



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0084460
(43) 공개일자 2009년08월05일

(51) Int. Cl.

G06T 7/40 (2006.01) G06T 7/00 (2006.01)

(21) 출원번호 10-2008-0010667

(22) 출원일자 2008년02월01일

심사청구일자 없음

(71) 출원인

삼성전자주식회사

경기도 수원시 영통구 매탄동 416

고려대학교 산학협력단

서울 성북구 안암동5가1 고려대학교 내

(72) 발명자

박영오

경기도 성남시 분당구 정자동 느티마을주공4단지
아파트 404동802호

송관용

서울특별시 도봉구 쌍문동 한양아파트1차 6동 40
7호

(뒷면에 계속)

(74) 대리인

이건주

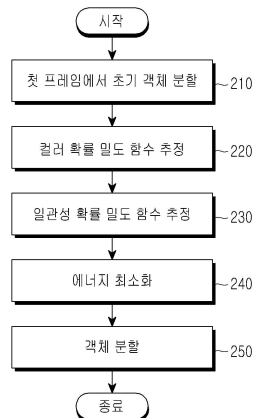
전체 청구항 수 : 총 15 항

(54) 동영상의 객체 경계 추정 방법

(57) 요약

본 발명은 동영상에서 배경과 사용자 객체의 경계 추정 방법에 있어서, 객체 분할을 위한 동영상의 첫 프레임에서 초기 객체 경계를 추정하여 객체와 배경을 분할하는 과정과, 첫 프레임에서 분할된 객체와 배경의 컬러(color) 확률 밀도 함수를 각각 추정하는 과정과, 현재 프레임에서, 이전 프레임의 객체 경계를 따라 상기 이전 프레임의 객체 경계를 중심으로 미리 설정된 일정한 너비를 가지는 경계 스트립(strip)을 생성하여 상기 경계 스트립의 일관성(coherence) 확률 밀도 함수를 추정하는 과정과, 에너지 최소화(Energy Minimization)를 수행하여 현재 프레임의 객체를 분할하는 과정을 포함한다.

대표도 - 도2



(72) 발명자

주영훈

경기도 용인시 기흥읍 영덕리 태영아파트 206동
1403호

최광표

경기도 안양시 만안구 박달2동 우성아파트 103-705

오윤제

경기도 용인시 기흥구 마북동 삼성래미안1차아파트
107동 1801호

김창수

서울특별시 서초구 양재동 우성아파트 111동 1401
호

안재균

서울특별시 성북구 안암동5가 156번지 정진고시원
2450호

특허청구의 범위

청구항 1

동영상에서 배경과 사용자 객체의 경계 추정 방법에 있어서,

객체 분할을 위한 동영상의 첫 프레임에서 초기 객체 경계를 추정하여 객체와 배경을 분할하는 과정과,

상기 첫 프레임에서 분할된 상기 객체와 상기 배경의 컬러(color) 확률 밀도 함수를 각각 추정하는 과정과,

현재 프레임에서, 이전 프레임의 객체 경계를 따라, 상기 이전 프레임의 객체 경계를 중심으로 미리 설정된 일정한 너비를 가지는 경계 스트립(strip)을 생성하여 상기 경계 스트립의 일관성(coherence) 확률 밀도 함수를 추정하는 과정과,

에너지 최소화(Energy Minimization)를 수행하여 현재 프레임의 객체를 분할하는 과정을 포함함을 특징으로 하는 객체 경계 추정 방법.

청구항 2

제 1항에 있어서, 상기 초기 객체 경계 추정은 사용자의 상호 작용을 이용하는 그랩컷(Grabcut) 알고리즘을 사용하여 수행하는 것을 특징으로 하는 객체 경계 추정 방법.

청구항 3

제 1항에 있어서, 상기 컬러 확률 밀도 함수는 하기의 수학적 식 10과 같음을 특징으로 하는 객체 경계 추정 방법.

[수학적 식 10]

$$p_{\text{color}}(\mathbf{x}|M_{\alpha}) = \frac{1}{h^d |M_{\alpha}|} \sum_{\mathbf{p} \in M_{\alpha}} K\left(\frac{\mathbf{x} - \mathbf{x}_{\mathbf{p}}}{h}\right), \quad \alpha = 0, 1$$

상기 수학적 식 10에서 상기 M_{α} 는 픽셀의 집합이며, 상기 α 는 배경일 경우 0이고 객체일 경우 1이 되어, 상기 M_1 은 객체의 픽셀 집합, 상기 M_0 은 배경의 픽셀 집합임, 상기 $\mathbf{x}_{\mathbf{p}}$ 는 픽셀 $\mathbf{p}$ 의 컬러(R, G, B)를 표시하는 삼차원 벡터이고, 상기 K 는 커널 함수(kernel function)이고, 상기 h 는 커널의 대역폭이고, 상기 d 는 $\mathbf{x}$ 의 차원(dimension)이고, 상기 $|M_{\alpha}|$ 는 상기 M_{α} 내의 픽셀의 수임.

청구항 4

제 3항에 있어서, 상기 커널은 다변수(multivariate)의 에파네치니코프 커널(Epanechnikov kernel)이며, 하기의 수학적 식 11과 같음을 특징으로 하는 객체 경계 추정 방법.

[수학적 식 11]

$$K(\mathbf{x}) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \mathbf{x}^t \mathbf{x}) & \text{if } \mathbf{x}^t \mathbf{x} < 1 \\ 0 & \text{otherwise} \end{cases}$$

상기 수학적 식 11에서 상기 C_d 는 d -차원 영역의 부피($C_1=2$, $C_2=\pi$, $C_3=4\pi/3$)이고, 상기 d 는 $\mathbf{x}$ 의 차원임.

청구항 5

제 1항에 있어서, 상기 컬러 확률 밀도 함수를 추정한 후 미리 설정된 룩업 테이블(Look-up table)에 저장하는 것을 특징으로 하는 객체 경계 추정 방법.

