

물체별 부호화를 위한 이동영역의 분할 및 움직임 추정

正會員 崔在覺*, 李時雄*, 金聖大*, 李炳旭**

Segmentation and Motion Estimation of Moving Objects for Object-Oriented Coding

Jae Gark Choi*, Si Woong Lee*, Seong Dae Kim*, Byung Wook Lee** Regular Members

要 約

본 논문은 물체별 부호화를 위한 이동물체 분할 및 움직임 추정에 관한 것이다. 일반적인 움직임(general motion)을 추정할 때 가장 큰 어려움은, 높은 신뢰성의 추정을 위해서는 추정영역이 커야 하지만 추정영역이 커질 때 그 영역은 균일한 움직임의 영역이 아닐 가능성이 커진다는 것이다. 이러한 문제를 해결하기 위해 이동벡터에 근거한 다단계 분할 알고리즘(multi-stage segmentation algorithm)을 제안한다. 이 방법은 영역확장(region growing)을 사용하여 복잡도가 낮은 모형에 따라 분할된 작은 영역(subregion)을 복잡도가 더 큰 모형의 관점에서 균일한 영역(region)으로 병합해 나간다. 본 논문에서 제안한 다단계 분할 방식은 3가지 매핑 파라미터 모형에 따라 3단계의 분할 과정을 점진적으로 거치게 된다. 첫 단계에서는 비슷한 이동벡터를 갖는 이웃영역의 화소를 모아서 2차원 병진운동이 균일한 영역으로 분할한다. 둘째 단계에서는 앞서 분할한 병진운동 영역을 균일한 움직임을 갖는 3차원 평면체 영역(planar patch)으로 병합하고, 마지막 단계에서 평면체 영역을 곡면체 영역(parabolic patch) 분리하게 된다. 따라서 머리-어깨 영상(head-and-shoulder image)에서 사람의 얼굴이 하나의 이동물체로 분리될 수 있다. 머리-어깨 영상에 대한 모의 실험 결과, 움직임영역 전체가 1-3개의 곡면체 영역으로 분할되었다.

ABSTRACT

This paper presents a segmentation and motion estimation method for object-oriented analysis-synthesis coding. A major difficulty in estimating general motion is that it requires a large area of support in order to achieve a good estimation. Unfortunately, when the supporting area is large it is very likely to have multiple moving objects. To solve this problem, we propose a multi-stage segmentation method which is based on optical flow. The basic concept is to group homogeneous subregions with respect to sim-

*한국과학기술원 전기 및 전자공학과

**대우전자(주)

論文番號 : 94321-1119

接受日字 : 1994年 11月 19日

pler mapping model into large homogeneous regions with respect to more complex mapping model. By applying a hierarchy of mapping parameter model progressively, we can segment the whole changed region into several parabolic patches. Especially person's face in head-and-shoulder images can be described as one object.

1. 서 론

H. 261과 같은 기존의 부호화 방법은 영상을 블럭으로 나누고 각 블럭을 운동보상 예측(motion compensated prediction)에 의하여 부호화한다. 따라서 저진송률 부호화시 블럭화 현상(block effect)이 나타난다. 블럭화 현상과 같은 영상 왜곡을 방지하기 위하여 최근 물체별 부호화(object-oriented analysis-synthesis coding)라 불리는 새로운 부호화 방법이 Musman 및 Hötter에 의해 제안되었다[1]. 이러한 물체별 부호화는 이동물체 영역을 추출하고 물체별로 움직임을 보상하는데 근거하고 있다.

물체별 부호화의 기본적인 문제 중 하나가 이동물체 영역을 추출하고 그 영역의 움직임을 추정하는 것이다. 즉 영상 해석(image analysis)이다. 이러한 영상 해석법 중 하나로 시공간 경사치를 이용하여 이동물체 분할을 계층적으로 수행하는 방법이 Hötter에 의해 제안되었다[2]. 이 방법은 분리된 각 움직임을 영역(isolated changed region)을 하나의 이동물체 영역이라 가정하고 매핑 파라미터(mapping parameter)를 추정한다. 이때 다수의 이동물체가 한 움직임을 영역에 포함되어 있는 경우 추정된 파라미터는 움직임을 영역의 주된 물체(dominant object)만을 기술하게 된다. 따라서 추정된 파라미터로 기술되지 못하는 영역은 다음 계층의 움직임을 영역이 되어 위의 과정을 계층적으로 수행하게 된다. 여기서 3차원 평면체(planar rigid object)의 일반적인 움직임(general motion)을 8개의 파라미터를 이용하여 기술하였다[5]. 이와 유사한 영상 해석법이 Diehl에 의해 제안되었다[3]. 이 방법은 시공간 경사치를 이용하여 계층적으로 이동물체를 분할하는 것은 Hötter방법과 유사하나 이동물체 분할의 신뢰성을 높이기 위해 정지 영상을 분할(intra-frame segmentation)한 정보를 이용하였다. 그리고 3차원 곡면체를 고려하는 일반적인 매핑 파라미터 모형을 사용하였다. Hötter와 Diehl의 영상 해석법은 모두 움직임을 영역 전체를 하나의 이동물체라 가

정하고 매핑 파라미터를 추정하므로 추정된 매핑 파라미터는 움직임을 영역 내의 크기가 가장 큰 주된 물체만을 기술하게 된다. 그러나 움직임을 영역내에 주된 물체가 없으면 이 방법들은 물체 분할 및 움직임 추정을 하지 못하는 단점이 있다. 다른 한편으로 이동 벡터장을 이용하는 이동물체 분할 방법이 Adiv에 의해 제안되었다[4]. 이 방법은 Hough 변환과 8파라미터 모형을 사용하여, 이동 벡터장을 근사적 평면물체의 움직임과 일치하는 영역으로 분할한다. 즉, 영상좌표와 이동벡터를 매핑 파라미터 공간으로 변환한뒤 가장 많은 벡터들이 속하는 파라미터 공간 영역의 좌표값들이 큰 영역의 이동물체에 대한 변환식을 나타냄을 이용하는 방식이다. 이러한 Hough 변환 기법들은 다른 방식에 비해 잡음 특성은 좋으나 메모리와 계산량이 많은 단점이 있다.

