論文94-31B-8-10

물체지향 분석 및 합성 부호화에서 가산투영을 이용한 영상분석기법

(An image Analysis Technique Using Integral Projections in Object-Oriented Analysis-Synthesis Coding)

金 俊 植*, 朴 來 弘**

(Joon-Seek Kim and Rae-Hong Park)

要 約

물체지향 분석 및 합성 부호화는 영상내의 움직이는 물체를 분할하여 각 물체단위로 움직임을 검출/보상하는 방법으로, 저전송율로 동영상 전송시 기존의 움직임 검출/보상방법들보다 자연스러운 움직임을 복원할 수 있다. 기존의 물체지향 분석 및 합성 부호화에서 각각의 물체에 대한 움직임을 검출할 때 변환 파라메터방법을 사용하는데, 이때 gradient 연산자를 사용하기 때문에 잡음이나 빛의 미세한 변화에 민감한 결과를 초래한다.

본 논문에서는 좀 더 정확한 변환 파라메터를 구하기 위해 gradient 계산시 가산투영을 사용한 새로운 영상분석 방법을 제안하였다. 또한 물체의 국부적인 움직임을 정확히 검출하기 위해 국부적인 움직임을 갖는 물체단위로 영역분할을 수행하여, 복잡한 물체를 하나의 물체로 간주할 때 발생하는 움직임 검출시의 오차를 줄였다. 여러 영상에 대한 실험결과를 볼 때 제안한 방법은 움직임이 복잡하거나 물체가 복잡한 형태를 가질수록 존의 방법보다 성능 향상이 커진다.

Abstract

Object-oriented analysis-synthesis coding subdivides each image of a sequence into moving objects and compensates the motion of each object. Thus it can reconstruct real motion better than conventional motion-compensated coding techniques at very-low-bit-rates. It uses a mapping parameter technique for estimating motion information of each object. Since a mapping parameter technique uses gradient operators, it is sensitive to redundant details and noise.

To accurately determine mapping parameters, we propose a new analysis method using integral projections for estimation of gradient values. Also to reconstruct correctly the local motion, the proposed algorithm divides an image into segmented objects, each of which having uniform motion information, while the conventional one assumes a large object having the same motion information. Computer simulation results with several test sequences show that the proposed image analysis method in object-oriented analysis-synthesis coding shows better performance than the conventional one.

* 正會員, 湖西大學校 電子工學科
(Dept. of Elec. Eng., Hoseo University)
** 正會員, 西江大學校 電子工學科
(Dept. of Elec. Eng., Sogang University)

※본 연구는 한국과학재단의 특정연구비 지원으로 이루어졌음.
接受日字 : 1993年 10月 6日

Ⅰ. 서 론

정보를 전송하고 받는 일은 우리 생활과 매우 밀접한 관계를 맺고 있다. 그중 영상정보는 음성정보보다 더 많은 내용의 정보전달 효과가 있으나 음성정보에 비해 전송대역폭이 큰 단점을 가지므로 한정된 채널을 사용하여 정보를 전달하는데 큰 어려움이 있다. 그러나 최근 영상처리 및 반도체 기술의 발전에 힘입어 많은 양의 영상정보를 효율적으로 전송할 수 있게 되었다. 다시 말해 압축기술의 발전으로 영상정보를 제한된 전송 대역폭에 맞게 압축하여 전송하는 것이 가능해졌다. 이러한 영상정보 압축기술은 현재도 많은 연구가 진행중이며 많은 부분이 표준화되고 있다.

프레임간 부호화 (interframe coding) 방법은 시간축상에서 연속적으로 변하는 영상신호를 압축하여 전송하는데 매우 효율적인 방법이다.[1,2] 특히, 움직임 보상 부호화 (MCC: Motion Compensated Coding) 방법은 연속한 영상내에서 움직임을 검출/보상하여 부호화하는 방법으로 높은 압축률과 고화질의 영상을 얻을 수 있다. 움직임 보상 부호화방법은 크게 두가지 부분으로 구성되어 있는데 하나는 영상내에 존재하는 움직임을 검출/보상하는 움직임 검출부와 움직임 예측에 의해 발생하는 예측오차를 부호화하는 부분이다. 예측오차 전송부의 성능은 그 부분 나름대로의 방법에 의해 성능이 결정되지만 움직임 검출부의 성능에 의해서도 많은 영향을 받는다. 즉, 영상내에 존재하는 움직임을 정확히 검출하면 예측오차가 줄어들 수 있기 때문에 정확한 움직임 검출은 상당히 중요하다.

움직임 검출/보상 방법은 크게 두가지로 나눌 수 있는데 첫째는 화소단위의 움직임을 검출하여 보상하는 화소순환 알고리듬 (PRA: Pel Recursive Algorithm)이며, 둘째는 영상을 일정한 크기의 블럭으로 나눈 후 각 블럭에 대한 움직임을 검출하여 움직임을 보상하는 블럭정합 알고리듬 (BMA: Block Matching Algorithm)이다. 전자는 비교적 정확한 움직임을 검출할 수 있지만 과다한 계산량으로 인해 실시간 처리에 많은 어려움이 있으며, 후자는 전자에 비해 화질은 조금 떨어지지만 계산량이 전자에 비해 월등히 적어 실시간 처리가 보다 용이하여 영상회의, 영상전화기, HDTV (High Definition TV) 등과 같은 많은 응용분야에서 사용되고 있다.

BMA는 영상을 일정한 크기의 블럭단위로 분할하여 처리하므로 전송률이 낮아질 경우 복원된 영상의 인접한 블럭들 사이의 불연속과 같은 blocking 효과나 mosquito 효과가 발생하여 시각적으로 심각한 화질저하를 초래한다. 이러한 화질저하를 극복하기 위해서는 영상내에 존재하는 물체단위로 움직임을 검출, 보상하여 BMA보다 시각적으로 우수한 화질의 영상을 복원할 수 있으며 전송정보량면에서도 더 많은 정보를 압축할 수 있다. 이러한 부호화 방법이 물체지향 분석 및 합성 (object oriented analysis synthesis) 부호화방법으로 영상을 각각의 물체로 영역분할하여 물체단위로 움직임 정보와 형태정보 그리고 밝기정보를 전송한다.

본 논문에서는 물체지향 분석 및 합성 부호화 기법에 대해 연구하였으며 물체지향적 분석 및 합성 부호화기법에서 움직이는 물체의 이동정보를 구할 때 가산투영기법을 적용하여 기존의 방법보다 잡음에 강하며 좀 더 정확한 움직임을 검출할 수 있는 방법을 제안하였다. 또한 기존 방법의 문제점을 보완하기 위하여 영역분할방법을 사용하여 좀 더 세밀하게 물체를 분할한 후 움직임을 검출하는 방법을 제안하였다.