청구항 6

제 1항에 있어서, 상기 경계 스트립의 너비는, 상기 이전 프레임에서 추정된 객체의 경계에 겹치는 블록들의 음

직접 벡터(motion vector) 값들의 중간(median) 움직임 벡터를 이용하여 계산하며, 하기의 수학식 12와 같음을 특징으로 하는 객체 경계 추정 방법.

[수학식 12]

$$w = w_{\min} + \|\mathbf{v}\|$$

상기 수학식 12에서 상기 $w_{\min}$ 은 상기 경계 스트립의 너비의 최소값이고, 상기 v 는 객체의 경계에 접치는 블록들의 중간 움직임 벡터임.

청구항 7

제 6항에 있어서, 각 블록의 움직임 벡터는 'Three-step search' 알고리즘을 이용하여 구하는 것을 특징으로 하는 객체 경계 추정 방법.

청구항 8

제 6항에 있어서, 상기 일관성 확률 밀도 함수는 하기의 수학식 13과 같음을 특징으로 하는 객체 경계 추정 방법.

[수학식 13]

$$p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1}) = \begin{cases} 0.5 + \frac{d_{\mathbf{p}}}{w} & \text{if } \alpha_{\mathbf{p}}^{t-1} = 1 \\ 0.5 - \frac{d_{\mathbf{p}}}{w} & \text{if } \alpha_{\mathbf{p}}^{t-1} = 0 \end{cases}$$

상기 수학식 13에서 α^t 는 프레임 t 의 이미지의 라벨로서 픽셀 P 의 $\alpha_{\mathbf{p}}^t$ 는 P 가 객체일 경우 1이고 P 가 배경일 경우 0임, 상기 w 는 상기 경계 스트립의 너비이고, 상기 $d_{\mathbf{p}}$ 는 이전 프레임($t-1$)의 객체 경계로부터 픽셀 P 까지의 최단거리로서 $0 \leq d_{\mathbf{p}} \leq \frac{w}{2}$ 임.

청구항 9

제 6항에 있어서, 상기 일관성 확률 밀도 함수는, 현재 프레임의 상기 중간 움직임 벡터를 이용하여 현재 프레임의 경계 스트립을 이동시켜서 계산하며, 하기의 수학식 14와 같음을 특징으로 하는 객체 경계 추정 방법.

[수학식 14]

$$p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1}) = \begin{cases} 0.5 + \frac{d_{\mathbf{p}-\mathbf{v}}}{w} & \text{if } \alpha_{\mathbf{p}-\mathbf{v}}^{t-1} = 1 \\ 0.5 - \frac{d_{\mathbf{p}-\mathbf{v}}}{w} & \text{if } \alpha_{\mathbf{p}-\mathbf{v}}^{t-1} = 0 \end{cases}$$

상기 수학식 14에서 α^t 는 프레임 t 의 이미지의 라벨로서 픽셀 P 의 $\alpha_{\mathbf{p}}^t$ 는 P 가 객체일 경우 1이고 P 가 배경일 경우 0임, 상기 w 는 상기 경계 스트립의 너비이고, 상기 $d_{\mathbf{p}}$ 는 이전 프레임($t-1$)의 객체 경계로부터 픽셀 P 까지의 최단거리로서 $0 \leq d_{\mathbf{p}} \leq \frac{w}{2}$ 이고, 상기 v 는 객체의 경계가 접치는 블록들의 중간 움직임 벡터임.

청구항 10

제 1항에 있어서, 상기 에너지 최소화는 컬러 에너지 함수와 일관성 에너지 함수와 평탄(smoothness) 에너지 함수를 이용하여, 하기의 수학식 15와 같이 계산하는 것을 특징으로 하는 객체 경계 추정 방법.

[수학식 15]

$$E(\alpha^t) = E_{\text{color}}(\alpha^t) + E_{\text{cohere}}(\alpha^t) + E_{\text{smooth}}(\alpha^t)$$

상기 수학식 15에서 $E_{\text{color}}(\alpha^t)$ 는 컬러 에너지 함수이고, 상기 $E_{\text{cohere}}(\alpha^t)$ 는 일관성 에너지 함수이고, 상기 $E_{\text{smooth}}(\alpha^t)$ 평탄 에너지 함수임.

청구항 11

제 10항에 있어서, 상기 컬러 에너지 함수는 하기의 수학식 16과 같이 계산하는 것을 특징으로 하는 객체 경계 추정 방법.

[수학식 16]

$$E_{\text{color}}(\alpha^t) = - \sum_{\mathbf{p} \in \mathcal{S}} \log p_{\text{color}}(\mathbf{x}_{\mathbf{p}} | M_{\alpha_{\mathbf{p}}^t})$$

상기 수학식 16에서 $\mathcal{S}$ 는 결합 스트립의 픽셀 집합이고, 상기 $\mathbf{x}_{\mathbf{p}}$ 는 픽셀 $\mathbf{p}$ 의 컬러를 표시하는 삼차원 벡터이고, 상기 M_{α} 는 픽셀의 집합이고, 상기 α^t 는 프레임 t 의 이미지의 레이블로서 픽셀 $\mathbf{p}$ 의 $\alpha_{\mathbf{p}}^t$ 는 $\mathbf{p}$ 가 객체일 경우 1이고 $\mathbf{p}$ 가 배경일 경우 0임.

청구항 12

제 10항에 있어서, 상기 일관성 에너지 함수는 하기의 수학식 17과 같이 계산하는 것을 특징으로 하는 객체 경계 추정 방법.

