

# 피라미드 구조 및 국부 오차 보상을 이용한 물체지향 부호화

正會員 趙 大 星\*, 朴 來 弘\*

## Object-Oriented Coder Using Pyramid Structure and Local Residual Compensation

Dae-Sung Cho\*, Rae-Hong Park\* *Regular Members*

※본 연구는 정보통신부 연구비 지원에 의해 수행되었음

### 요 약

본 논문에서는 저전송률 선로에서 피라미드 구조와 오차영상의 보상을 이용한 물체 지향부호화 기법을 제안한다. 움직임 추정단계에서 전역적인 움직임은 피라미드 구조에서의 다해상도 영상을 이용하여 추정한다. 그래디언트(gradient) 값을 이용하여 영상을 두영역으로 나눈다. 움직임 성분이 큰 영역은 잡음성분 및 움직임 constraint식을 만족하기 위하여 저해상도에 측정화소를 결정하고 반면에 눈, 코, 입 등의 국부 움직임 영역은 현재의 해상도에서 측정화소를 결정한다. 영상내의 국부 움직임 및 밝기값 변화 성분은 이전오차 영상을 추가로 이용하여 보정한다. 끝으로 MF (Model Failure) 영역은 이전의 DFD (Displaced Frame Difference) 피라미드의 정합에 의해 보상한다. 컴퓨터 모의실험 결과 PSNR (Peak Signal to Noise Ratio)과 CR(Compression Ratio) 및 계산량의 복잡성 측면에서 제안한 방법이 기존의 방법보다 우수함을 보인다.

### ABSTRACT

In this paper, we propose an object-oriented coding method in low bit-rate channels using pyramid structure and residual image compensation. In the motion estimation step, global motion is estimated using a set of multiresolution images constructed in a pyramid structure. We split an input image into two regions based on the gradient value. Regions with large motions obtain observation points at low resolution level to guarantee robustness to noise and to satisfy a motion constraint equation whereas regions with local motions such as eye, and lips get obseravation points at the original resolution level. Local motion variations and intensity variations of an image

\*西江大學校 電子工學科  
Dept. of Elec. Eng., Sogang Univ.  
論文番號:96064-0215  
接受日字:1996年 2月 15日

reconstructed by the global motion are compensated additionally by using the previous residual image component. Finally, the model failure (MF) region is compensated by the pyramid mapping of the previous displaced frame difference (DFD). Computer simulation results show that the proposed method gives better performance than the conventional one in terms of the peak signal to noise ratio (PSNR), compression ratio (CR), and computational complexity.

## I. 서론

최근 통신 시스템의 발달 및 정보화의 급진전으로 멀티미디어, 화상회의, 화상전화 등 다양한 서비스가 요구되고 있다. 이러한 디지털 영상의 정보량은 많아서 제한된 전송로<sup>(1)</sup>에서 효율적인 데이터 압축이 필요하다. 특히 64kbps 이하에서는 기존의 MPEG-1 (Moving Picture Experts Group-1)<sup>(2)</sup>이나 MPEG-2의 경우 블럭단위로 물체의 움직임을 처리하여 시각적으로 복원화질 저하가 심하고 블럭효과 (blocking effect)가 크기 때문에 저전송을 부호화에 적합하지 않다. 이러한 화질 저하를 개선하기 위해 인간의 시각적 특성을 고려한 소위 제2세대 부호화 기법<sup>(3)</sup>이 제안되었다. 특히 최근에는 초저속 부호화를 위한 MPEG-4의 표준화 후보안의 하나로 물체지향 부호화기법이 제안되었다.<sup>(4-6)</sup> 이 방법은 물체를 부호화의 기본단위로 하여 움직임이 있는 물체부분과 움직임이 없는 배경부분으로 나누어, 배경부분은 이전 프레임의 정보를 그대로 이용하고 물체부분은 기하학적(geometric) 특징을 바탕으로 부호화한다.