본 논문의 구성은 Ⅱ장에서 가산투영의 기본개념 및 수학적 배경을 설명하고, Ⅲ장에서 물체지향 분석 및 합성 부호화기법에 대해 간단히 알아본다. Ⅳ장에서는 기존의 물체지향 분석 및 합성 부호화에서의 영상분석방법에 대해 알아보고, Ⅴ장에서는 가산투영과 영역분할을 사용한 제안한 영상분석기법에 대해 설명하고 Ⅵ장에서 실험결과에 대해 분석하고 Ⅶ장에서 결론을 맺는다.

Ⅱ. 가산투영의 기본개념

영상을 2차원 공간 함수 $f(x, y)$로 나타낼 때 투영의 개념을 그림 1에 보였다. 여기서 영상을 지나는 선 t를 투영선 (projection ray)이라 하고 투영선상

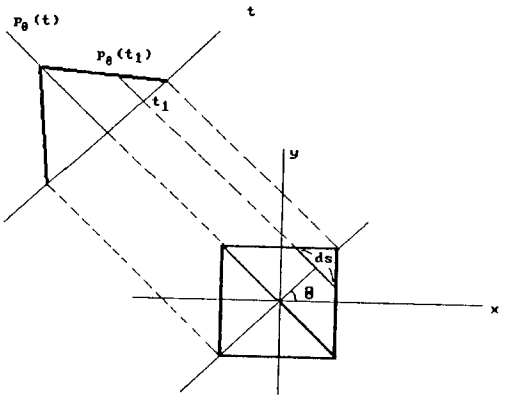


그림 1. 가산투영의 개념
Fig. 1. Concept of integral projections.

의 함수 $f(x, y)$의 적분을 투영선적분 (projection line integral)이라 부른다. 투영선 t와 원점으로부터 공간상의 좌표 (x, y)에 그은 선분과 이루는 각도를 θ 라고 하면 함수 $f(x, y)$의 투영선 t에 대한 투영선적분 [3] 은

$$P_\theta(t_1) = \int_{ray\ t_1} f(x,y)ds = \iint f(x,y)\delta(x\cos\theta + y\sin\theta - t_1)dx\,dy \quad (1)$$

와 같다. 여기서 $P_\theta(t_1)$를 각도 θ 에 대한 가산투영이라고 정의한다. 주어진 영상에 대한 가산투영은 어느 한 방향으로 존재하는 화소값을 더한 합으로 정의되며, 수직 (수평)방향으로 더해진 1차원 정보를 수직 (수평) 가산투영이라 한다. [4-6]

영상신호 $f(x, y)$에 대한 x축 방향 (θ = 0)의 1차원 가산투영 $P_0(y)$가

$$P_0(x = t_1) = \int_{-\infty}^{\infty} f(x,y)dx \quad (2)$$

와 같이 주어지며, $f(x, y)$에 대한 2차원 Fourier 변환이

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+uy)}dx\,dy \quad (3)$$

와 같이 주어지고, $f(x, y)$에 대한 가산투영 $P_\theta(t)$의 1차원 Fourier 변환이

$$S_\theta(w) = \int_{-\infty}^{\infty} P_\theta(t)e^{-j2\pi wt}dt \quad (4)$$

와 같을 때, uv-평면상에서의 u = 0인 선상의 $F(u, v)$를 고려하면

$$F(0,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,e^{-j2\pi vy}dx\,dy = \int P_0(y)\,e^{-j2\pi vy}dy = S_0(v) \quad (5)$$

이 된다. 위의 결과로부터 u = 0인 선상에 존재하는 2차원 Fourier 변환값 $F(u, v)$는 영상신호의 수평 가산투영의 1차원 Fourier 변환에 의해 얻어질 수 있다. [3]

가산투영연산은 적분형태의 연산이므로 가산형 잡음에도 강한 특성을 지니고 있다. 원래의 영상신호를 $f(x, y)$라 하고 평균이 0이고 표준편차가 σ 인 가산형 가우스 잡음을 $n(x, y)$라 하면 잡음에 왜곡된 신호 $f_n(x, y)$는

$$f_n(x,y) = f(x,y) + n(x,y) \quad (6)$$

와 같이 주어지며, $f_n(x, y)$에 대한 가산투영값을 구해보면

$$g(y) = \int f_n(x,y)dx = \int f(x,y)dx + \int n(x,y)dx \quad (7)$$

가 되어 두번째 항에 대한 결과는 평균이 0인 가우스 잡음에 대한 적분이므로 그 결과는 0에 가까운 값을 갖게된다. 따라서 가산형 잡음이 첨가된 신호의 가산투영값은 원래 원래 신호의 가산투영값과 비슷하다고 할 수 있다. 이와 같은 결과로 볼 때 가산투영 연산은 이산 영상에 적용될 수 있으며, 가산투영을 사용한 움직임 검출방법이 잡음에 강한 특성을 갖는다는 것이 참고문헌 [4-6] 에서 보여졌다.

Ⅲ. 물체지향 분석 및 합성 부호화 방법

물체지향 분석 및 합성 부호화방법은 contour-texture 영상부호화 개념을 동영상으로 확장한 것으로, 영상내에 존재하는 움직이는 물체를 추출하여 각 물체의 움직임 정보와 형태정보 및 밝기정보를 전송하기 때문에 복원된 영상에서 물체의 움직임이 자연스러우며, 영상내에 존재하는 몇개의 물체들에 대해서만 부호화하기 때문에 전송정보량면에서도 많은 감축을 할 수 있기 때문에 저전송율의 동영상 전송시 적합한 방법이다.

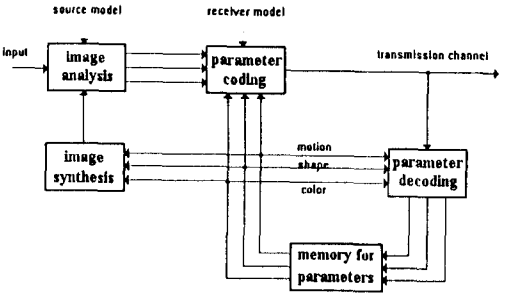


그림 2. 물체지향 분석 및 합성 부호화의 블럭도
Fig. 2. Block diagram of object-oriented analysis-synthesis coding.

물체지향 분석 및 합성 부호화 방법은 연속영상의 각 영상을 움직이는 물체 단위로 분할한 후 각 물체의 이동정보와 형태정보 및 밝기정보를 묘사하는 파

라메터들을 전송하는 방법으로 각 물체에 대한 이동정보와 형태정보 그리고 밝기정보가 수신단으로 전송된다.[7] 수신단에서는 이들에 대한 정보를 받아 영상을 복원한다. 그림 2는 물체지향 분석 및 합성 부호화 방법의 블럭선도이다. 그림에서 영상분석부분은 입력영상을 각각의 물체단위로 분할하여 이전영상에 대한 합성영상을 사용해 현재영상내에 있는 물체의 움직임 정보와 형태정보 그리고 밝기정보를 추정하는 부분이며, 파라메터 부호화 (역부호화)부분은 추정된 파라메터들을 각각의 정보에 적합하게 부호 (역부호)화하는 부분이다. 또한 영상합성부분은 검출된 파라메터들을 사용하여 영상을 합성하는 부분으로 현재영상에 대한 파라메터들을 추출하기 위한 이전영상의 합성된 영상을 저장하고 있는 부분이다.