[수학식 17]

$$E_{\text{cohere}}(\alpha^t) = - \sum_{\mathbf{p} \in \mathcal{S}} \log p_{\text{cohere}}(\alpha_{\mathbf{p}}^t | \alpha^{t-1})$$

상기 수학식 17에서 $\mathcal{S}$ 는 결합 스트립의 픽셀 집합이고, 상기 α^t 는 프레임 t 의 이미지의 레이블로서 픽셀 $\mathbf{p}$ 의 $\alpha_{\mathbf{p}}^t$ 는 $\mathbf{p}$ 가 객체일 경우 1이고 $\mathbf{p}$ 가 배경일 경우 0임.

청구항 13

제 10항에 있어서, 상기 평탄 에너지 함수는 하기의 수학식 18과 같이 계산하는 것을 특징으로 하는 객체 경계 추정 방법.

[수학식 18]

$$E_{\text{smooth}}(\alpha^t) = \sum_{\substack{(\mathbf{p}, \mathbf{q}) \in \mathcal{N} \\ \alpha_{\mathbf{p}} \neq \alpha_{\mathbf{q}}}} \frac{1}{\|\mathbf{p} - \mathbf{q}\|} \exp \left(- \frac{(z_{\mathbf{p}} - z_{\mathbf{q}})^2}{2\sigma^2} \right)$$

상기 수학식 18에서 상기 α^t 는 프레임 t 의 이미지의 레이블로서 픽셀 $\mathbf{p}$ 의 $\alpha_{\mathbf{p}}^t$ 는 $\mathbf{p}$ 가 객체일 경우 1이고 $\mathbf{p}$ 가 배경일 경우 0임, 상기 $\mathcal{N}$ 은 8방향의 인접한 픽셀의 쌍의 집합이고, $z_{\mathbf{p}}$ 는 컬러 $X_{\mathbf{p}}$ 로부터 계산된 픽셀 $\mathbf{p}$ 의 강도(픽셀 데이터의 크기)이고, $z_{\mathbf{q}}$ 는 컬러 $X_{\mathbf{q}}$ 로부터 계산된 픽셀 $\mathbf{q}$ 의 강도이고, σ 는 상기 강도의 표준 편차임.

청구항 14

제 1항에 있어서, 상기 현재 프레임의 객체 분할 후, 상기 경계 스트립에 포함된 픽셀을, 배경의 컬러 확률 밀도 함수와 객체의 컬러 확률 밀도 함수 둘 다와 비교하여, 상기 비교 결과 0의 확률을 가진 픽셀이 전체 픽셀의

7퍼센트(%) 이상이면, 컬러 확률 밀도 함수를 업데이트 하는 것을 특징으로 하는 객체 경계 추정 방법.

청구항 15

제 14항에 있어서, 상기 컬러 확률 밀도 함수를 업데이트하는 경우, 과거 세 프레임의 컬러를 참조하는 것을 특징으로 하는 객체 경계 추정 방법.

명세서

발명의 상세한 설명

기술 분야

- <1> 본 발명은 휴대 단말을 이용한 모바일 환경에서 화상 통신 및 영상 통화시 동영상에서 배경과 사용자 객체를 구분하기 위한 객체 경계 추정 방법에 관한 것이다.

배경 기술

- <2> 최근에 동영상의 압축 및 전송 기술의 발달로 인하여 화상 통신 또는 영상 통화가 대중화되었다. 하지만 화상 통신이 갖는 역기능에 대한 기술적 개선은 미미한 상태이다. 대표적 역기능 중 하나는 사용자가 있는 장소가 노출되어, 사용자의 사생활이 침해된다는 것이다. 이러한 문제점을 극복하고 화상통신을 원활하게 이용하기 위해서 화상 통신의 객체와 배경을 구분하여 배경을 따로 처리하는 기술이 필요하다. 객체와 배경을 구분하면, 배경을 제어하여 다른 배경으로 대체함으로써 사용자의 위치를 감출 수 있다.
- <3> 현재까지 동영상의 객체와 배경을 구분하는 방법으로 다양한 기법들이 제시되었다. 가장 대표적인 방법으로 배경을 모델링하는 방법을 들 수 있다. 이 방법은 고정된 카메라를 사용하는 경우 일정 시간 동안 변화가 없는 화소를 배경으로 설정하는 방법이다. 지정된 배경을 기준으로 화소 값이 급격히 변하는 화소를 객체로 인식하여 객체와 배경을 구분하는 방법이다. 경우에 따라 배경이 바뀔 수 있으므로 특정 시간마다 배경을 갱신하는 방법으로 변화에 대한 민감도를 줄인다.
- <4> 상기 설명한 기술은 배경을 모델링하여 객체와 배경을 구분하는 방법이다. 대체로 배경에 대한 앞, 뒤 프레임은 유사하므로 이 방식을 사용하면 배경을 효과적으로 모델링할 수 있다. 또한 모델링된 배경을 이용하여, 빠른 속도로 객체를 구분할 수 있다. 즉 배경 모델링은 배경이 아닌 모든 변화를 객체로 인식하는 방법이다. 대표적 방법인 가우시안 혼합 모형(Gaussian mixture model)이나 커널 밀도 추정(kernel density estimation) 등을 이용한 배경 모델링은 순간적 또는 지속적인 변화에 적절하게 대응하도록 배경을 갱신한다.
- <5> 상기 설명한 종래 기술의 가장 큰 문제점은 카메라가 고정되어 있어야 한다는 것이다. 시간적으로 인접한 프레임에 대해 각 화소별 모델링이 이루어지므로, 화소 위치가 변화할 경우에는 오류가 발생한다. 심지어 카메라의 위치가 한 화소만큼만 이동해도 엉뚱한 정보가 객체로 인식되는 문제점이 있다. 또한 장치의 연산량과 저장 용량도 결과에 크게 영향을 미칠 수 있다. 이런 제한 때문에, 기존의 방법은 감시 카메라 및 컴퓨터 화상 통신으로만 사용이 국한되었다. 이에 차세대 통신기술인 휴대 단말의 화상 통신에 적용할 수 있는 객체 분할 방법을 요구된다.