본 논문에서는 이동 벡터장을 구한뒤 이를 이동물체별로 분할하는 방법을 사용한다. 이를 위하여 추정영역(supporting area)의 크기와 추정영역의 균일성(region homogeneity)의 관계를 먼저 분석하였다[6]. 일반적인 움직임을 추정할 때 가장 큰 어려움은 다음과 같다. 복잡한 매핑 파라미터 모형을 사용할수록 다양한 물체표면 및 물체움직임을 기술할 수 있으므로 하나의 매핑 파라미터 집합에 의하여 표현될 수 있는 영역은 커진다. 하지만 추정하려는 매핑 파라미터의 신뢰성과 수렴성을 고려할 때 복잡한 매핑 파라미터 모형일수록 추정영역이 커야한다. 추정영역이 커지면 그 영역이 추정하려는 매핑 파라미터 모형의 관점에서 균일한 움직임(uniform motion)이 아닐 확률이 커진다. 이러한 모형의 복잡도(the complexity of model) 및 추정영역의 크기(the size of supporting area)와 관련된 영역 균일성의 문제를 해결하기 위해 본 논문에서는 다단계 이동물체 분할방법(multi-stage segmentation algorithm)을 제안한다. 이 방법은 영역확장(region growing)을 사용하여 복잡도가 낮은 모형에 따라 분할된 작은 영역을 복잡도가 더 큰 모형의 관점에서 균일한 영역으로 병합해 나간다. 본 논문에서는 3가지 매핑 파

라미터 모형에 따라 3단계의 분할 과정을 점진적으로 거치게 되는데 첫 단계에서는 2차원 병진운동이 균일한 영역으로, 둘째 단계에서는 균일한 움직임을 갖는 3차원 평면체 영역으로, 마지막 단계에서는 균일한 움직임을 갖는 3차원 곡면체 영역으로 분리하게 된다. 따라서 머리-어깨 영상(head-and-shoulder image)에서 사람의 얼굴이 하나의 이동물체로 분리될 수 있다. 머리-어깨 영상에 대한 모의실험 결과, 움직임 영역 전체가 1-3개의 곡면체 영역으로 분할되었다. 또한 제안된 방식은 (2) 및 (3)의 방식과는 달리 이동벡터장을 매핑 파라미터 모형에 따라 점진적으로 영역확장을 행하기 때문에 움직임영역내에 주된 물체(dominant object)가 없는 경우에도 이동물체를 분리할 수 있다. 그리고 영역확장에 의한 분할 방법을 사용하기 때문에 변형된 Hough 변환을 이용한 방식(4)에 비해 계산량과 메모리량이 훨씬 작아도 되는 장점이 있다.

본 논문의 구성은 다음과 같다. 2절에서 다단계 분할 방식에 사용될 매핑 파라미터 모형에 대해 기술한다. 3절에서 이동물체 분할 및 움직임 추정을 위한 다단계 분할 알고리즘에 대해 설명하고 4절에서 다단계 분할 방식의 성능 평가를 위한 모의실험 결과를 제시한다. 마지막으로 5절에서 결론을 맺는다.

2. 매핑 파라미터 모형

연속된 영상에서 시간에 따른 밝기 변화가 물체나 카메라의 움직임에 기인한다고 가정할 때 이러한 밝기 변화를 기술하기 위해 매핑 파라미터 모형이 필요하다. 매핑 파라미터 모형을 이용하면 현재 프레임, $I_n(x, y)$ 의 한 영역을 다음 프레임, $I_{n+1}(x, y)$ 의 대응 영역으로 매핑시킬 수 있다. 지금부터 매핑 파라미터 모형을 전개할 때 사용하는 좌표의 표기는 다음처럼 하기로 한다.

- (X, Y, Z) = 움직임 전의 물체 공간상의 한 점 P의 좌표
- (X', Y', Z') = 움직임 후의 물체 공간상의 한 점 P의 좌표
- (x, y) = 움직임 전의 점 P의 영상 평면상의 좌표
- (x', y') = 움직임 후의 점 P의 영상 평면상의 좌표

매핑 파라미터 모형은 아래와 같은 운동모형, 투영모형, 그리고 물체모형에 의해 결정된다(3).

• 운동모형(motion model):

운동모형의 경우, 물체(rigid body)의 일반적인 3차원 움직임(general motion)은 회전행렬 R과 병진벡터

T에 의해 다음처럼 기술될 수 있다

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = R \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + T = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}. \quad (1)$$

• 투영모형(projection model):

투영모형은 3차원 물체공간을 영상평면으로 투영하는 과정을 나타낸다. 이러한 투영은 집중투영(central projection)이나 평행투영(parallel projection)으로 모형화 할 수 있다(7). 집중투영의 경우 물체공간상의 좌표와 영상평면상의 좌표는 다음과 같은 대응 관계를 갖는다.

$$x = F \frac{X}{Z}, \quad y = F \frac{Y}{Z}. \quad (2)$$

여기서 F는 카메라의 촛점거리를 나타낸다.

그리고 평행투영의 경우는 식(3)과 같은 관계가 성립한다.

$$x = mX, \quad y = mY. \quad (3)$$

여기서 상수 m은 촛점거리와 물체까지의 평균거리에 의해 결정된다.

집중투영이 평행투영보다 영상화 미커니즘(imaging mechanism)을 더 정확히 모형화할 수 있지만 만일 카메라로부터의 거리에 비해 물체의 크기가 충분히 작다면 집중투영의 효과가 감소하기 때문에 평행투영을 사용할 수 있다.

• 물체모형(object model):

물체모형은 물체의 표면을 기술한다. 머리-어깨 영상에서 머리와 어깨는 아주 복잡한 형태를 보이지만 계산상의 간단함과 구현상의 용이성을 고려하면 평면체(planar surface)와 곡면체(parabolic surface)의 형태로 가정할 수 있다. 평면체의 경우 물체표면은 다음처럼 기술 될 수 있다.

$$Z = p_1X + p_2Y + p_3. \quad (4)$$

그리고 곡면체의 경우 물체표면은 식(5)와 같이 기술된다.

$$Z = q_1X_2 + q_2XY + q_3Y^2 + q_4X + q_5Y + q_6. \quad (5)$$

위의 운동모형, 투영모형, 물체모형을 조합하면 하나의 매핑 파라미터 모형을 구성할 수 있다. 본 논문에서는 2차원 병진운동, 3차원 평면체 운동, 그리고 3차원 곡면체 운동을 각각 기술할 수 있는 3가지 매핑 파라미터 모형을 사용한다. 이 3가지 모형이 움직임에 따른 밝기 변화를 계층적으로 기술한다. 그리고 단계 분할 알고리즘에 따라 3가지 모형을 점진적으로 적용한다. 아래에 이동벡터, $(x'-x, y'-y)$ 를 표시하기 위해 (u, v) 를 사용한다.

2.1 2차원 병진 운동모형

2차원 병진 운동모형(2-D translation model)은 영상평면상의 물체움직임을 기술하기 때문에 투영모형이나 물체모형은 필요치 않고 2차원상에서의 운동모형만이 사용된다. 2차원 영상 평면상에서의 병진운동의 경우 아래와 같은 2파라미터로 표현된다.

$$x' = x + a_1, \quad y' = y + b_1. \quad (6)$$

따라서 이동벡터는 다음과 같은 단순 변환(simple transformation)형태로 표시된다.

$$u(x, y) = a_1, \quad v(x, y) = b_1. \quad (7)$$

2.2 평면체 운동모형

평면체 운동모형(planar patch model)은 3차원 평면체(a rigid planar surface), 평행투영 그리고 3차원 움직임을 가정할 때 성립하는 모형이다. 따라서 식(1), 식(3), 식(4)를 결합하면 다음과 같은 식을 얻을 수 있다.

$$\begin{aligned} x' &= (r_3 p_3 m' + m T_X) + \frac{m'}{m} (r_1 + r_3 p_1) x \\ &\quad + \frac{m'}{m} (r_2 + r_3 p_2) y, \\ y' &= (r_6 p_3 m' + m T_Y) + \frac{m'}{m} (r_4 + r_6 p_1) x \\ &\quad + \frac{m'}{m} (r_5 + r_6 p_2) y. \end{aligned} \quad (8)$$