앞에서 설명한 물체지향 분석 및 합성 부호화 개념에 따라 영상내의 물체는 각 물체의 움직임 정보, 형태정보, 밝기정보를 나타내는 파라메터들에 의해 묘사된다. 이때 수신단으로 전송되는 정보량은 물체나 물체의 움직임을 묘사하는 모델을 어떤 것을 사용하느냐에 따라 달라진다.

물체지향 분석 및 합성 부호화에서 가장 중요한 부분이 영상분석부분이며 이 부분은 영상내의 움직이는 물체를 분리하여 각 물체의 묘사 파라메터를 검출하는 역할을 한다. 움직이는 물체를 검출하고 각 물체의 속도를 추정하는 여러가지 방법들이 제안되었지만 이들 대부분은 optical flow에 의해 이동정보를 검출하기 때문에 물체의 경계와 움직임 사이의 의존성이 모호하다는 것이 단점이다. 물체의 정확한 움직임 검출은 물체의 경계와 상관성을 가지고 있다. 다시 말해 정확한 움직임 검출은 물체의 경계에 대한 정보의 정확한 검출이 사전조건으로 보장되어야 하며, 이와 반대로 정확한 물체경계검출은 정확한 움직임 묘사가 필수적이다. 즉, 움직임 검출과 물체경계의 검출은 서로 연관성이 있으며 서로에게 영향을 미친다.

물체의 움직임을 묘사하는 모델에는 크게 두가지가 있다. 하나는 3차원 공간상의 움직임을 2차원의 평면 강체로 묘사하는 모델이며, 또 다른 하나는 3차원 공간상의 움직임을 3차원 강체로 묘사하는 모델이다.[8] 전자는 영상평면상에서 추출한 특징들의 대응여부를 조사하여, 강체 움직임 등과 같은 가정으로부터 유도되는 제한조건을 만족하는 해를 구하여 이동정보를 검출하는 방식이며, 후자는 영상평면상의 모든 화소들에 대해 2차원 속도를 구하고 부가적인 제한조건을 사용하여 3차원 이동정보를 계산하는 방식이다.[8] 후자의 경우 3차원 움직임을 묘사하므로 비교적 실제의 움직임과 유사한 이동정보를 검출할 수 있으나 계산량이 방대하여 동영상 부호화에 응용하기에는 어렵다. 따라서 본 논문에서는 전자의 방식에 속하는 변환 파라메터[7]를 사용한 이동정보 검출모델을 사용한 방식에 대해 연구하였다.

Ⅳ. 기존의 영상분석기법

기존의 영상분석방법은 그림 3에 보인 것과 같이 계층적 구조를 가지며, 변환 파라메터를 계산하는 부분과 구해진 변환 파라메터에 의해 영상을 합성하는 부분, 그리고 구해진 변환 파라메터들이 정확한지를 검사하는 부분의 세 부분으로 구성되어 있으며[7] 연속 영상에서의 움직이는 물체는 8개의 변환 파라메터[9-11]에 의해 묘사될 수 있고, 효율적인 영상분해 및 합성을 위해서는 정확한 변환 파라메터 추정이 필요하다.

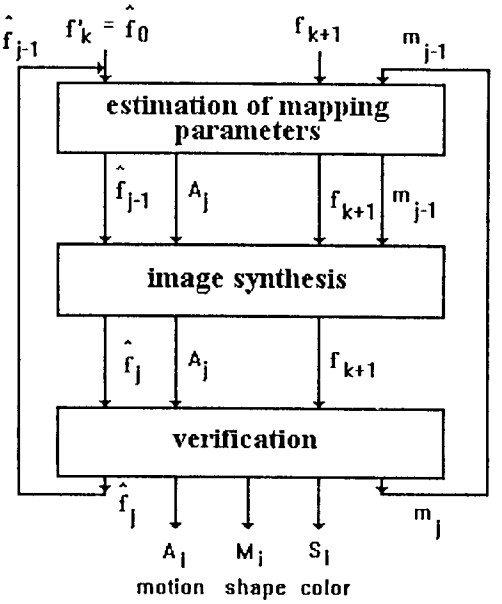


그림 3. 기존의 영상분석 방법 (f_{k+1}, (k+1)번째 영상; $\hat{f}_j$, f_{k+1}을 묘사하는 합성된 영상; m_j, 분석될 영역; 분석된 물체의 번환 파라메터; j, 계층의 인덱스; i, 물체의 인덱스)

Fig. 3. Conventional image analysis method (f_{k+1}, (k+1)th luminance image; $\hat{f}_j$, synthesized image describing f_{k+1}; m_j, areas to ble analyuzed; a, mapping parameters of analyzed areas; j, index of hierarchy; i, index of an object).

변환 파라메터를 계산하는 부분에서는 인접한 두장의 영상으로부터 차영상을 구해 움직임이 있는 부분에 대해 앞절에서 설명한 변환 파라메터를 사용하여 이동정보를 구하며, 이때의 변환 파라메터를 사용하여 영상합성부에서 합성영상을 만들어 낸다. 모델 검증부에서는 합성된 영상과 원영상의 오차를 계산해 정확한 변환 파라메터가 구해졌는지를 판단하여 구해진 변환 파라메터가 적합한지를 검사하여 그 물체에 적합한 정보를 전송한다.

연속적인 영상들 중 두장의 인접한 영상이 입력으로 주어지며, 초기의 가정인 영상내에는 하나의 물체만이 존재하며 움직임은 없다는 전제하에 모델검증을 한다. 모델검증시 영상내에 움직임이 존재하지 않는 부분은 초기의 가정에 부합하는 부분이 되며, 움직임이 존재하는 부분은 가정에 부합되지 않은 부분이 된다. 즉, 두영상의 차신호가 임의의 임계치 T_c 보다 작은 부분은 초기에 가정된 모델에 의해 정확한 변환 파라메터가 구해진 부분이며, 차신호가 임계치 T_c 보다 큰 부분은 가정된 모델이 적합하지 않은 부분이 된다. 이러한 부분은 하나의 물체로 간주하여 다음 단계에서 새로운 모델가정에 의해 변환 파라메터가 구해지며 다시 검증부를 통해 모델검증이 이루어 진다.

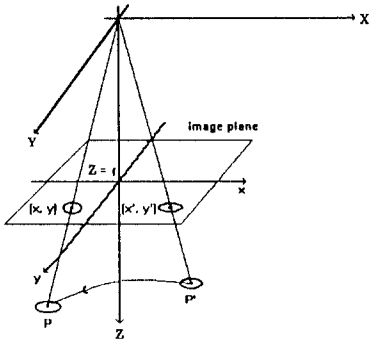


그림 4. 3차원 움직임 추정에 대한 기본적인 기하형태

Fig. 4. Basic geomety for three-dimensional motion estimation.