발명의 내용

해결 하고자하는 과제

- <6> 본 발명은 휴대 단말을 이용한 화상 통신 및 영상 통화시 배경과 사용자 객체의 경계를 추정할 수 있는 방법을 제공하고자 한다.

과제 해결수단

- <7> 이를 달성하기 위한 본 발명의 일 형태에 따르면, 동영상에서 배경과 사용자 객체의 경계 추정 방법에 있어서, 객체 분할을 위한 동영상의 첫 프레임에서 초기 객체 경계를 추정하여 객체와 배경을 분할하는 과정과, 상기 첫 프레임에서 분할된 상기 객체와 상기 배경의 컬러(color) 확률 밀도 함수를 각각 추정하는 과정과, 현재 프레임에서, 이전 프레임의 객체 경계를 따라 상기 이전 프레임의 객체 경계를 중심으로 미리 설정된 일정한 너비를 가지는 경계 스트립(strip)을 생성하여 상기 경계 스트립의 일관성(coherence) 확률 밀도 함수를 추정하는 과정

과, 에너지 최소화(Energy Minimization)를 수행하여 현재 프레임의 객체를 분할하는 과정을 포함함을 특징으로 한다.

효 과

<8> 본 발명은 화상 통신시 동영상에서 사용자 객체와 배경의 경계를 추정하여 배경과 객체를 구분하는 방법을 제공한다. 본 발명은 복잡한 배경이나 이동으로 인한 배경의 변화에 대해 민감하게 반응하지 않으며, 객체의 움직임에 대해서도 효과적으로 대응하여 객체와 배경의 구분이 가능하다. 또한 객체 경계 추정의 연산이 간단하여 실시간으로 객체와 배경을 분할할 수 있는 효과가 있다.

발명의 실시를 위한 구체적인 내용

<9> 이하 첨부된 도면을 참조하여 본 발명을 구성하는 장치 및 동작 방법을 본 발명의 실시 예를 참조하여 상세히 설명한다. 하기 설명에서는 구체적인 구성 소자 등과 같은 특정 사항들이 나타나고 있는데 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐 이러한 특정 사항들이 본 발명의 범위 내에서 소정의 변형이나 혹은 변경이 이루어질 수 있음은 이 기술분야에서 통상의 지식을 가진 자에게는 자명하다 할 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다.

<10> 본 발명은 휴대용 단말에서 사용하는 화상 통신이나 영상 통화와 같이 배경이 고정되지 않는 동영상으로부터 배경과 사용자 객체의 경계를 추정하여 객체와 배경을 분할할 수 있는 실시간 동영상 객체 경계 추정 방법을 제안한다. 본 발명에서 객체는 화상 통신이나 영상 통화를 사용하는 사용자를 의미한다. 본 발명은 먼저, 첫 번째 프레임에서 초기 객체 경계를 분할하고, 분할된 객체와 배경의 컬러에 근거하여 확률 밀도 함수를 추정한다. 다음, 이후 프레임에서 시공간적(spatio-temporal) 상호관계(correlation)를 이용하기 위해 사용자 객체의 경계를 포함하는 경계 스트립(strip)을 설정한다. 마지막으로, 컬러와 경계 스트립, 평활도(smoothness)를 더해 에너지 함수 최소화(Minimizing)를 수행하여 객체와 배경의 분할을 수행한다. 도면을 참조하여 상세히 설명하기로 한다.

<11> 도 1은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 장치의 블록 구성도이다. 본 발명의 일 실시 예에 따른 객체 경계 추정 장치는 컬러 모델링부, 시공간 모델링부, 객체 분할부를 포함하여 구성되며, 도 1에 도시한 바와 같이 영상을 송신하는 측에 구비되어 객체를 분할하는 동작을 수행한다. 도 1을 참조하면, 송신기(110)는 초기 분할된 객체와 배경의 컬러를 이용하여 컬러 확률 밀도 함수를 추정하는 컬러 모델링부(111)와, 시공간 상관관계를 이용하여 일관성(Coherence) 확률 밀도 함수를 추정하는 시공간 모델링부(112)와, 에너지 최소화를 수행하여 객체를 분할하는 객체 분할부(113)와, 영상 정보를 인코딩하는 비디오 인코더(114)와, 영상 비트스트림을 송신하는 송신부(115)를 포함한다. 상기 송신기(110)는 휴대폰, PDA, 스마트폰등 휴대용 모바일 기기 및, 노트북, 데스크탑 등의 고정형 통신 장치를 포함하며, 통신망을 통해 영상 정보를 송수신하여 다른 기기와 화상 통신이 가능한 통신 기기를 의미한다.

<12> 도 2는 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법의 동작 흐름도이다. 도 2를 참조하면, 먼저 210단계에서 동영상의 첫 프레임에서 초기 객체 경계를 추정한다. 본 발명의 일 실시 예에서는 초기 객체의 경계를 추정하기 위해 에너지 최소화(Energy Minimization)가 반복되며 사용자의 상호 작용이 필요한 그래프 컷(GrabCut) 알고리즘(C.Rother, V.Kolmogorov, and A.Black, "GrabCut-interactive foreground extraction using iterated graph cuts" in ACM Transactions on Graphics, 2004, vol.23, no.3, pp.309-314)을 사용한다.

<13> 도 3은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법에서 초기 객체 분할 동작의 단계를 보여주는 예시도이다. 도 3은 'foreman' 영상의 첫프레임을 그래프 컷 알고리즘을 이용하여 초기 객체 분할하는 과정을 도시한다. 처음 사용자는 객체를 분할하기 위해 객체가 포함된 대략적인 사각형을 선택한다. 그러면 반복적인 최소화 방식이 임시 레이블(label)을 할당한다. 임시 레이블은 각 픽셀이 객체에 속하는지 배경에 속하는지 표시한다. 만약 임시 레이블이 정확하지 않으면 사용자의 표시가 더 필요하다. 이 경우 사용자는 원하는 윤곽 주변에 화소 붓 형태를 사용하여 표시할 수 있으며, 에너지 최소화는 계속 수행된다. 상기과 같이 그래프 컷 알고리즘을 이용하여 초기 객체 분할을 수행할 수 있다.