식(8)은 6개의 파라미터를 이용하여 아래식처럼 간단히 표현된다.

$$x' = a_1 + (1 + a_2)x + a_3 y, \quad y' = b_1 + b_2 x + (1 + b_3)y. \quad (9)$$

따라서 이동벡터는 다음과 같은 affine 변환을 갖는다

$$u(x, y) = a_1 + a_2 x + a_3 y, \quad v(x, y) = b_1 + b_2 x + b_3 y. \quad (10)$$

2.3 곡면체 운동모형

곡면체 운동모형(parabolic patch model)은 3차원 곡면체(parabolic surface), 평행투영 그리고 3차원 움직임을 가정할 때 성립하는 모형이다. 따라서 식(1), 식(3), 식(5)를 결합하면 다음과 같은 식을 얻을 수 있다.

$$\begin{aligned} x' &= m'(T_X + r_3 q_6) + \frac{m'}{m} (r_1 + r_3 q_4) x + \frac{m'}{m} (r_2 + r_3 q_5) y \\ &\quad + \frac{m'}{m^2} r_3 q_1 x^2 + \frac{m'}{m^2} r_3 q_3 y^2 + \frac{m'}{m^2} r_3 q_2 xy, \end{aligned} \quad (11)$$

$$\begin{aligned} y' &= m'(T_Y + r_6 q_6) + \frac{m'}{m} (r_4 + r_6 q_4) x + \frac{m'}{m} (r_5 + r_6 q_5) y \\ &\quad + \frac{m'}{m^2} r_6 q_1 x^2 + \frac{m'}{m^2} r_6 q_3 y^2 + \frac{m'}{m^2} r_6 q_2 xy. \end{aligned}$$

식(11)을 간략히 표현하면 아래와 같은 12파라미터로 표현된다.

$$\begin{aligned} x' &= a_1 + (1 + a_2)x + a_3 y + a_4 x^2 + a_5 y^2 + a_6 xy, \\ y' &= b_1 + b_2 x + (1 + b_3)y + b_4 x^2 + b_5 y^2 + b_6 xy. \end{aligned} \quad (12)$$

따라서 이동벡터는 다음과 같은 quadratic 변환을 갖는다.

$$\begin{aligned} u(x, y) &= a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 y^2 + a_6 xy, \\ v(x, y) &= b_1 + b_2 x + b_3 y + b_4 x^2 + b_5 y^2 + b_6 xy. \end{aligned} \quad (13)$$

3. 단계 분할 알고리즘

본 절에서는 움직임 추정 및 이동물체 분할을 위한 단계 분할 알고리즘에 대해 기술한다. 이동물체의 움직임을 추정할 때 신뢰성있는 파라미터 추정을 위해서 추정영역은 다음 조건을 만족해야 한다.

- 추정영역이 고려 중인 매핑 파라미터 모형의 관점에서 균일한 움직임을 가져야 한다.
- 복잡한 매핑 파라미터 모형일수록 움직임 추정에 사용하는 추정영역은 커야 한다.

따라서 일반적인 움직임을 추정할 때 가장 큰 어려움은, 높은 신뢰성의 추정을 위해서는 추정영역이 커야 하지만 추정영역이 커질 때 그 영역은 균일한 움직임의 영역이 아닐 가능성이 커진다는 것이다. 이러한 문제를 해결하기 위한 다단계 분할 알고리즘의 기본 개념은 다음과 같다. 먼저 복잡도가 낮은 모형에 따라 움직임영역을 분할 한다. 이렇게 분할된 영역은 복잡도가 낮은 모형의 관점에서 균일한 움직임을 갖는다. 이때 추정영역은 크지 않아도 되므로 균일한 움직임의 영역이 될 가능성이 커진다. 그리고 추정된 파라미터도 충분한 신뢰도를 가진다. 다음은 복잡도가 더 큰 파라미터 모형을 이용하여 움직임영역을 분할한다. 파라미터 모형이 복잡하면 추정영역이 커야 하고 동시에 그 영역은 균일한 움직임을 가져야 한다. 따라서 앞단계의 복잡도가 낮은 모형에 따라 분할된 영역을 복잡도가 더 큰 모형의 관점에서 균일한 영역으로 병합해가면 위의 두 조건을 동시에 만족하게

된다.

본 논문에서는 3가지 매핑 파라미터 모형에 따라 3단계의 분할 과정을 거치게 되는데 첫 단계에서는 비슷한 이동벡터를 갖는 이웃영역의 화소를 모아서 2차원 병진운동의 관점에서 균일한 영역으로 분할하는 것이다. 둘째 단계에서는 앞서 분할한 병진운동 영역을 균일한 움직임을 갖는 3차원 평면체 영역으로 병합하는 것이다. 마지막 단계는 평면체 영역을 균일한 움직임을 갖는 3차원 곡면체 영역으로 병합하는 것이다. 따라서 머리-어깨 영상에서 사람의 얼굴이 하나의 이동물체로 분리될 수 있다. 이는 사람의 얼굴을 곡면체로 근사화하고 곡면체는 평면체의 집합으로 근사화된다는 가정하에 성립하는 것이다. 다단계 분할 알고리즘의 흐름도는 그림 1과 같다. 아래에 각 과정에 대해 상세히 기술한다.

3.1 움직임 판정기

움직임 판정기(change detector)는 연속된 두 영상에 대해 시간적으로 변화가 있는 움직임 영역과 변화가 없는 영역으로 구분하기 위하여 사용한다. 여기서 사용하는 움직임 판정기는 참고문헌(2)에 주어진 것과 유사하다. 먼저 연속된 두 영상의 차 신호에 대해 3*3 마스크를 이용한 저대역 통과 필터를 거친 뒤, 그 결과를 주어진 임계값(threshold)과 비교하여 움직임 영역과 배경으로 이진화한다. 이진화된 영상은 5*5 크기의 미디안 필터(median filter)를 통과하여 고립된 화소를 주위영역으로 할당된 뒤 크기가 작은 영역들을 제거하는 과정을 거치게 된다. 이동물체의 움직임 추정과 분할은 움직임영역에 대해서만 이루어지고 밝기 변화가 없는 영역은 배경(background)으로 간주한다.