변환 파라메터에 의한 이동검출 방식은 연속한 두장의 2차원 영상내의 각 화소간의 대응관계를 계산하여 이들사이의 이동정보를 검출하는 것으로, 화소간의 대응관계는 8개의 변환 파라메터에 의해 표현된다. 동영상에서 연속되는 영상간의 시각적 변화는 전적으로 영상내에 있는 물체의 움직임으로부터 기인된다고 볼 수 있으며, 3차원 물체공간상에서의 이동물체 각 점의 실제 밝기는 변화가 없다고 가정하면 이동정보의 계산이 간단해진다. 이때 움직임에 따른 영상내에서의 밝기의 변화를 해석하기 위해 3차원 물체공간을 2차원 영상평면으로 투영하는 투영모델이 필요하며 물체의 움직임을 묘사하는 3차원 이동모델 그리고 물체의 표면형태를 나타내는 물체에 대한 모델등이 필요하다.

카메라로부터 취득한 영상은 일반적으로 central projection에 의해 설명될 수 있다. 그림 4에서와 같이 물체내의 어떤 점 P를 고려할 때 (X, Y, Z)는 t_1 시각에서의 점 P의 물체공간상에서의 좌표이며, (X', Y', Z')는 t_2 시각에서의 점 P의 물체공간상에서의 좌표라 가정하고 카메라에 의해 영상평면으로 투영된 좌표를 각각 (x, y)와 (x', y')이라고 가정하면 각 좌표간의 관계는

$$(x', y') = \left(\frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1}, \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1} \right) \tag{8}$$

로 표현될 수 있으며, 이때 a_i, $(1 \le i \le 8)$를 8개의 mapping 파라메터라 하고, 식 (8)의 변환을 mapping 변환[8] 이라 한다.

인접한 영상사이의 물체 움직임을 변환 파라메터에 의해 묘사하는 수식에 대해 설명하면 다음과 같다. 영상내의 한점 (x, y)가 다음 영상에서 (x', y')인 점으로 이동하였다고 가정하면 이때 물체의 움직임에 의한 빛의 변화와 영상사이의 차를 mapping 변환에 의해 나타내면

$$\begin{aligned} FD(x,y) &= f_{k+1}(x,y) - f_k(x,y) \\ &= f_k(x',y') - f_k(x,y) \\ &= f_k\left(\frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1}, \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1} \right) - f_k(x,y) \end{aligned} \tag{9}$$

로 나타낼 수 있다. 여기서 $FD(\cdot, \cdot)$는 차영상을 나타내며, $f_{k+1}(\cdot, \cdot)$와 $f_k(\cdot, \cdot)$는 $(k+1)$번째 영상과 k번째 영상을 나타낸다. 즉, 움직이는 화소를 묘사하는 방정식[8] 은

$$\begin{aligned} FD(x,y) &= G_x x \Delta a_1 + G_x y \Delta a_2 + G_x \Delta a_3 + G_y x \Delta a_4 \\ &\quad + G_y y \Delta a_5 + G_y \Delta a_6 - x(G_x x + G_y y)\Delta a_7 \\ &\quad - y(G_x x + G_y y)\Delta a_8 \\ &\equiv H \cdot \Delta a \end{aligned} \tag{10}$$

로 표시할 수 있다. 여기서 $a = (a_1\text{-}1, a_2, a_3, a_4, a_5\text{-}1, a_6, a_7, a_8)$로 변환 파라메터의 벡터이며, G_x

(x, y)와 $G_y(x, y)$는 국부 gradient로

$$G_x = \frac{1}{2}\left\{\frac{\partial f_{k+1}(x,y)}{\partial x} + \frac{\partial f_k(x,y)}{\partial x}\right\} \quad (11)$$

와

$$G_y = \frac{1}{2}\left\{\frac{\partial f_{k+1}(x,y)}{\partial y} + \frac{\partial f_k(x,y)}{\partial y}\right\} \quad (12)$$

로 주어진다.

영상합성부분에서는 변환 파라메터 추정부분에 의해 얻어진 변환 파라메터들을 사용하여 영상을 합성하며, 이들에 대한 정확성의 판단은 검증부에서 이루어 진다.[7] 검증부에서의 정확성 판단기준은 평균 자승 DFD (mean squared displaced frame difference)를 사용하며 물체의 실제 움직임과 추정된 움직임 사이의 오차의 최소화가 이루어졌는지를 검증하게 된다. 이때 검증부를 통과하지 못한 물체들은 다음 단계에서 분석의 대상이 되어 변환 파라메터가 새로 구해지게 된다.

Ⅴ. 제안한 영상분석방법

기존의 영상분석방법[7]에는 두가지 문제점이 있다. 첫째는 변환 파라메터를 계산할 때 영상의 화소값을 사용하기 때문에 식 (11)과 (12)의 x축 방향과 y축 방향의 gradient 계산시 잡음이나 조그마한 빛의 변화에 민감한 결과를 초래하여 실제의 움직임과 다른 이동정보를 검출하게 된다는 단점이 있으며, 둘째로 인접한 영상간의 차신호가 존재하는 인접해 있는 영역들을 하나의 물체로 간주하여 움직임을 검출하기 때문에 변환 파라메터를 이용한 움직임 검출시 부정확한 움직임 검출이 많다. 본 연구에서는 이러한 문제점들을 해결하기 위해 가산투영기법과 영역분할 방법을 사용하여 기존의 물체지향 분석 및 합성 부호화방법의 단점들을 개선하였다.

그림 5는 제안한 물체지향 분석 및 합성 부호화 방법의 블럭선도이다. 가산투영변환부에서는 입력영상을 변환하여 수평가산투영 데이타와 수직가산투영 데이타를 얻으며, 물체 분할부에서는 이전영상과 입력영상사이의 차신호가 임의의 임계값 T_c보다 큰 값을 갖는 영역을 검출하여 두 영상사이의 움직임이 존재하는 부분과 움직임이 없는 배경부분을 분리하며, 분리된 영역의 물체를 세분화하여 하나의 움직임을 갖는 강체로 분리한다. 변환 파라메터 계산부에서는 가산투영 데이타와 움직임 물체 분할부에 의해 얻어진 강체의 움직임 영역의 정보를 사용하여 움직임이 있

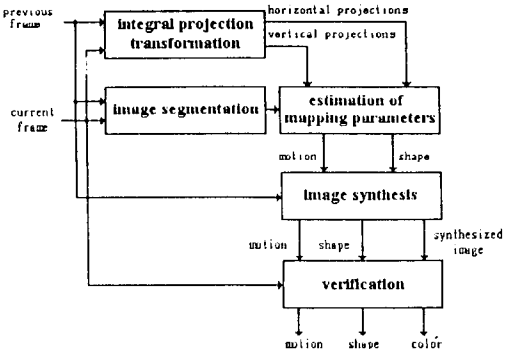


그림 5. 제안한 영상분석방법
Fig. 5. Proposed image analysis method.

