 움직임 추정 방법.

청구항 6.

제4항에 있어서, 주어진 조절된 움직임 파라미터 세트 결정(PE_i)에서, 또 다른 조절된 움직임 파라미터 세트 결정들(PE1-PE_n)의 국부적 정합 에러가 주어진 임계값 아래로 떨어지는 화상 부분들에 대해서 상기 가중치들이 감소되는, 움직임 추정 방법.

청구항 7.

비디오 데이터를 움직임-보상하는 방법으로서,

입력 비디오 데이터(n, n-1)로부터 적어도 2개의 움직임 파라미터 세트들을 제공하는 단계(PE1-PE_n)로서, 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터들이 계산될 수 있으며, 움직임 파라미터 세트 각각은 화상 내의 모든 화상 부분들에 대한 대응하는 국부적 정합 에러들을 가지며, 상기 적어도 2개의 움직임 파라미터 세트들은, 화상 내의 모든 화상 부분들에 대한 제로 속도를 나타내는 하나의 움직임 파라미터 세트를 포함하는, 상기 적어도 2개의 움직임 파라미터 세트들을 제공하는 단계; 및

상기 적어도 2개의 움직임 파라미터 세트들에 의존하여, 상기 입력 비디오 데이터($n, n-1$)로부터 출력 비디오 데이터($n-\frac{1}{2}$)를 보간하는 단계(UC2)로서, 상기 출력 비디오 데이터($n-\frac{1}{2}$)를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요도(가중치 인자들(W)에 의해 결정됨)는 상기 움직임 파라미터 세트들의 국부적인 정합 에러들에 의존하는, 상기 출력 비디오 데이터($n-\frac{1}{2}$)를 보간하는 단계(UC2)를 포함하는, 비디오 데이터의 움직임-보상 방법.

청구항 8.

제7항에 있어서, 상기 보간 단계는,

움직임-보상된 평균(mcav)을 계산하는데 사용된 상기 움직임 벡터(D)의 정합 에러가 상기 제로 움직임 벡터의 정합 에러를 초과하는 경우, 이전 필드($n-1$)에서의 대응하는 화소(l), 다음 필드(n)에서의 대응하는 화소(r), 및 상기 이전 필드($n-1$)와 다음 필드(n) 양자 모두로부터의 상기 움직임-보상된 평균(mcav) 중 중앙값(MED2), 또는

상기 움직임-보상된 화소들을 제공하는데 사용된 상기 움직임 벡터(D)의 정합 에러가 상기 제로 움직임 벡터의 정합 에러 아래로 떨어지는 경우, 상기 이전 필드($n-1$)로부터의 움직임-보상된 화소(mcl), 상기 다음 필드(n)로부터 움직임-보상된 화소(mcr), 및 상기 이전 필드($n-1$)와 상기 다음 필드(n) 양자 모두로부터의 움직임-보상되지 않은 평균(av) 중 중앙값(MED1)을 제공하는, 움직임-보상 방법.

청구항 9.

비디오 데이터에서 움직임을 추정하기 위한 장치로서,

입력 비디오 데이터($n, n-1$)로부터 적어도 2개의 움직임 파라미터 세트들을 제공하기 위한 수단(PE1-PE n)으로서, 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터들이 계산될 수 있으며, 움직임 파라미터 세트 각각은 화상 내의 모든 화상 부분들에 대한 대응하는 국부적 정합 에러들을 가지며, 상기 적어도 2개의 움직임 파라미터 세트들은, 화상 내의 모든 화상 부분들에 대한 제로 속도를 나타내는 하나의 움직임 파라미터 세트를 포함하는, 상기 적어도 2개의 움직임 파라미터 세트들을 제공하기 위한 수단(PE1-PE n); 및

상기 적어도 2개의 움직임 파라미터 세트들에 의존하여, 상기 입력 비디오 데이터($n, n-1$)로부터 출력 움직임 데이터(P(n))를 결정하기 위한 수단(MPE)으로서, 상기 출력 움직임 데이터(P(n))를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요도(가중치 인자들(W)에 의해 결정됨)는 상기 움직임 파라미터 세트들의 국부적인 정합 에러들에 의존하는, 상기 출력 움직임 데이터(P(n))를 결정하기 위한 수단(MPE)을 포함하는, 움직임 추정 장치.

청구항 10.