Tsai와 Huang<sup>(7)</sup>은 3차원 움직임을 2차원에서 8-파라미터로 표현하였다. 영상의 차영상을 구해 그래디언트 기반 선형 회귀법(linear regression)으로 물체의 움직임을 나타냈다. Musmann 등<sup>(4)</sup>은 이를 얼굴과 어깨(head and shoulder) 영상에 계층적으로 비트율에 따라 적응부호화하였다. 그러나 급격한 움직임이나 잡음이 존재하면 그래디언트 방법에 의한 움직임은 정확하지 않고 부호화시 누적오차가 크다. 이를 보완하기 위해 투영(projection)의 개념으로 잡음에 강한 파라미터를 얻는 방법이 제안되었다.<sup>(8-9)</sup> 또한 블럭 기반 알고리즘<sup>(10)</sup>을 이용한 움직임 벡터를 움직임 파라미터 모델로 정합하는 방법도 제안되었다.<sup>(11-13)</sup>

본 논문에서는 그래디언트 기반 방법을 보완하고 또한 계산량 및 누적 오차를 줄이기 위해 피라미드

구조 및 오차영상의 보정을 이용하였다. Moon과 Kim<sup>(14)</sup>은 파라미터 수에 따른 복원 영상의 오차를 실험하였다. 여기서 8-파라미터는 다른 모델에 비해 가산형 가우시안 잡음(additive Gaussian noise)에 민감하고 6-파라미터는 8-파라미터에 근접한 성능을 보여 본 논문에서는 이 모델을 사용하였다. 제안한 방법은 움직임 파라미터 추정시 피라미드 구조를 적용하였고 전역 움직임 보정후 생긴 국부 움직임 오차는 이전오차를 이용하여 보정한다. MF 영역은 이전의 DFD 피라미드로 보정하였다.

본 논문에서는 II장에서 기존의 물체지향 부호화기를, III장에서는 피라미드 구조와 이전오차 정합을 이용한 제안한 물체지향 부호화기를 설명하였다. IV장에서는 제안한 방법과 기존의 방법을 컴퓨터 모의 실험으로 성능을 비교하였다. 끝으로 V장에서 결론을 맺었다.

## II. 기존의 물체지향 부호화 방법

3차원 공간상의 물체의 움직임을 2차원 평면상에 투영하여 나타내면 물체의 직선운동 및 회전성분 등을 8개의 파라미터  $a_i (1 \leq i \leq 8)$ 를 사용하여  $(X, Y)$ 에서  $(X', Y')$ 로의 움직임을

$$\begin{aligned} X' &= \frac{(a_1 + 1)X + a_2Y + a_3}{a_7X + a_8Y + 1} \\ Y' &= \frac{a_4X + (a_5 + 1)Y + a_6}{a_7X + a_8Y + 1} \end{aligned} \quad (1)$$

로 쓸 수 있다.<sup>(4, 6, 7)</sup> 물체가 카메라에서 멀고 초점거리가 크면 이 움직임은 6개의 파라미터로

$$\begin{aligned} X' &= (a_1 + 1)X + a_2Y + a_3 \\ Y' &= a_4X + 1(a_5 + 1)Y + a_6 \end{aligned} \quad (2)$$

과 같이 표현할 수 있다.<sup>(14)</sup>

그림 1은 물체지향 부호화기의 기본 구조이다.<sup>(4)</sup> 영상 분석(image analysis)은 입력 영상을 물체와 배경으로 나누어 움직임 파라미터를 추정하며 MF 영역을 검출하고 보상한다. 파라미터 부호화(parameter coding)는 전단계에서 추정된 형태(shape), 움직임(motion), 색(color) 성분을 부호화하고 파라미터 복호화(parameter decoding)는 다음 프레임의 움직임 보상을 위한 영상 합성(image synthesis)을 수행한다. 파라미터 부호화는 파라미터 메모리(memory for parameters)에 저장된 이전의 파라미터들을 사용한다.