는 각각의 물체에 대한 변환 파라메터를 계산한다. 계산된 변환 파라메터들은 기존의 방법과 유사하게 영상 합성부에 의해 합성된 영상을 만들고 검증부를 통해 변환 파라메터들의 정확성을 판단한다. 이때 검증을 통과한 물체들에 대해서는 물체의 움직임 정보와 형태 정보가 얻어지며, 잘못 구해진 변환 파라메터를 갖는 물체에 대해서는 형태 정보와 밝기정보가 얻어져 수신단으로 전송된다. 그림에서 보듯이 제안한 방법은 기존의 방법과 달리 계층적 구조를 갖지 않으며, 한 단계의 처리로 물체들에 대한 움직임 추정 및 보상을 수행하게 된다.

가산투영변환부에서는 입력영상을

$$H(x,y) = \frac{1}{N}\sum_{p=-N/2}^{N/2} f(x+p,y)$$
$$V(x,y) = \frac{1}{N}\sum_{q=-N/2}^{N/2} f(x,y+q) \quad (13)$$

와 같이 각각 x축 방향과 y축 방향으로 N개의 화소씩 가산투영변환하여 수평방향에 대한 가산투영 데이타와 수직방향에 대한 가산투영 데이타를 얻으며, 위 식에서 $H(\cdot,\cdot)$와 $V(\cdot,\cdot)$는 각각 수평가산투영값과 수직가산투영값을 나타내며, $f(\cdot,\cdot)$는 원영상을 나타낸다. 식 (13)에 의해 얻어진 수평가산투영 데이타는 수평방향으로 존재하는 물체의 경계를 보존하면서 수평방향의 가산형 잡음성분을 감소시키는 효과가 있으며 이와 반대로 수직가산투영 데이타는 수직방향의 경계는 잘 보존하면서 수직방향으로 존재하는 가산형 잡음을 제거하는 효과가 있다. 이를 수식으로 전개해 보면 다음과 같다 원영상 신호를 $f(x, y)$라 하고, 평균이 0이고 표준편차가 인 가산형 가우스 잡음신 호를 $n(x, y)$라 하면, 가산형 잡음에 의해 열화된 영상

신호는 식 (6)과 같이 표현되며, 열화된 영상신호의 x축 및 y축 방향에 대한 식 (11)과 (12)의 국부 gradient는

$$G_x(x,y)=\frac{1}{2}\left\{\frac{\partial f_{n,k+1}(x,y)}{\partial x}+\frac{\partial f_{n,k}(x,y)}{\partial x}\right\} \tag{14}$$

$$G_y(x,y)=\frac{1}{2}\left\{\frac{\partial f_{n,k+1}(x,y)}{\partial y}+\frac{\partial f_{n,k}(x,y)}{\partial y}\right\} \tag{15}$$

와 같이 구해진다. 이때 $f_{n,k+1}(\cdot,\cdot)$는 잡음에 의해 열화된 현재 영상신호를 나타내며, $f_{n,k}(\cdot,\cdot)$는 다른 형태의 가산형 잡음에 열화된 이전 영상신호이다. 식 (14)와 (15)에서 $f(x, y)$가 투영방향으로 균일하다고 가정하고, 잡음에 의해 열화된 영상신호 대신 가산투영에 의한 수평 가산투영신호 $H(x, y)$와 수직 가산투영신호 $V(x, y)$를 사용하면,

$$\begin{aligned}\hat{G}_x(x,y)&=\frac{1}{2}\left\{\frac{\partial V_{n,k+1}(x,y)}{\partial x}+\frac{\partial V_{n,k}(x,y)}{\partial x}\right\}\\&=\frac{1}{2N}\left\{\frac{\partial\left[\sum_q f_{k+1}(x,y+q)+\sum_q n_{k+1}(x,y+q)\right]}{\partial x}\right.\\&\left.+\frac{\partial\left[\sum_q f_k(x,y+q)+\sum_q n_k(x,y+q)\right]}{\partial x}\right\}\end{aligned} \tag{16}$$

이며, 같은 과정에 의해

$$\begin{aligned}\hat{G}_y(x,y)&=\frac{1}{2N}\left\{\frac{\partial\left[\sum_p f_{k+1}(x+p,y)+\sum_p n_{k+1}(x+p,y)\right]}{\partial y}\right.\\&\left.+\frac{\partial\left[\sum_p f_k(x+p,y)+\sum_p n_k(x+p,y)\right]}{\partial y}\right\}\end{aligned} \tag{17}$$

이 된다. 식 (16)과 (17)에서 $\Sigma n_k(\cdot,\cdot)$와 $\Sigma n_{k+1}(\cdot,\cdot)$는 투영되는 화소의 개수가 많고, x축 신호와 y축 신호가 분리가능하다고 가정하면 0에 가까운 값을 갖게 되므로

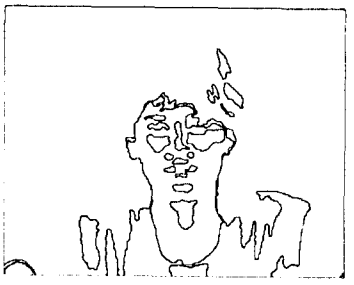
$$\begin{aligned}\hat{G}_x(x,y)&\cong\frac{1}{2N}\left\{\frac{\partial\sum_q f_{k+1}(x,y+q)}{\partial x}+\frac{\partial\sum_q f_k(x,y+q)}{\partial x}\right\},\\\hat{G}_y(x,y)&\cong\frac{1}{2N}\left\{\frac{\partial\sum_p f_{k+1}(x+p,y)}{\partial y}+\frac{\partial\sum_p f_k(x+p,y)}{\partial y}\right\}\end{aligned} \tag{18}$$

와 같은 결과를 얻을 수 있다. 따라서 가산투영 평면상에서 변환 파라메터를 계산하면 잡음이나 조그마한 빛의 변화에 민감한 gradient의 잘못된 계산결과를 초래할 경우를 줄일 수 있다.

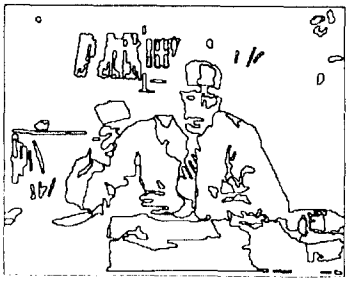
물체 분할부는 차신호 검출부와 영역분할부로 구성되어 있다. 차신호 검출부는 인접한 영상에서 움직임이 있는 부분을 검출하는 역할을 하는데, 이때 사용되는 조건식은

$$Diff(x,y)\begin{cases}\le T_c & \text{움직임이 없는 영역이 화소}\\ > T_c & \text{움직임 영역의 화소}\end{cases} \tag{19}$$

와 같이 인접한 두 영상의 차신호 $Diff(x, y)$가 임의의 임계값 T_c보다 큰 값을 갖는 부분은 움직임 영역으로 판단하고, 임계값보다 작은 화소는 배경부분에 속하는 화소로 판단한다. 이때 임계값 T_c는 영상에 포함된 잡음성분의 크기에 따라 다르지만 보통 3정도의 값을 사용한다.