3.2 이동벡터의 추정

이동벡터 추정은 이동벡터 검출과 이동벡터 보정으로 구성된다. 이동벡터 검출을 위해 먼저 수평, 수직방향으로 각각 4 화소마다 하나의 이동벡터를 추정한 뒤, 그 사이에 위치한 화소에 대한 이동벡터는 선형보간(bilinear interpolation)을 사용하여 구한다. 4 화소마다 하나의 이동벡터를 구할때, 9*9 측정 윈도우(measurement window)를 사용한 전탐색 BMA(full search BMA)를 사용하여 추정하고 추정된 벡터를 9*9 윈도우의 가운데 화소의 이동벡터로 하였다. 그리고 수평, 수직방향의 최대 이동벡터가 각각 ±10이 되

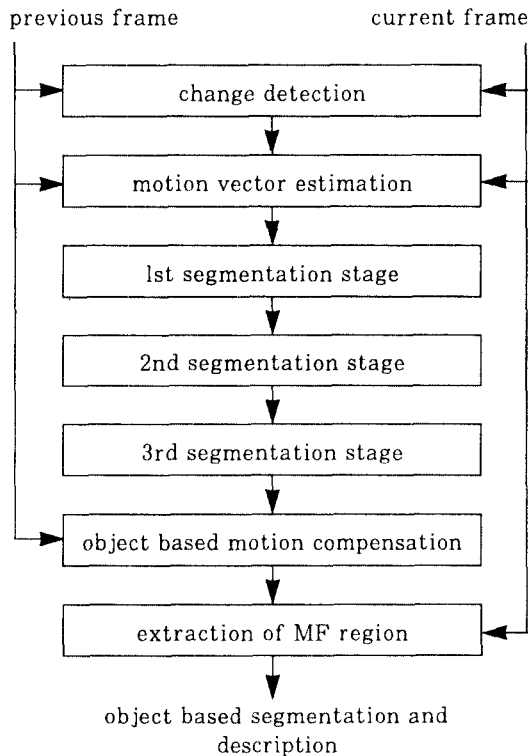


그림 1. 다단계 분할 알고리즘의 구조.
Fig. 1. Structure of the multi-segmentation algorithm.

도록 탐색영역(search area)을 정하였다. 이렇게 검출된 이동벡터는 다음과 같은 사실 때문에 실제 움직임(real motion)이 아닐 수 있다.

- 추정영역이 드러난 배경(uncovered background)이나 많은 잡음을 포함하는 경우

- 최소화하려는 왜곡함수(distortion function)로 실제 움직임을 검출하지 못하는 경우

따라서 이러한 이동벡터를 보정하기 위해 수평, 수직 방향의 이동벡터 성분에 대해 각각 미디안 필터를 적용하였다.

3.3 첫번째 분할단계

첫번째 분할단계(the first segmentation stage)에서는 앞에서 구한 움직임영역을 이동벡터가 유사한 영역으로 분할하는 것이다. 이렇게 분할된 영역을 2차원 병진운동 영역(2-D translational patch)이라 하고 각 영역은 2차원 병진운동 관점에서 균일한 움직임을 갖는다. 이러한 분할을 위해 다음과 같은 화소 단위의 영역확장 방법(region-growing method by pixel-aggregation)을 사용하였다. 이동벡터장을 상하-좌우로 스캔(raster scan)하면서 각 화소를 이동벡터가 유사한 영역에 할당하게 된다. 현재의 화소가 (k, l) 위치에 있다고 하면 이 화소 주위에는 이웃한 영역들 R_j 가 존재하게 된다. 이웃한 영역의 정의는 화소 (k, l) 의 이웃한 화소 $(k-1, l-1), (k-1, l), (k-1, l+1), (k, l-1)$ 이 속한 영역으로 한다. 따라서 화소 (k, l) 은 이웃한 영역들 R_j 중 하나로 병합을 시도하게 되며 만일 병합할 만한 이웃영역이 없으면 이 화소는 새로운 영역으로 할당된다. 화소 (k, l) 을 이웃한 영역 R_j 로 병합하기 위한 기준으로 다음과 같이 정의된 유사성 측도(similarity measure) $S_{(k, l), R_j}$ 를 사용하였다.

$$S_{(k, l), R_j} = |d_{x, R_j} - d_{x, (k, l)}| + |d_{y, R_j} - d_{y, (k, l)}| \quad (14)$$

여기서 d_{x, R_j} 및 d_{y, R_j} 는 각각 이웃한 영역 R_j 의 평균 이동벡터의 x축, y축 값을 나타낸다. 그리고 $d_{x, (k, l)}$ 및 $d_{y, (k, l)}$ 은 각각 이동벡터 $d_{(k, l)}$ 의 x축, y축 값을 나타낸다. 만일 유사성 측도가 주어진 임계값보다 작으면 이웃한 영역 R_j 는 병합후보가 된다. 그리고 점 (k, l) 은 병합후보 중 오차가 가장 적은 후보로 병합된다.

3.4 두번째 분할단계

두번째 분할단계(the second segmentation stage)는 앞단계에서 구한 2차원 병진운동 영역을 동일한 움직임을 갖는 3차원 평면체 영역으로 영역확장을 행하는 것이다. 이렇게 분할된 영역을 평면체 영역(planar patch)이라 하자. 3차원 평면체의 움직임은 식(10)으로 표현되는 평면체 운동모형에 의하여 설명될 수 있다. 따라서 두번째 분할 단계는 동일한 affine 변환을 갖는 영역들을 영역확장을 통하여 동일한 움직임을 갖는 3차원 평면체의 영역으로 병합해 나가는 것이다. 병합(grouping)문제를 해결하기 위해 본 논문에서는 다음과 같은 영역확장법을 사용하였다. 여기서 사용하는 영역확장법이란 간단한 모형에 따라 분할된 작은 영역을 보다 복잡한 모형의 관점에서 균일한 큰 영역으로 병합하는 과정을 말한다. 자세한 과정은 다음과 같다.

1. 병합을 위한 기준영역(seed region)을 찾는다.

아직 새로운 영역으로 할당되지 않은 영역들 중 가장 큰 영역을 기준영역으로 한다. 균일한 움직임을 갖는 영역 중 가장 큰 영역을 기준영역으로 선택하기 때문에 신뢰성이 큰 파라미터를 추정할 수 있다.

2. 기준영역 R_i 로 병합할 후보영역 R_j 를 찾는다.

여기서 후보영역(merging candidates)은 기준영역에 이웃한 영역 중 아직 새로운 영역으로 할당되지 않은 영역으로 한다.

3. 후보영역을 기준영역으로 병합할 것인지를 테스트한다.

후보영역 R_j 를 기준영역 R_i 로 병합할 것인지는 $R_i \cup R_j$ 에 대해 식(15)의 영역 균일성(region homogeneity)을 확인하여 결정한다.

$$P(R_i \cup R_j) = \begin{cases} \text{True if } R_i \cup R_j \text{ is homogeneous} \\ \text{False if } R_i \cup R_j \text{ is nonhomogeneous} \end{cases} \quad (15)$$

여기서 $P(R_i \cup R_j)$ 는 영역 $R_i \cup R_j$ 에 대한 논리 판단자(logical predicate)이다. 영역 $R_i \cup R_j$ 에 대하여 최소자승오차 방식(least mean square method)으로 최적 affine 변환을 계산한 다음, 계산된 변환계수로부터 후보영역의 오차를 구하여 이 오차로부터 병합 여부를 결정한다.