<14> 다음 220단계에서는 상기 210단계에서 분할된 객체와 배경에 대하여 각각 컬러 확률 밀도 함수를 추정한다. 본 발명에서 컬러는 (R,G,B)공간을 표현한다. M_1 을 객체의 픽셀 집합이라 하고, M_0 은 배경의 픽셀 집합이라 하고,

x_p 는 픽셀 P의 컬러를 표시하는 삼차원 벡터라 하면 M_0 , M_1 집합을 사용하여 배경과 객체 각각 파라미터 없이 확률 밀도 함수를 하기의 수학적 식 1과 같이 추정할 수 있다.

수학적 식 1

$$p_{\text{color}}(\mathbf{x}|M_\alpha) = \frac{1}{h^d |M_\alpha|} \sum_{\mathbf{p} \in M_\alpha} K\left(\frac{\mathbf{x} - \mathbf{x}_p}{h}\right), \quad \alpha = 0, 1$$

<15>

<16> 상기 수학적 식 1에서 K는 커널 함수(kernel function)이고, h는 커널의 대역폭이고, d는 x의 차원(dimension)이고, $|M_\alpha|$ 는 M_α 내의 픽셀의 수를 표시하고, 레이블(label) α 는 배경일 경우 0을 객체일 경우 1이 된다. 본 발명에서는 d는 3이다.

<17> 본 발명의 일 실시 예에서는 다변수(multivariate)의 에파네치니코브 커널(Epanechnikov kernel)(B.W.Silverman, "Density Estimation for Statistics and Data Analysis", in Chapman and Hall, 1986)이 사용되며, 하기의 수학적 식 2와 같다.

수학적 식 2

$$K(\mathbf{x}) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \mathbf{x}^t \mathbf{x}) & \text{if } \mathbf{x}^t \mathbf{x} < 1 \\ 0 & \text{otherwise} \end{cases}$$

<18>

<19> 상기 수학적 식 2에서 상기 c_d 는 d-차원 영역의 부피이다($C_1=2$, $C_2=\pi$, $C_3=4\pi/3$). 본 발명의 일 실시 예에서 사용하는 에파네치니코브 커널은 방사상으로(radially) 대칭(symmetric)이며, 단봉(uni-modal)이다. 또한 가우시안 커널(Gaussian kernel)보다 빠르게 계산되는 이점을 가진다. 계산적인 복잡성을 줄이기 위해 수학적 식 1의 컬러 확률 밀도 함수는 처음 한번 계산되고 룩 업 테이블(Look-up table)에 저장된다.

<20>

도 4는 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법에서 컬러 확률 밀도 함수를 3차원 좌표로 나타낸 예시도이다. 도 4의(a)는 'foreman' 동영상의 첫 프레임의 객체의 컬러 확률 밀도 함수를 나타내며, 도 4의(b)는 배경의 컬러 확률 밀도 함수를 나타낸다. 도 4에서 각 점의 휘도는 확률 값을 나타낸다.

<21>

다음 230단계에서는 일관성(coherence) 확률 밀도 함수를 추정한다. 상기 220단계에서 추정한 컬러 정보에 더하여, 본 발명에서는 일반적으로 연속된 프레임 사이에서는 객체의 위치가 비슷한 특성을 고려한다. 따라서 본 발명의 일 실시 예에서는 본 발명의 특징에 따라 현재 프레임 상에서 이전 프레임의 객체의 윤곽의 유사한 위치에 객체 스트립(strip)을 생성한다. 객체 스트립의 개념을 이용하면 시공간적으로 인접한 영상의 객체 경계 특성을 이용하기 때문에 계산의 복잡성을 효과적으로 줄일 수 있다.

<22>

도 5의(a)는 'foreman' 동영상의 프레임에서 분할된 객체를 도시한다. 도 3의(b)는 본 발명의 일 실시 예에 따른 객체 윤곽을 따르는 객체 스트립을 나타내는 예시도이다. 예를 들어 도 5의 (a)와 같은 분할된 객체에서는 도 5의(b)와 같은 객체 스트립을 생성할 수 있다.

<23>

만약 객체가 적은 움직임을 가지면 다음 프레임의 윤곽은 현재 프레임의 윤곽과 유사할 것이고, 스트립 안에 그대로 있을 확률이 매우 크다. 그러므로, 시공간 일관성(spatio-temporal coherence)을 이용하여 객체 스트립에 속하는 픽셀이 객체일 확률을 선형적 지식으로 추정한다. 도 3의(c)는 픽셀의 밝기로 표현된 선형 확률 밀도 함수를 도시한다. 도 3의(c)를 참조하면 각 픽셀의 밝기는 객체에 속할 확률을 나타낸다. α^t 는 프레임 t의 이미지의 바이너리(binary) 라벨을 나타내고, 픽셀 P의 α_p^t 는 P가 객체일 경우 1이고, P가 배경일 경우 0이 된다. 경계 스트립의 픽셀의 선형 확률 밀도 함수는 하기의 수학적 식 3과 같다.

수학식 3

$$p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1}) = \begin{cases} 0.5 + \frac{d_{\mathbf{p}}}{w} & \text{if } \alpha_{\mathbf{p}}^{t-1} = 1 \\ 0.5 - \frac{d_{\mathbf{p}}}{w} & \text{if } \alpha_{\mathbf{p}}^{t-1} = 0 \end{cases}$$

<24>

<25>

상기 수학식 3에서 w 는 경계 스트립의 너비이고, d_p 는 이전 프레임($t-1$)의 객체 경계로부터 픽셀 P 까지의 최단 거리이며, $0 \leq d_p \leq \frac{w}{2}$ 이다. 픽셀 p 가 바깥쪽 방향의 윤곽으로부터 멀어지게 되면 $p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1})$ 는 값이 점점 작아진다.