(a)



(b)

그림 6. 제안한 방법의 영역분할 결과
(a) "Miss America" 영상
(b) "Salesman" 영상

Fig. 6. Segmentation results by the proposed method (a) "Miss America" ,(b) "Salesman"

영역분할부에서는 입력영상을 각각의 강체로 판단할 수 있는 여러 개의 물체로 분할한 후 분할된 영역과 차신호 검출부에서 검출된 움직임이 존재하는 영역을 함께 고려하여 각각의 움직인 물체에 대해 변환 파라메터를 계산하므로 여러가지 움직임이 존재하는 영역을 하나의 물체로 간주하여 발생하는 오차를 제거할 수 있다. 영역분할부는 입력된 현재영상의 비슷한 성질을 갖는 화소들을 한 영역으로 모으는 것으로 영역의 개수가 결정되어야 하고 한 영역을 다른 영역과 구분하는 기준이 있어야 하며, 이를 위해 유사성을 평가할 수 있는 함수가 결정되어야 한다. 제안한 방법의 영역분할시 사용한 방법은 region growing 방법[12]으로, 어느 화소를 기준으로 그 주위화소의 빛의 밝기, 색채, texture, 경사 등의 특징간의 유사성에 의해 주위의 화소들을 연결시키며, 단순한 유사도 측정방법은 알려진 화소들와 주위화소간에 Euclidean 거리를 측정하는 것이다. 유사성 검사시 사용되는 임계값은 영상내의 화소값의 분포나 요구되는 압축율에 따라 결정된다. 만일 큰값을 사용한다면 분할된 영역의 수는 작아지고 압축율은 커지는 효과를 가진다. 그림 6은 제안한 방법에서의 실험영상들에 대한 영역분할된 결과를 나타내며, 사용된 임계값은 30으로 "Miss America" 영상의 경우 분할된 영역의 수는 34개이고, "Salesman" 영상은 96개이다.

Ⅵ. 모의실험 및 결과

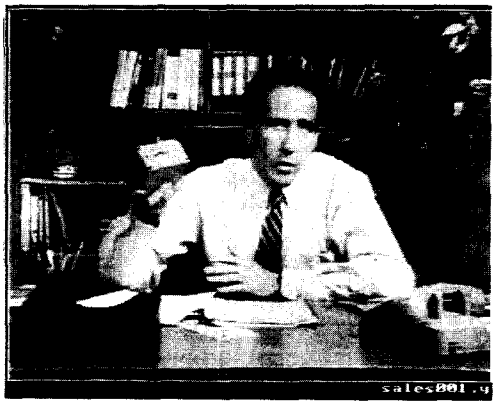
본 절에서는 제안한 움직임 검출방법의 성능분석을 위해 컴퓨터 모의실험결과를 분석하였다. 움직임 검출방법들의 성능분석을 위한 실험에 사용된 영상은 simulation input format (SIF)의 "Miss America" 영상 150장과 "Salesman" 영상 200장을 사용하였으며, 영상의 크기는 352 × 288이다. 사용된 영상은 8bits로 양자화된 영상으로 그림 7에 휘도영상을 보였다. 실험영상에 대한 성능분석 측정치로

$$\mathrm{PSNR} = -\log_{10}\frac{\frac{1}{PQ}\sum\sum\left[f(x,y)-f_r(x,y)\right]^2}{255^2} \qquad (20)$$

로 주어지는 PSNR을 사용하였다. 여기서 $f(\cdot,\cdot)$는 원영상의 화소값을 나타내고 $f_r(\cdot,\cdot)$는 복원된 영상의 화소값이며, P과 Q은 영상의 가로와 세로의 크기이다.

먼저 가산투영변환을 사용한 제안한 방법의 움직임 검출성능을 분석하기 위해 원영상과 잡음영상에 대해

(a)

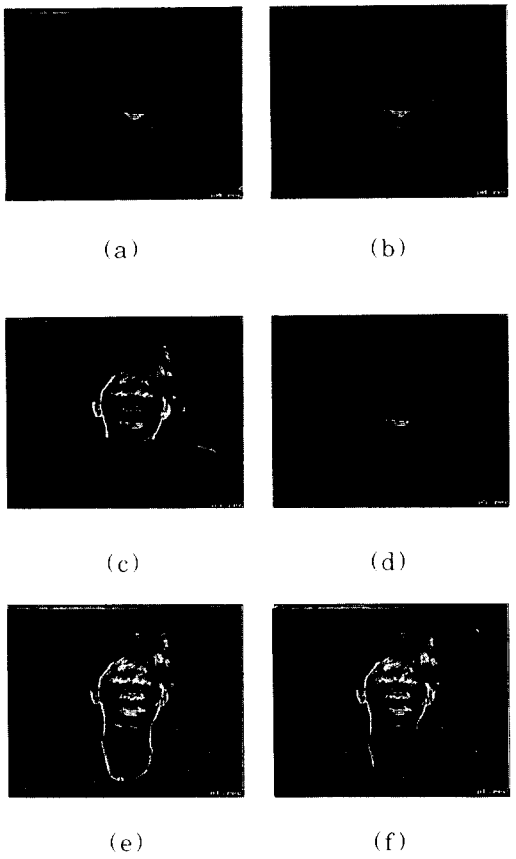


(b)

그림 7. 실험영상
(a) "Miss America" 영상
(b) "Salesman" 영상
Fig. 6. Test sequences.
(a) "Miss America" .
(b) "Salesman"

각각 실험하였다. 이 때 움직임 검출성능만을 분석하기 위해 움직임 보상이 불가능한 물체에 대한 정보는 부호화 오차없이 전송되었다고 가정하였다. 실험에 사용한 잡음영상은 평균이 0이고 표준편차가 각각 5와 10인 가산형 가우스 잡음을 원영상에 섞은 잡음영상이다. 실험에 사용한 기존의 방법[7]은 움직임이 존재하는 영역을 검출한 후 각각의 움직이는 물체에 대해 기존의 변환 파라메터 계산방법에 의해 움직임을 검출하여 보상하는 방법이고, 제안한 방법은 가산투영변환에 의해 수평 및 수직가산투영 영상을 구한 후 투영 데이타와 영역분할된 결과를 사용하여 변환 파

라메터를 계산하는 방법이다. 이때 제안한 방법에서 투영 데이타의 개수는 $N = 10$을 사용하였다. 성능측정을 위한 정량적 평가로는 식 (20)의 PSNR을 사용하였다.