(a) 최적 affine 변환의 계산:

먼저 기준영역 R_i 와 후보영역 R_j 는 동일한 affine 변환을 갖고 $R_i \cup R_j$ 는 균일한 영역이라 가정한다. 영역 $R_i \cup R_j$ 에 대한 이동벡터로부터 최소자승오차 방식을 이용하여 최적의 affine 변환계수를 계산한다. 이때 최소화해야 할 오차는 다음과 같다.

$$E_{ij}(a_1, \dots, b_3) = \sum_k [(u(x_k, y_k) - a_1 - a_2x_k - a_3y_k)^2 + (v(x_k, y_k) - b_1 - b_2x_k - b_3y_k)^2]. \quad (16)$$

여기서 k 는 $R_i \cup R_j$ 의 원소이다. 따라서 이 오차를 변환계수 $\{a_1, a_2, a_3, b_1, b_2, b_3\}$ 에 대하여 편미분한 뒤 그 식을 0으로 두면 6개의 선형 방정식이 얻어진다. 이 방정식이 상호 독립이라면 그 해가 최적 affine 변환계수가 된다.

(b) 병합의 결정 (merging decision):

계산된 최적 변환계수 $\{a_1^*, a_2^*, a_3^*, b_1^*, b_2^*, b_3^*\}$ 와 후보영역 R_j 의 이동벡터를 이용하여 식(17)에 σ_j 를 계산한다.

$$\sigma_j = \sqrt{E_j(a_1^*, \dots, b_3^*) / \sum_j 1}. \quad (17)$$

σ_j 는 최적 변환으로부터 후보영역 R_j 의 이동벡터 분포의 표준편차를 의미한다. 병합 결정은 σ_j 를 기준으로 하고 다음처럼 논리 판단자가 결정된다.

$$P(R_i \cup R_j) = \begin{cases} \text{True if } \sigma_j \leq \text{threshold} \\ \text{False if } \sigma_j > \text{threshold} \end{cases} \quad (18)$$

따라서 $P(R_i \cup R_j)$ 가 논리적 참이면 후보영역 R_j 는 기준영역 R_i 로 병합된다. 기준영역은 후보영역을 병합함에 따라 점차 영역이 확대된다. 따라서 추정된 변환계수는 점점 신뢰성이 높아진다.

4. 기준영역 주위에 더 이상 병합할 후보영역이 없을 때까지 단계 2-3을 반복적으로 수행하게 된다. 기준영역 주위에 더 이상 병합할 후보가 없으면 새로운 기준영역을 찾아 병합을 수행하는 단계 1-4를 반복한다.

3.5. 세번째 분할단계

세번째 분할단계 (the third segmentation stage)는 평면체 영역을 3차원 곡면체의 움직임과 일치하는 영역으로 병합하는 단계이다. 이렇게 분할된 영역을 곡면체 영역(parabolic patch)이라 하자. 3차원 곡면체의

움직임은 식(13)의 quadratic 변환으로 표현되는 곡면체 운동모형에 의해 설명될 수 있다. 따라서 동일한 quadratic 변환을 갖는 평면체 영역을 영역확장을 통해 곡면체 영역으로 병합하는 것이다. 병합을 위해 사용하는 영역확장법은 앞단계에서 사용한 것과 동일하나 다만 추정하는 매핑 파라미터 모형이 affine 변환이 아니라 quadratic 변환을 이용한다. 기준영역 R_i 와 후보영역 R_j 의 병합 여부를 결정할 때 $R_i \cup R_j$ 에 대하여 최소자승오차 방식으로 최적 quadratic 변환을 계산한 다음, 계산된 변환계수로부터 후보영역의 오차를 구하여 이 오차로부터 병합 여부를 결정한다. 이때 최소화해야 할 오차는 다음과 같다.

$$E_{ij}(a_1, \dots, a_6, b_1, \dots, b_6) = \sum_k [(u(x_k, y_k) - a_1 - a_2x_k - a_3y_k - a_4x^2 - a_5x^2 - a_6xy)^2 + (v(x_k, y_k) - b_1 - b_2x_k - b_3y_k - b_4x^2 - b_5y^2 - b_6xy)^2]. \quad (19)$$

여기서 k 는 영역 $R_i \cup R_j$ 의 원소이다. 따라서 이 오차를 변환계수 $\{a_1, \dots, a_6, b_1, \dots, b_6\}$ 에 대하여 편미분한 뒤 그 식을 0으로 두면 12개의 선형 방정식이 얻어진다. 이 방정식이 상호 독립이라면 그 해가 최적 quadratic 변환계수가 된다.

3.6 고립된 작은 영역의 제거

이 단계의 목적은 앞단계에서 어느 기준영역으로도 병합되지 못한 고립된 작은 영역을 제거하는 것이다. 이와 같은 영역의 발생은 잘못된 이동벡터 추정에 기인한 것이다. 따라서 quadratic 변환과의 일치성 테스트는 의미가 없다. 여기서 고립된 작은 영역은 이웃한 곡면체 영역으로 병합함으로써 제거하였다. 최종적으로 움직임 영역은 몇 개의 곡면체 영역으로 분할 된다.

3.7 움직임 보상이 잘못된 영역의 추출

최종적으로 분할된 각 곡면체 영역은 추정된 quadratic 변환을 이용하여 움직임 보상이 가능하다. 따라서 원영상과 물체별 움직임 보상된 영상의 차신호에 대해 주어진 임계값을 이용하여 움직임 보상이 잘못된 영역을 추출한다. 이때 추출된 영역을 MF(motion fail) 영역이라 한다. 머리-어깨 영상에서 MF 영역은 주로 움직임 보상을 할 수 없는 입(mouth), 눈(eyes) 그리고 드러난 배경에서 발생한다. 따라서 이 MF 영역에

대해서는 움직임 보상을 할 수 없기 때문에 화소값을 부호화하여 전송해야 한다. MF영역의 판정시 사용하는 임계값에 따라 MF 영역의 발생량을 조절할 수 있으므로 영상 부호화 적용시 비트 발생량을 조절하는 파라미터로 사용할 수 있다.

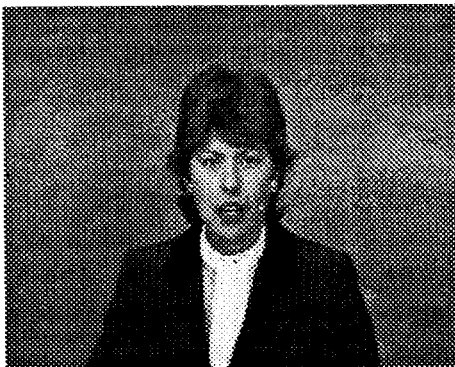
4. 모의실험 결과 및 고찰

본 절에서는 다단계 분할 알고리즘을 이용하여 모의실험한 결과를 제시하고 그 결과를 분석하였다. 모의실험에 사용된 test sequence는 "Claire" 영상(그림2)과 "Trevor" 영상(그림5(a))으로 머리와 어깨의 움직임을 갖는 대표적인 영상 전화용 영상이다. 이 영상들은 CIF(352*288)의 크기를 가지며 8비트로 양자화 되었다. 그리고 움직임 추정과 분할에는 밝기 성분(luminance)만을 사용하였다. 먼저 다단계 분할 알고리즘을 이용한 영상 해석(image analysis)과 영상 합성(image synthesis)결과를 제시한다. 그리고 그 결과를 Adiv[4]에 의해 제안된 방법과 비교하였다. 다단계 분할 알고리즘의 성능을 평가하기 위한 기준으로 분할된 이동물체 및 파라미터 갯수, MF 영역의 크기, 그리고 PSNR을 사용하였다. 이동물체 분할 결과, 전송해야할 맵핑 파라미터 수는 이동물체 갯수와 맵핑 파라미터의 복잡도와의 곱으로 표시된다. MF 영역은 움직임을 할 수 없는 영역이므로 이 영역에 대해서는 화소값을