비디오 데이터를 움직임-보상하기 위한 디바이스로서,

입력 비디오 데이터($n, n-1$)로부터 적어도 2개의 움직임 파라미터 세트들을 제공하기 위한 수단(PE1-PE n)으로서, 움직임 파라미터 세트는 화상 내의 움직임을 기술하는 파라미터들의 세트이고, 상기 움직임 파라미터 세트에 의해 움직임 벡터들이 계산될 수 있으며, 움직임 파라미터 세트 각각은 화상 내의 모든 화상 부분들에 대한 대응하는 국부적 정합 에러들을 가지며, 상기 적어도 2개의 움직임 파라미터 세트들은, 화상 내의 모든 화상 부분들에 대한 제로 속도를 나타내는 하나의 움직임 파라미터 세트를 포함하는, 상기 적어도 2개의 움직임 파라미터 세트들을 제공하기 위한 수단(PE1-PE n); 및

상기 적어도 2개의 움직임 파라미터 세트들에 의존하여, 상기 입력 비디오 데이터($n, n-1$)로부터 출력 비디오 데이터($n-\frac{1}{2}$)를 보간하기 위한 수단(UC2)으로서, 상기 출력 비디오 데이터($n-\frac{1}{2}$)를 계산하는데 있어서 움직임 파라미터 세트 각각의 중요도(가중치 인자들(W)에 의해 결정됨)는 상기 움직임 파라미터 세트들의 국부적인 정합 에러들에 의존하는, 상기 출력 비디오 데이터($n-\frac{1}{2}$)를 보간하기 위한 수단(UC2)을 포함하는, 비디오 데이터의 움직임-보상 장치.

청구항 11.

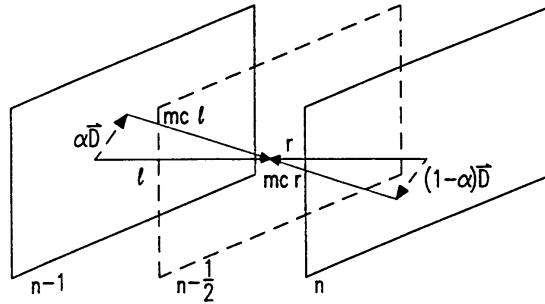
비디오 디스플레이 장치로서,

제10항에 청구된 비디오 데이터의 움직임-보상을 위한 디바이스; 및

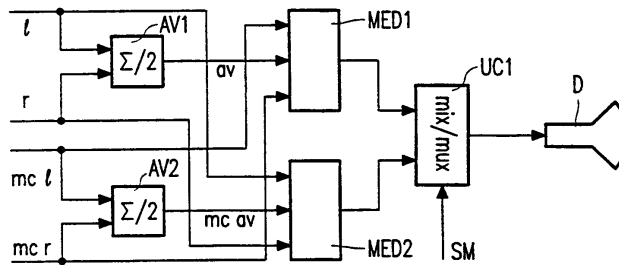
입력 비디오 데이터($n, n-1$) 및 출력 비디오 데이터($n-\frac{1}{2}$)를 디스플레이하기 위한 디스플레이 유닛(D)을 포함하는, 비디오 디스플레이 장치.

도면

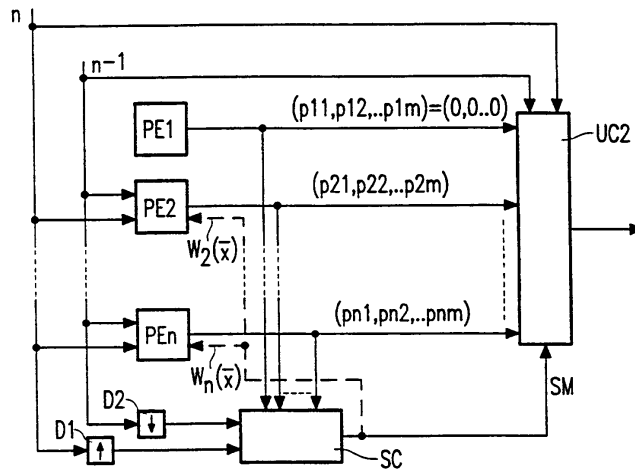
도면1



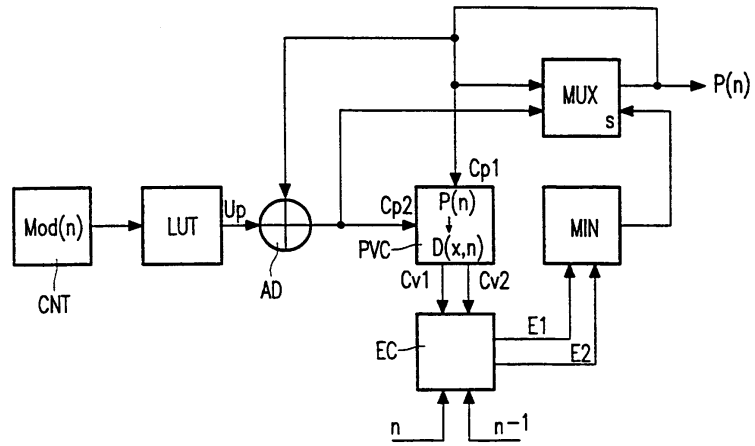
도면2



도면3



도면4



도면5