<26>

스트립의 너비 w 는 분할의 정확도와 계산 복잡도를 모두 고려하여 설정한다. 만약 너비가 매우 넓으면, 많은 픽셀들이 분류되어 더 긴 처리 시간이 걸리게 되고, 반대로 너비가 매우 좁으면, 객체가 빠르게 움직일 경우 객체 윤곽이 경계 스트립 바깥으로 이동할 수 있다.

<27>

그래서, 객체의 윤곽에 접치는 블록들의 움직임 벡터(motion vector)를 이용하여 너비 w 를 결정한다. 이전 프레임($t-1$)을 16×16 사이즈의 블록으로 나눈 뒤, 한 블록이 객체의 경계와 겹치면 'Three-step search' 알고리즘을 이용하여 현재 프레임(t)의 움직임 벡터를 추정한다. 다음 일정 수준을 벗어난 특이 값을 없애기 위해 경계와 겹치는 블록들의 중간(median) 움직임 벡터 v 를 계산한다. 마지막으로 v 의 절대값에 비례하여 w 를 설정한다. w 를 계산하는 식은 하기의 수학식 4와 같다.

수학식 4

$$w = w_{\min} + \|v\|$$

<28>

<29>

상기 수학식 4에서 $w_{\min}$ 은 스트립의 너비의 최소값으로서 바람직하게는 8로 설정한다.

<30>

게다가, 중간 움직임 벡터 v 는 이전 프레임($t-1$)에서 현재 프레임(t)으로의 객체의 전체 움직임 벡터로서 고려된다. 상기 전체 움직임 벡터에 의해 경계 스트립이 쉬프트(shift)된다. 쉬프트된 경계 스트립은 하기의 수학식 5와 같다.

수학식 5

$$p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1}) = \begin{cases} 0.5 + \frac{d_{\mathbf{p}-v}}{w} & \text{if } \alpha_{\mathbf{p}-v}^{t-1} = 1 \\ 0.5 - \frac{d_{\mathbf{p}-v}}{w} & \text{if } \alpha_{\mathbf{p}-v}^{t-1} = 0 \end{cases}$$

<31>

<32>

$p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 0 | \alpha^{t-1}) = 1 - p_{\text{cohere}}(\alpha_{\mathbf{p}}^t = 1 | \alpha^{t-1})$ 이다. 쉬프트된 확률 밀도 함수는 특히 빠른 움직임을 가진 객체에서 상기 수학식 3의 확률 밀도 함수보다 더 정확하다. 다시 말하면 전체 움직임을 고려하여 계산한 확률 밀도 함수가 더욱 개선된 객체 분할 결과를 얻을 수 있다.

<33>

다음 240단계에서 에너지 최소화(Energy minimization)를 수행한다. 에너지 최소화를 이용하여 결합 스트립 내의 각 픽셀을 분류한다. 스트립 바깥쪽의 픽셀은 배경으로 분류되며, 스트립 안쪽의 픽셀은 객체로 분류된다. 전체 프레임에 대하여 적용하지 않고 오직 경계 스트립에만 에너지 최소화를 적용하기 때문에 계산 복잡도가 많이 줄어들게 된다.

<34>

에너지 최소화에는 컬러, 일관성(coherence), 평탄(smoothness)의 세 가지 에너지 함수를 사용한다. 첫 번째 컬러 에너지 함수는 하기의 수학식 6과 같이 계산한다.

수학식 6

$$E_{\text{color}}(\alpha^t) = - \sum_{\mathbf{p} \in \mathcal{S}} \log p_{\text{color}}(\mathbf{x}_{\mathbf{p}} | M_{\alpha_{\mathbf{p}}^t})$$

<35>

<36> 상기 수학식 6에서 S는 결합 스트립의 픽셀 집합이다. 상기 함수는 만약 픽셀 P가 레이블되어 있거나 $\alpha_{\mathbf{p}}^t$ 로 분류되어 있다면, 상기 수학식 1의 컬러 확률이 높게 나타낸다.

<37> 유사한 방식으로 두 번째 일관성 에너지 함수는 하기의 수학식 7과 같다.

수학식 7

$$E_{\text{cohere}}(\alpha^t) = - \sum_{\mathbf{p} \in \mathcal{S}} \log p_{\text{cohere}}(\alpha_{\mathbf{p}}^t | \alpha^{t-1})$$

<38>

<39> 상기 수학식 7은 상기 수학식 5의 선형 확률 밀도 함수를 사용한다.

<40> 세번째, 평탄 에너지 함수는 유사한 강도(intensity)의 이웃하는 픽셀이 같은 레이블로 할당되도록 제한한다.(Y. Boykov and M.Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images" in Proc. Int. Conf. on Computer Vision, July 2001, vol.1, pp.105-112) 평탄 에너지 함수는 하기의 수학식 8과 같다.

수학식 8

$$E_{\text{smooth}}(\alpha^t) = \sum_{\substack{(\mathbf{p}, \mathbf{q}) \in \mathcal{N} \\ \alpha_{\mathbf{p}} \neq \alpha_{\mathbf{q}}}} \frac{1}{\|\mathbf{p} - \mathbf{q}\|} \exp \left(- \frac{(z_{\mathbf{p}} - z_{\mathbf{q}})^2}{2\sigma^2} \right)$$

<41>

<42> 상기 수학식 8에서 N은 8방향의 인접한 픽셀의 쌍의 집합이고, $z_{\mathbf{p}}$ 는 컬러 $X_{\mathbf{p}}$ 로부터 계산된 픽셀 p의 강도(픽셀의 데이터 크기)이고, σ 는 강도의 표준 편차이다.