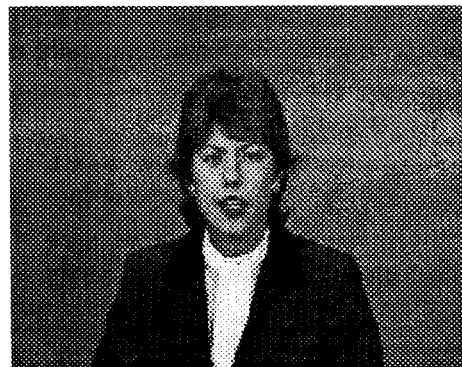
전송해야 한다. 따라서 MF 영역이 커지면 그 만큼 전송해야할 정보량이 많아진다. MF 영역의 크기를 전체 영상에 대한 MF영역의 비(%)로 표시한다. PSNR로 제안 알고리즘의 성능을 평가하기 위해 아래 3가지 경우를 비교하였다.

- PSNR 1: 연속된 두 원 영상 사이에 측정된 PSNR
- PSNR 2: 원 영상과 물체별 움직임 보상된 영상 사이에 측정된 PSNR
- PSNR 3: 원 영상과 물체별 움직임 및 MF 영역을 보상한 영상 사이에 측정된 PSNR

그림 3은 그림 2의 연속된 두 영상에 대해 다단계 분할 알고리즘을 적용한 결과이다. 그림 3(a)는 검출된 움직임영역과 움직임영역 내의 이동벡터를 나타낸다. 그림 3(b) - 그림 3(d)는 다단계 분할 알고리즘의 각 단계에서 분할된 결과이다. 그림 3(b)는 2차원 병진운동의 관점에서 균일한 움직임으로 분할된 결과이다. 그림 3(c)와 그림 3(d)의 각 영역은 각각 균일한 움직임을 갖는 평면체 및 곡면체 영역을 나타낸다. 그림 3(d)에서 알 수 있듯이 마지막 분할단계에서 움직임영역 전체가 3개의 곡면체 영역으로 분리되었다. 특히 사람의 얼굴이 하나의 곡면체 영역으로 분리됨을 알 수 있다. 분리된 3개의 곡면체 영역에 대해 12 파라미터로 표현되는 식(13)의 quadratic 변환을 사용하여 물체별 움직임 보상한 결과가 그림 3(e)이다. 그리고 움직임 보상이 잘못된 영역(MF region)을 그림 3(f)에서 검은색



(a)



(b)

그림 2. "Claire" sequence의 연속된 두 영상: (a) 첫번째 영상 (b) 두번째 영상.

Fig. 2. Two successive images from the "Claire" sequence: (a) the first image, (b) the second image

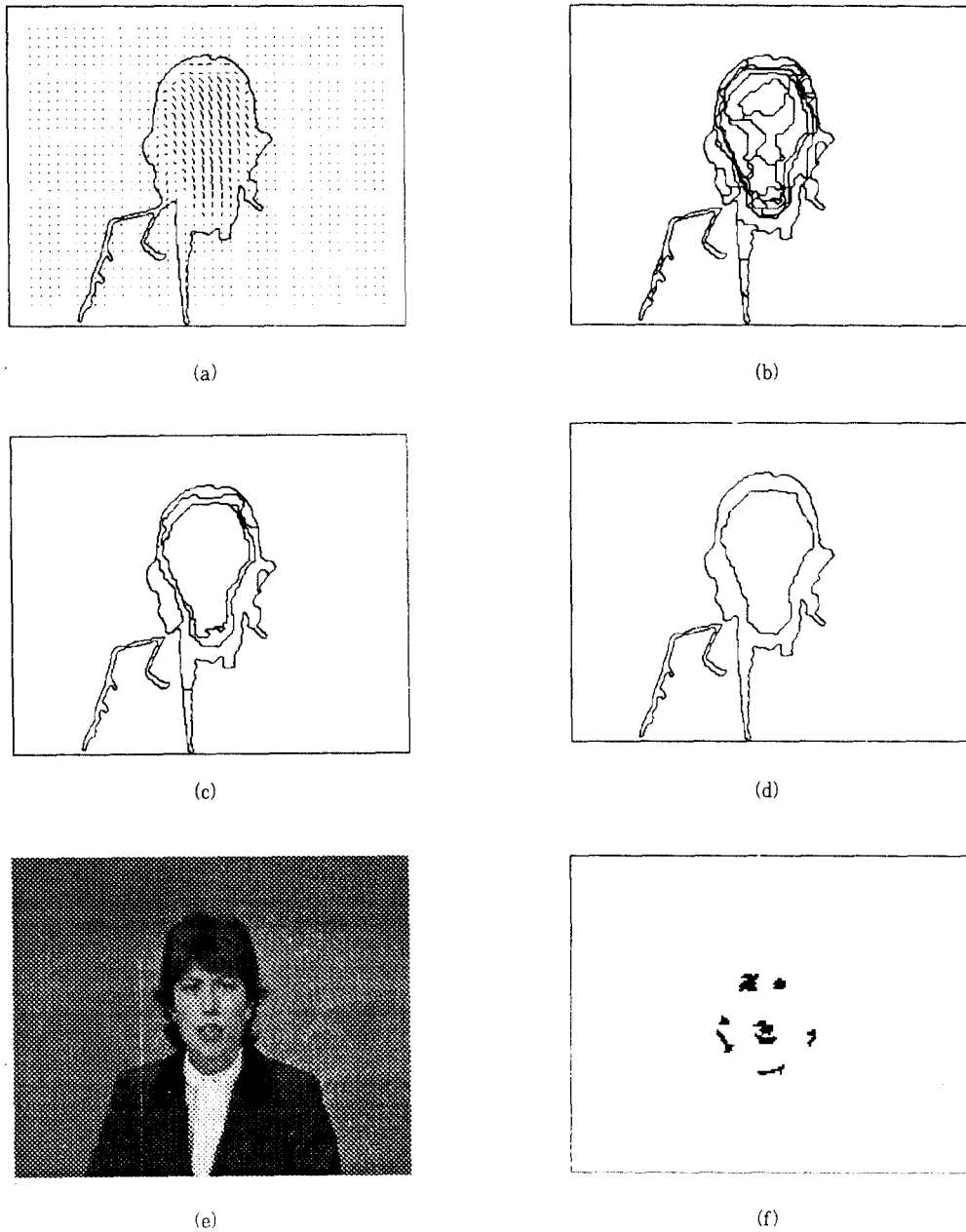


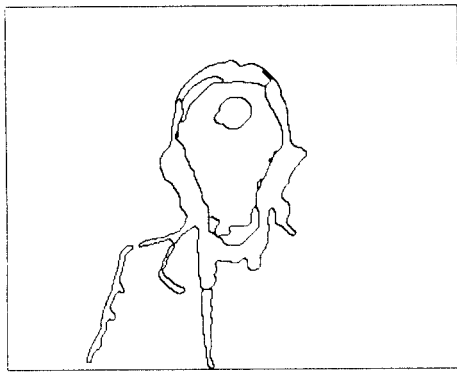
그림 3. 제안된 방식에 의한 모의실험 결과: (a) 움직임영역 및 이동 벡터장, (b) 첫번째 분할 결과, (c) 두번째 분할 결과, (d) 세번째 분할 결과, (e) 물체별 움직임 보상된 영상, (f) MF 영역.