(a) (b) (c) (d) (e) (f)

그림 8. 차영상에 의한 성능비교
(a) 기존의 방법(원영상) (b) 제안한 방법(원영상) (c) 기존의 방법(잡음영상, σ =5) (d) 제안한 방법(잡음영상, σ =5) (e) 기존의 방법(잡음영상, σ =10) (f) 제안한 방법(잡음영상, σ =10)

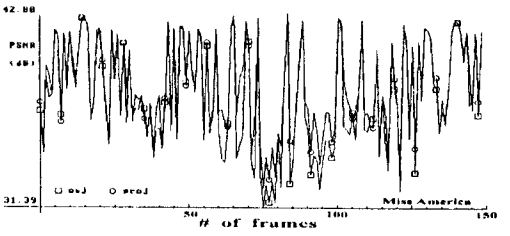
Fig. 8. Performance comparison by difference images. (a) conventional method(original sequence) (b) proposed method (original sequence) (c) con-ventional mthod(moisy sequence, σ =5) (d) Proposed method(noisy sequence, σ =5) (e) conventional mthod(moisy sequence, σ =10) (f) Proposed method (noisy sequence, σ =5).

표 1. "Miss America" 영상에 대한 평균 PSNRI영상분석방법의 성능)

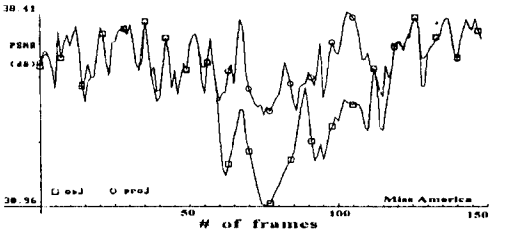
Table 1. Average PSNR for the "Miss America" sequence(performance of image analysis methods).

(단위: dB)

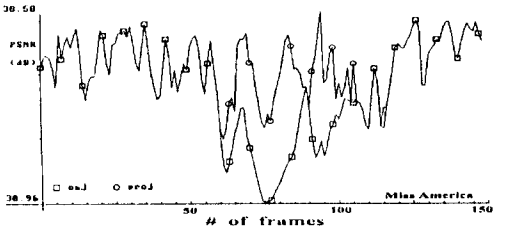
방법	실험 영상		
	원 영상	잡음 영상	
		σ = 5	σ = 10
기존의 방법	37.99	36.57	35.61
제안한 방법	38.52	36.70	36.41



(a)



(b)



(c)

그림 9. PSNR에 의한 성능비교 (a) 원영상 (b) 잡음영상(σ =5) (c) 잡음영상(σ =10)

Fig. 9. Performance comparison by the PSNR measure (a) original sequence, (b) noisy sequence(σ =5) (c) noisy sequence(σ =10).

그림 8은 기존의 방법과 제안한 방법에 의해 복원된 영상과 원영상과의 차이를 8배한 차영상을 나타낸다. 그림 8에서 보듯이 제안한 방법은 기존의 방법이 검출하지 못하는 눈과 입술부분의 움직임까지 잘 검출하여 보상함을 알 수 있다. 표 1과 그림 9에 전체 영상에 대한 실험의 평균 PSNR과 PSNR 그래프를 보였다. 표에서 보듯이 기존의 영상분석방법보다 제안한 방법이 정량적 평가면에서도 우수한 성능을 가짐을 알 수 있다 (원영상에 대한 실험결과 기존의 방법과 제안한 방법의 평균 PSNR의 차는 0.53dB이고, = 5인 잡음영상의 경우 0.13dB, = 10인 경우 0.8dB이다). 그림 9의 그래프에서 'obj'는 기존의 방법이고 'proj'는 제안한 방법이다. 그래프에서 볼 수 있듯이 제안한 방법이 기존의 방법보다 좋은 성능을 나타냄을 알 수 있으며, 움직임이 크거나 복잡한 영상 (영상 sequence중 중간부분)에 대해서는 성능향상이 더 커짐을 볼 수 있으며, 잡음이 강해질수록 제안한 방법의 성능향상이 커짐을 알 수 있다. 위의 실험결과를 종합해 볼 때 제안한 방법은 영상내에 포함되는 잡음의 정도가 커지거나 영상에 포함된 움직임이 복잡해질수록 기존의 방법에 비해 우수한 성능을 갖게됨을 알 수 있다.

두번째 실험은 물체지향 분석 및 합성 부호화에 대한 실험으로 물체의 경계에 대한 부호화[13, 14]와 더불어 움직임 보상이 불가능한 물체에 대한 영역 부호화[15-17]까지 고려한 전체적인 부호화 시스템에 대한 실험결과를 분석하였다. 실험에 사용한 영상은 "Miss America" 영상 150장과 "Salesman"영상 200장이다. 이 실험에서 물체의 경계 부호화에는 difference chain code[13, 14]를 사용하였으며, 영역에 대한 근사화는 다항식 근사법[15-17]을 사용하여 부호화하였다. 성능비교를 위한 기존의 방법[7] ('obj')으로는 IV장에서 설명한 방법을 사용하였으며, 제안한 방법은 좀 더 세밀한 성능분석을 위해 서로 다른 3가지 형태로 시스템을 구성하여 실험하였다. 첫번째 방법은 그림 5의 제안한 시스템에서 움직임 물체 검출시 영역분할 방법을 사용하지 않고 단지 변환 파라메터 계산시 가산투영변환을 사용하여 투영평면에 대한 데이타로부터 변환 파라메터를 구하는 방법 ('pobj1')이며, 두번째 방법은 움직임 물체 검출부에서 영역분할기법을 사용하여 좀 더 세밀한 물체검출과 더불어 기존의 방법에서 사용한 변환 파라메터 계산방법을 사용한 방법 ('pobj2')이다. 마지막으로 세번째 방법은 영역분할기법과 가산투영변환을 함께 결합하여 사용한 방법 ('pobj3')이다.

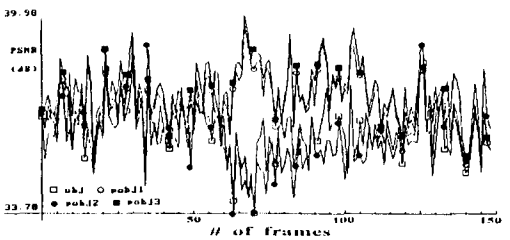
그림 10은 실험영상에 대해 4가지 방법의 PSNR

표 2. 평균 PSNR과 압축비(부호기의 성능)

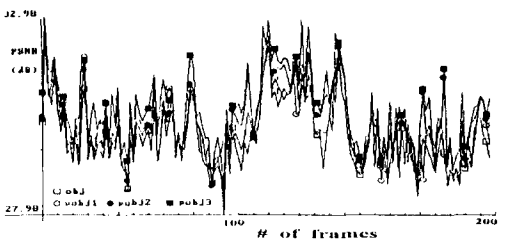
Table 2. Average PSNR and compression ratio(performance of coders).

(단위: dB)

방법	실험 영상			
	"Miss America"		"Salesman"	
	PSNR	압축률	PSNR	압축률
obj	36.29	110.68	30.49	22.18
pobj1	37.21	85.63	30.59	21.58
pobj2	36.50	83.26	30.89	55.03
pobj3	37.50	92.78	31.01	47.74



(a)



(b)

그림 10. 부호기 성능비교를 위한 PSNR 그래프
(a) "Miss America" 영상
(b) "Salesman" 영상

Fig. 10. PSNR graphs for comparison of coder performance (a) "Miss America" (b) "Salesman".