<43> 총괄 에너지 함수는 상기 세 가지 에너지 함수의 합이며, 하기의 수학식 9와 같다.

수학식 9

$$E(\alpha^t) = E_{\text{color}}(\alpha^t) + E_{\text{cohere}}(\alpha^t) + E_{\text{smooth}}(\alpha^t)$$

<44>

<45> 본 발명의 일 실시 예에서는 에너지 최소화를 위해 그래프 컷(graph cut) 기술(Y.Boykov and V.Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no.9, pp.1124-1137, 2004년 9월)을 사용한다. 그래프 G는 노드 집합 V와 에지 집합 E를 나타낸다. 노드 집합 V는 픽셀과 레이블로 구성된다. 이 경우 두 개의 에지 타입을 설정하며, 첫 번째 에지 타입은 픽셀과 그 레이블에 연결되고, 상기 수학식 6의 컬러 에너지 함수와, 상기 수학식 7의 일관성 에너지 함수로부터 에지 강도가 할당된다. 두 번째 에지 타입은 두 개의 인접한 픽셀에 연결되고, 상기 수학식 8의 평탄 에너지 함수로부터 에지 정도가 할당된다. 가중치 그래프의 min-cut은 배경으로부터 객체의 최적 분할을 제공한다.

<46> 각 프레임의 객체 분할 후, 경계 스트립의 픽셀의 컬러가 상기 수학식 1의 컬러 확률 밀도 함수와 매치하는지 검사한다. 이 작업에서 만약 두 개의 확률 밀도 함수 모두에서 0의 확률을 가진 픽셀이 7% 이상이면, 확률 밀도 함수의 업데이트를 수행한다. $t_k(k=1,2, \dots)$ 는 업데이트 프레임의 시간 인덱스를 나타낸다. t_k 프레임의 컬러 확

를 밀도 함수를 업데이트할 경우, 새로운 샘플과 같이 과거 샘플을 포함하도록 t_k , t_{k-1} , t_{k-2} 의 과거 세 프레임의 컬러를 사용한다.

<47> 도 6은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 동작의 수행 결과를 보여주는 프레임 예시도이다. 도 6의 (a),(b),(c)는 각각 'Akiyo', 'Mother and Daughter', 'Foreman' 동영상에서 본 발명의 일 실시 예에 따라 객체를 분할한 결과를 나타내는 예시도이다. 도 6에서 왼쪽 줄은 20번째 프레임, 중간 줄은 30번째 프레임, 오른쪽 줄은 40번째 프레임을 나타낸다.

<48> 도 7은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법을 이용하여 배경을 대체한 영상의 예시도이다. 본 발명의 일 실시 예에 따라 객체와 배경의 분할을 수행하여 화상 통신시 배경을 다른 배경으로 대체한 영상을 보여준다. 본 발명은 복잡한 배경이나 이동으로 인한 배경의 변화 및 객체의 움직임에 대하여 강력한 성능을 보여주며, 또한 프레임 전체가 아닌 시공간상 일관성을 판단하여 객체가 유력한 부분의 영역만을 고려하여 일관성 확률 밀도 함수를 계산하기 때문에 적은 연산량을 가지는 장점이 있다.

<49> 상기와 같이 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법의 동작이 이루어질 수 있으며, 한편 상기한 본 발명의 설명에서는 구체적인 실시 예에 관해 설명하였으나 여러 가지 변형이 본 발명의 범위를 벗어나지 않고 실시될 수 있다.

도면의 간단한 설명

<50> 도 1은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 장치의 블록 구성도

<51> 도 2는 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법의 동작 흐름도

<52> 도 3은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법에서 초기 객체 분할 동작의 단계를 보여주는 예시도

<53> 도 4는 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법에서 컬러 확률 밀도 함수를 3차원 좌표로 나타낸 예시도

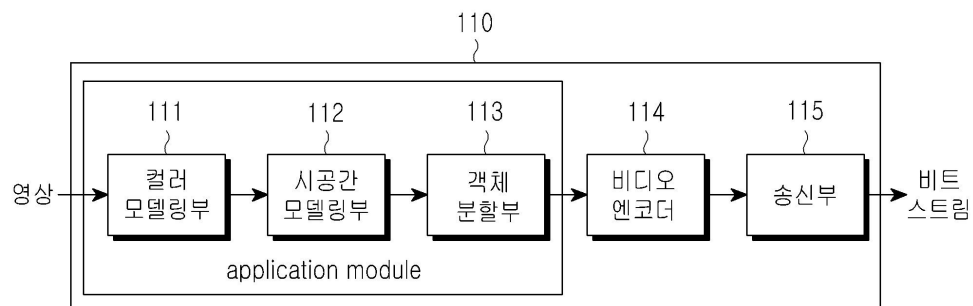
<54> 도 5는 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법에서 경계 스트립을 나타내는 예시도

<55> 도 6은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 동작의 수행 결과를 보여주는 프레임 예시도

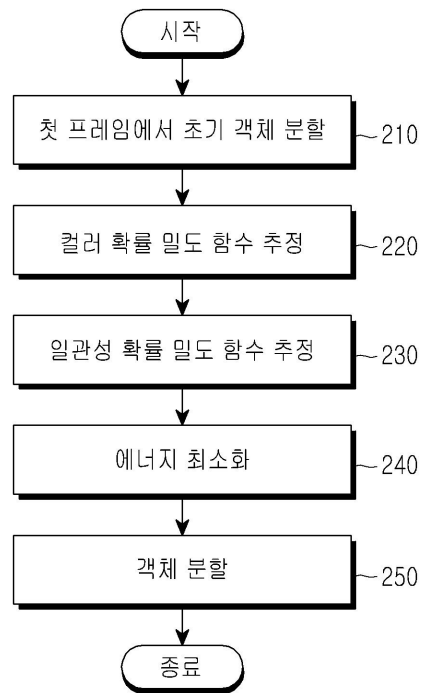
<56> 도 7은 본 발명의 일 실시 예에 따른 동영상의 객체 경계 추정 방법을 이용하여 배경을 대체한 영상의 예시도