Fig. 3. Results by the proposed method: (a) the changed region and motion vector field, (b) the result of the 1st segmentation, (c) the result of the 2nd segmentation, (d) the result of the 3rd segmentation, (e) reconstructed image by object-based motion compensation, (f) MF region

으로 표시하였다. 그림에서도 알 수 있듯이 MF 영역은 입, 눈, 그리고 드러난 배경의 일부에서 발생함을 알 수 있다. 그러나 MF 영역은 전체 영상에 비해 매우 작은 영역임을 알 수 있다. 움직임 추정 및 이동물체 분할 결과를 PSNR 관점에서 살펴보면 표 1에서도 알 수 있듯이 물체별 움직임 보상만으로 PSNR이 6.4dB(PSNR2 - PSNR1) 향상되었고 여기에 MF 영역에 대한 보상이 이루어질 경우 2.09dB(PSNR3 - PSNR2)가 향상됨을 알 수 있다. 따라서 물체별 움직임 보상으로 인한 PSNR 향상이 MF보상으로 인한 향상보다 훨씬 크므로 이동물체 분할 및 움직임 추정이 우

수함을 알 수 있다.

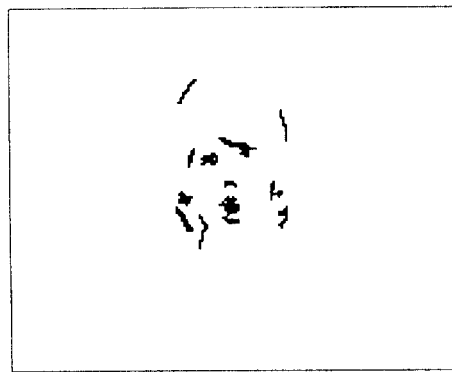
그림 4는 그림 2의 연속된 영상에 대해 Adiv가 제안한 방식(4)을 사용하여 처리한 결과이다. 여기서 사용한 이동 벡터장은 그림 3(a)와 동일한 것을 사용하였다. 그림 4(a)는 Adiv가 제안한 방식에 따라 분할된 결과이다. 여기서 각 영역은 집중투영의 가정하에 8파라미터로 표현되는 균일한 움직임을 갖는 3차원 평면체 영역을 나타낸다. 그림 4(b)는 분리된 각 영역별로 8파라미터로 표현되는 식(20)의 변환식을 사용하여 움직임 보상된 결과를 나타내고 MF 영역을 그림 4(c)에 검은색으로 표시하였다.



(a)



(b)



(c)

그림 4. Adiv 방식에 의한 모의실험 결과: (a) 분할 결과, (b) 물체별 움직임 보상된 영상, (c) MF 영역.
Fig. 4. Results by Adiv's method: (a) the segmentation result, (b) reconstructed image by object-based motion compensation, (c) MF region.

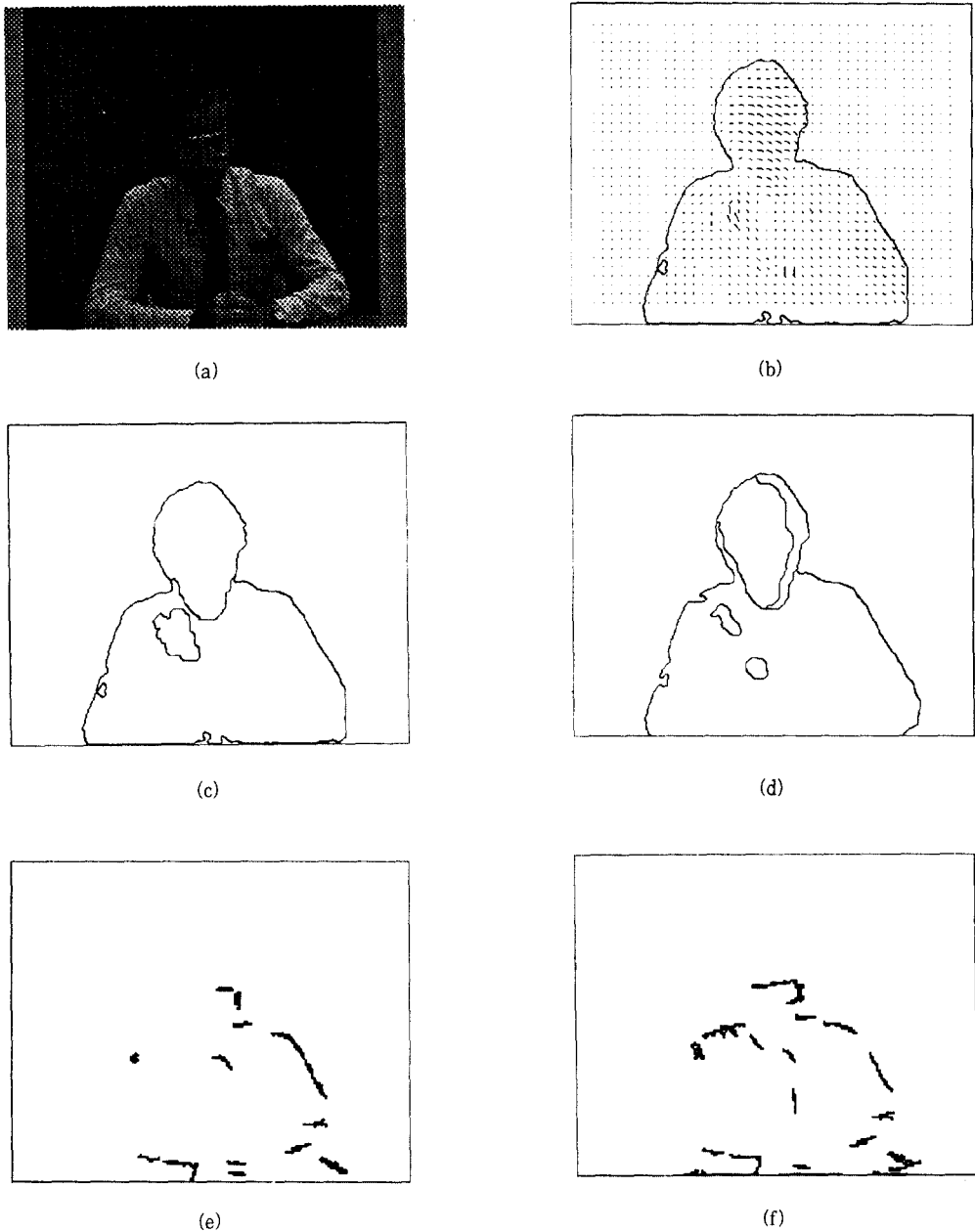


그림 5. Trevor 영상에 대한 모의실험 결과: (a) 원 영상, (b) 움직임 영역 및 이동 벡터장, (c) 제안 방식에 의해 최종 분할된 결과, (d) 기존 방식에 의해 최종 분할된 결과, (e) 제안 방식에 의해 추출된 MF 영역, (f) 기존 방식에 의해 추출된 MF 영역.