그래프를 나타내고, 표 2는 각 부호기에 대한 평균 PSNR과 압축율을 나타낸다. 표에서 보듯이 "Miss America" 영상에 대해 'obj'와 'pobj1' 사이의 평균 PSNR의 차이는 0.92dB이고, "Salesman" 영상에 대해서는 0.1dB이다. 또한 "Miss America" 영상에 대해 'obj'와 'pobj2' 사이의 차는 0.21dB이고,

"Salesman" 영상에 대해서는 0.4dB이다. 마지막으로 'obj' 와 'pobj3' 사이의 차는 "Miss America" 영상에 대해 1.21dB이고, "Salesman" 영상에 대해 0.52dB이다. 표와 그림에서 보듯이 제안한 방법들은 기존의 방법보다 우수한 성능을 가지며, 'pobj1 '과 ' pobj2 '에서의 결과를 보면 가산투영변환을 사용하거나 영역분할방법을 사용하므로 기존의 방법이 갖는 문제점을 어느 정도 해결하였으며, 가산투영을 사용함으로써 복원영상의 화질을 PSNR면에서 개선할 수 있음을 알 수 있으며, 영역분할방법을 사용하여 압축율면에서 높은 압축을 얻게 된다. 이 두가지 기법이 결합되어 사용됨으로써 더욱 좋은 성능을 가짐을 알 수 있다. 또한 제안한 방법은 "Miss America" 영상과 같이 배경이 단순하고, 커다란 몇개의 영역으로 구성된 영상에 대해 기존의 방법보다 성능향상이 커짐을 알 수 있다. ("Miss America" 영상의 평균 PSNR의 증가: 1.21 dB, "Salesman" 영상의 평균 PSNR의 증가: 0.52 dB)

Ⅶ. 결론

본 논문에서는 전송률이 낮아지면 BMA에서 발생하는 blocking 효과나 mosquito 효과에 의한 시각적 화질저하를 개선해줄 수 있는 물체지향 분석 및 합성 부호화방법에 대해 연구하였다. 물체지향 분석 및 합성 부호화 방법은 영상내에 존재하는 물체단위로 움직임을 검출하여 보상하여줌으로써 BMA보다 자연스러운 움직임을 복원할 수 있으며, 물체단위의 부호화를 적용하므로 전송정보량도 감축할 수 있는 방법이다.

물체지향 분석 및 합성 부호화에서 각각의 물체에 대한 움직임을 검출하기 위해 영상분석과정을 거쳐 변환 파라메터에 의해 움직임을 추출할 때 기존의 방법은 영상내에서 직접 파라메터를 계산하므로 움직임 검출결과가 영상내에 존재하는 잡음이나 조그마한 빛의 변화에 의해 많은 영향을 받는다. 그러나 제안한 방법은 가산투영변환을 사용하여 얻어진 가산투영 데이타를 사용하기 때문에 잡음이나 빛의 변화에 강한 특성을 갖고 비교적 정확한 움직임을 검출한다. 또한 기존의 방법이 갖는 또 다른 문제점으로 여러가지 움직임이 포함된 영역을 하나의 물체로 간주하여 발생하는 움직임 검출시의 오차를 줄이기 위해 영역분할 기법을 도입하여 이러한 문제를 해결하였다. 실험결과적으로 볼 때 제안한 방법은 기존의 방법이 갖는 두가지 문제점을 해결하여 보다 우수한 성능을 나타냄을 알 수 있었다.

물체지향 분석 및 합성 부호화방법은 BMA보다 저전송률의 영상전송시 복원 영상의 시각적 화질이 좋으나 계산량이 방대하여 하드웨어 구현시 상당한 어려움이 있다. 따라서 앞으로는 과다한 계산량을 줄여 실시간처리를 가능케 하는 연구가 계속되어야 할 것이다.

參 考 文 獻

[1] A. K. Jain, "Image data compression: A review," *Proc. IEEE*, vol. 69, pp. 349-389, Mar. 1981.

[2] A. N. Netravali and J. O. Limb, "Picture coding: A review," *Proc. IEEE*, vol. 68, pp. 366-406, Mar. 1980.

[3] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Academic Press, 2nd ed., vol. 1, pp. 353-430, 1982.

[4] J. S. Kim and R. H. Park, "A fast feature based block matching algorithm using integral projections," *IEEE J. Selected Areas Commun.*, vol. SAC-10, pp. 968-971, June 1992.

[5] J.-S. Kim and R.-H. Park, "Feature-based block matching algorithm using integral projections," *IEE Electron. Lett.*, vol. 25, pp. 29-30, Jan. 1989.

[6] J.-S. Kim, R.-H. Park, and B. U. Lee, "Two-stage fast block matching algorithm using integral projections," *Journ. Visual Commun. Image Representation*, to be published, vol. 4, Dec. 1993.

[7] H. G. Musmann, M. H tter, and J. Ostermann, "Object oriented analysis synthesis coding of moving images," *Signal Processing: Image Commun.*, vol. 1, pp. 117-138, Oct. 1989.

[8] M. Hötter and R. Thoma, "Image segmentation based on object oriented mapping parameter estimation," *Signal Processing*, vol. 15, pp. 315-334, Oct. 1988.

[9] J.-Q. Fang and T. S. Huang, "Some experiments on estimating the 3 D motion parameters of a rigid body from

two consecutive image frames," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI 6, pp. 545-554, Sep. 1984.

[10] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of three dimensional motion parameters of rigid objects with curved surface," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 13-27, Jan. 1984.

[11] R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of rigid planar patch," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, pp. 1147-1152, Dec. 1981.

[12] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphics, Image Processing*, vol. 29, pp. 100-132, Jan. 1985.

[13] H. Freeman, "Computer processing of line drawings," *Computer Surveys*, vol. 6, pp. 57-97, 1974.

[14] M. Eden and M. Kocher, "On the performance of a contour coding algorithm in the context of image coding Part I: Contour segment coding," *Signal Processing*, vol. 8, pp. 381-386, July 1985.

[15] M. Gilge, "Coding of arbitrarily shaped image segments using moment theory," *Signal Processing IV: Theories and Applications*, J. L. Lacoume, A. Chehikian, N. Martine, and J. Malbos (Eds.), Elsevier Science Pub. B. V., 1988.

[16] M. Gilge, T. Engelhardt, and R. Mehlan, "Coding of arbitrarily shaped image segments based on a generalized orthogonal transform," *Signal Processing: Image Commun.*, vol. 1, pp. 153-180, Oct. 1989.

[17] M. Gilge, "Region oriented transform coding (ROTC) of images," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing '90*, Albuquerque, New Mexico, pp. 2245-2248, Apr. 1990.

著者紹介

金 俊 植(正會員) 第 29卷 第 4號 參照
현재 호서대학교 전자공학과 전임강사

朴 來 弘(正會員) 第 23卷 第 6號 參照
현재 서강대학교 전자공학과 교수