도면

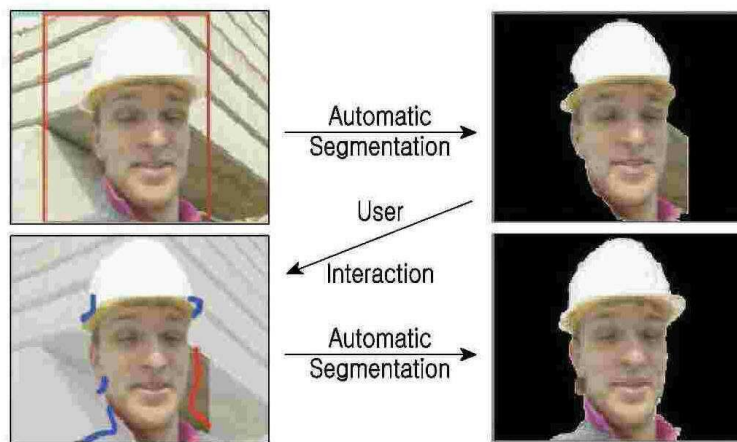
도면1



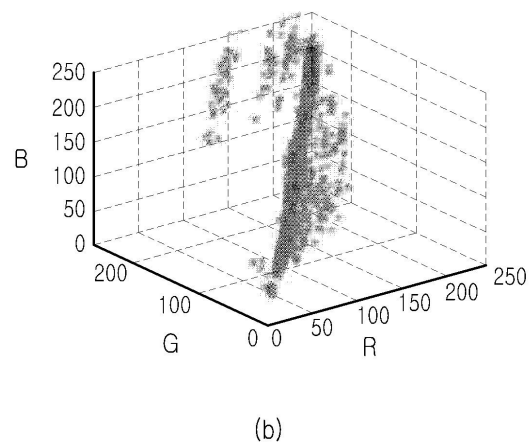
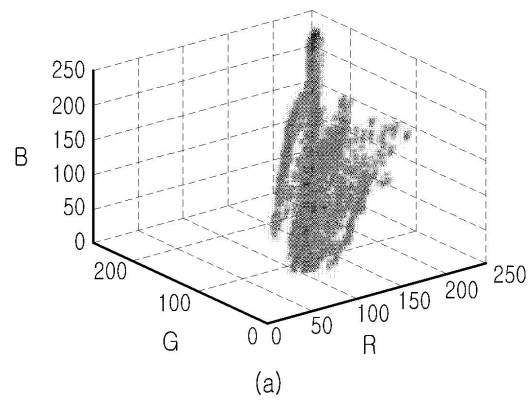
도면2



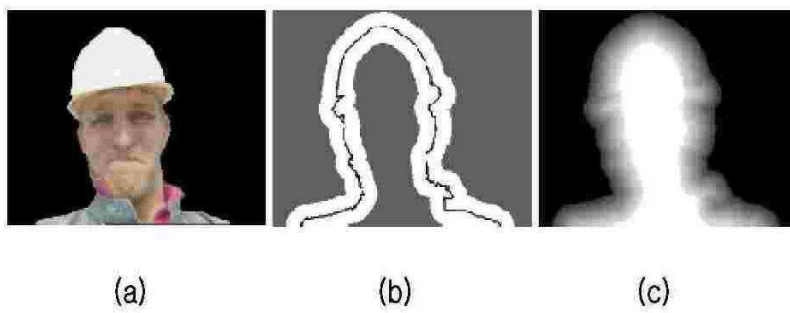
도면3



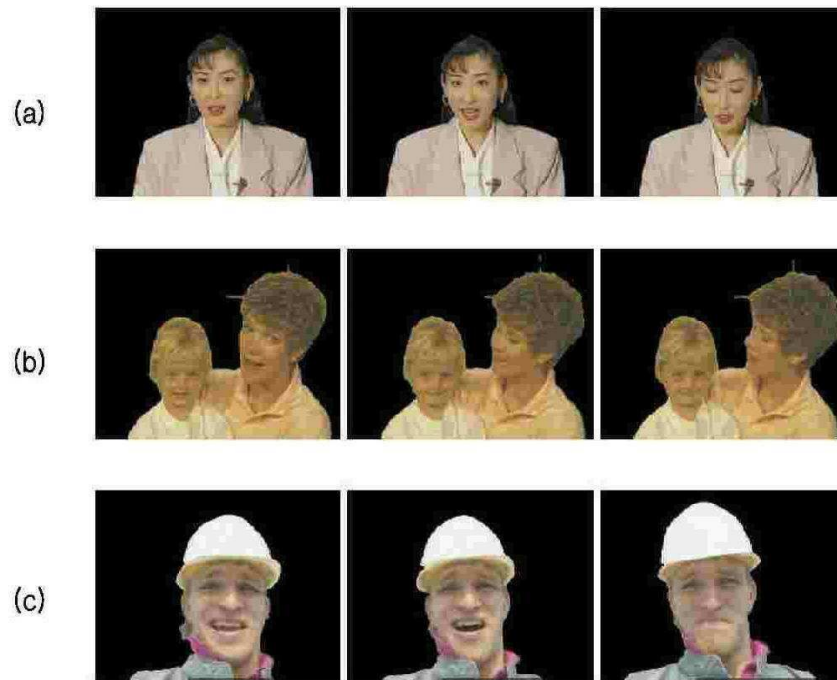
도면4



도면5



도면6



도면7