Fig. 5. Results for the Trevor image: (a) original image, (b) optical flow field, (c) segmentation result by the proposed method, (d) segmentation result by the conventional method, (e) MF parts by the proposed method, (f) MF parts by the conventional method.

표 1. 제안 방식과 Adiv 방식의 비교.

Table 1. The comparison of the proposed method and Adiv's method

(M1: Proposed method; M2: Adiv's method)

image	item	# of parameters	ratio of	PSNR1	PSNR2	PSNR3
			MF(%)	(dB)	(dB)	(dB)
"Claire"	M1	24	0.848	29.648	36.069	38.157
	M2	56	1.069	29.648	35.639	38.374
"Trevor"	M1	36	1.207	28.873	33.481	35.352
	M2	40	1.866	28.873	33.239	35.421

$$u(x, y) = a_1 + a_2x + a_3y + a_7x^2 + a_8xy$$

$$v(x, y) = a_4 + a_5x + a_6y + a_7xy + a_8y^2. \quad (20)$$

다양한 영상에 대한 실험 결과를 보이기 위해 "Trevor" 영상에 대한 결과를 그림 5에 나타내었다. 그림 5(a)는 "Trevor" sequence의 3번째 원 영상이고 그림 5(b)는 "Trevor" sequence의 2번째 영상과 3번째 영상에 대한 움직임영역 및 이동 벡터장을 나타낸다. 그리고 그림 5(c)와 그림 5(d)가 각각 제안된 알고리즘과 기존방식(Adiv 방식)에 의해 최종 분할된 결과이다. 그림에서 알 수 있듯이 제안된 방식에서는 움직임영역이 머리와 어깨로 분리되었다. 제안된 방식과 기존방식으로 각각 물체별 움직임 보상을 한 뒤 MF 영역을 추출한 결과가 그림 5(e)와 그림 5(f)이다. 제안된 방법에 의해 추출된 MF 영역이 기존 방식에 의해 추출된 MF 영역보다 훨씬 작음을 알 수 있다. 따라서 움직임영역 분할 및 매핑 파라미터 추정이 우수함을 알 수 있다.

본 논문에서 제안한 다단계 분할 방식과 Adiv가 제안한 방식을 표 1에 나타난 항목별로 비교하였다. 분리된 이동물체 갯수와 이에 따른 파라미터 갯수 면에서 제안된 방식이 훨씬 우수함을 알 수 있다. 그리고 MF 영역의 경우도 제안된 방식이 Adiv 방식보다 훨씬 적음을 알 수 있다. 필요한 메모리 및 계산량의 경우 정량적인 비교는 힘들지만 알고리즘의 특성에서 알 수 있듯이 영역확장을 사용하는 본 논문의 제안 방식이 변형된 Hough 변환을 사용하는 Adiv 방식보다 훨씬 적다.

5. 결 론

본 논문에서는 물체별 부호화를 위한 이동물체의 분할 및 움직임 추정 방법을 제안하였다. 이 방법은 간단한 매핑 모형 관점에서 균일한 작은 영역을 보다 복잡한 매핑 모형 관점에서 균일한 영역으로 영역확장을 하는 것에 근거하고 있다. 머리-어깨 영상에 대한 모의실험 결과 움직임영역 전체가 몇 개의 곡면체 영역으로 분할되었다. 특히 사람의 얼굴이 하나의 곡면체 영역으로 분할됨을 알 수 있다. 추후 연구는 제안된 움직임 물체 분할 및 움직임 추정을 물체별 부호화에 응용하는 것에 초점이 맞추어질 것이다. 물체별 부호화의 효율은 한 매핑 파라미터 세트에 표현될 수 있는 물체의 크기에 달려있기 때문에 높은 부호화 효율을 기대할 수 있다.

참고문헌

1. H. G. Musman, H. tter and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images," Signal Processing: Image communication, vol. 1, no. 2, pp.117-138, Oct., 1989.
2. M. Hotter and R. Thoma, "Image segmentation based on object oriented mapping parameter estimation," Signal Process., vol. 15, no. 3, pp.315-334, Oct., 1988.
3. N. Diehl, "Object oriented motion estimation

and segmentation in image sequences." Signal Processing: Image Communication, vol. 3, no. 1, pp.23-56, Feb., 1991

4. G. Adiv, Determining three-dimensional motion and structure from optical flow generated by several moving objects." IEEE. Trans. PAMI, vol. 7, no. 4, pp.384-401, July, 1985.

5. R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch." IEEE. Trans. ASSP, vol. 29,

no. 6, pp.1147-1152, Dec., 1981.

6. J. G. Choi, S. W. Lee and S. D. Kim, "Segmentation and motion estimation of moving objects for object-oriented coding." ICAS-SP'95 Proceedings, pp.2431-2434, Detroit, May, 1995.

7. B. K. P. Horn, Robot vision. The MIT Electrical Engineering and Computer Science Series, pp.19-22, 1987.



崔在覺(Jae Gark Choi) 정희원

1961년 5월 9일생
 1984년 : 경북대학교 전자공학과(공학사)
 1987년 : 한국과학기술원 전기 및 전자공학과(공학석사)

1992년~현재 : 한국과학기술원 전기 및 전자공학과 박사과정
 1987년~현재 : 한국전자통신연구원
 ※주관심 분야 : 영상처리, 영상통신등



李時雄(Si Woong Lee) 정희원

1969년 1월 10일생
 1991년 : 경북대학교 전자공학과(공학사)
 1993년 : 한국과학기술원 전기 및 전자공학과(공학석사)

1993년~현재 : 한국과학기술원 전기 및 전자공학과 박사과정
 1995년~현재 : 삼성전자 연구원
 ※주관심 분야 : 영상처리, 영상통신등



金聖大(Seong Dae Kimi) 정희원

1953년 12월 26일생
 1977년 : 서울대학교 전자공학과(공학사)
 1979년 : 한국과학기술원 전기 및 전자공학과(공학석사)

1983년 : 프랑스 INPT ENSEEIHT(공학박사)
 1984년~현재 : 한국과학기술원 전기 및 전자공학과 교수
 ※주관심 분야 : 영상처리, 영상통신, 컴퓨터 비전, VLSI 구현등



李炳旭(Byung Wook Lee) 정희원

1975년 3월~1979년 2월 : 서울대학교 전자공학과 학사
 1979년 3월~1981년 8월 : 한국과학기술원 전기 및 전자공학과 석사

1981년 3월~1985년 8월 : 대우전자 중앙연구소 Video 개발부
 1985년 8월~1991년 6월 : Stanford University, Dept. of Electrical Engineering, 박사
 1991년 7월~현재 : 대우전자 영상연구소
 ※주관심 분야 : Computer vision image processing, computer graphics, digital signal processing, image compression