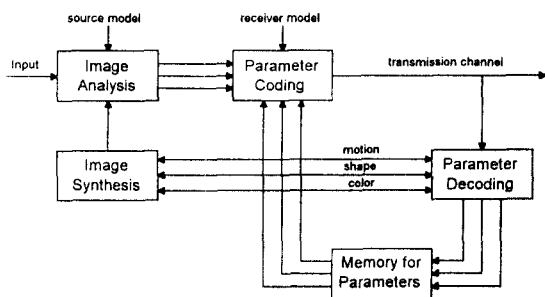


그림 1. 기존의 물체지향 부호화기의 블럭선도  
Fig 1. Block diagram of the conventional object-oriented analysis-synthesis coder.

### 1. 움직임 파라미터 추정

Hötter와 Thoma<sup>(6)</sup>는 8-파라미터를 사용하여 움직임을 나타냈다. 연속된 두영상의  $k+1$ 번째 프레임의  $(X, Y)$ 에서의 FD(Frame Difference)  $FD_{k+1}(X, Y)$ 는 Taylor 전개를 이용하여

$$FD_{k+1}(X, Y) = H^T \Delta A \quad (3)$$

와 같이 나타낼 수 있다. 여기서

$$\begin{aligned} \Delta A &= (\Delta a_1, \Delta a_2, \Delta a_3, \Delta a_4, \Delta a_5, \Delta a_6, \Delta a_7, \Delta a_8)^T = A - \Phi = A, \\ A &= (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8)^T, \\ H &= (G_X X, G_X Y, G_X, G_Y X, G_Y Y, G_Y - X(G_X X + G_Y Y), \\ &\quad -Y(G_X X + G_Y Y))^T \end{aligned}$$

로  $\Delta A$ 는 움직임 파라미터의 변화 성분이고 움직임이

작은 경우에 이 움직임 벡터는 영벡터  $\Phi$ 이므로 8-파라미터 벡터  $A$ 와 같다.  $H$ 는 영상의 그래디언트와 위치정보로 이루어졌다.  $G_X$ 와  $G_Y$ 는  $X, Y$  방향의 그래디언트 성분으로서  $k$ 번째와  $k+1$ 번째의 입력 영상으로부터

$$\begin{aligned} G_{k+1, X} &= \frac{1}{2} \left\{ \frac{\partial f_{k+1}(X, Y)}{\partial X} + \frac{\partial f_k(X, Y)}{\partial X} \right\} \\ G_{k+1, Y} &= \frac{1}{2} \left\{ \frac{\partial f_{k+1}(X, Y)}{\partial Y} + \frac{\partial f_k(X, Y)}{\partial Y} \right\} \end{aligned} \quad (4)$$

과 같이 된다. 식 (3)으로부터 실제 FD와 추정된 FD 사이의 오차  $E[(FD_{k+1}(X, Y) - \Delta H^T A)^2]$ 를 최소화하여 움직임 파라미터 벡터  $A$ 를

$$A = (H^T \cdot H)^{-1} \cdot H^T \cdot FD \quad (5)$$

와 같이 구한다. 여기서  $p$ 개의 관측 데이터로부터  $H$ 와  $FD$ 는

$$\underline{H} = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_p \end{bmatrix}, \quad \underline{FD} = \begin{bmatrix} FD(X_1, Y_1) \\ FD(X_2, Y_2) \\ \vdots \\ FD(X_p, Y_p) \end{bmatrix}$$

로 표현된다. 6-파라미터의 경우  $A$ 와  $H$  벡터의 앞의 6항만을 고려하여 구한다.

### 2. 예측 오차 부호화

움직임 파라미터에 의한 물체의 움직임 보상후 모델에 맞지 않는 MF 영역 부호화를 위하여 Schiller와 Hötter<sup>(5)</sup>는 MF 영역을  $8 \times 8$  크기의 블럭으로 나누고 블럭내에 움직임이 있는 화소의 비율이 크면 DCT (Discrete Cosine Transform)를, 그렇지 않으면 DPCM (Differential Pulse Code Modulation)을 사용하여 부호화하는 적응 DCT/DPCM 방법을 사용하였다. 움직이는 화소의 비율이 작은 블럭에 대해 DCT를 행할 경우 실제 움직임이 없는 화소의 밝기값 정보도 부호화하므로 비트량이 많다. 따라서 이 경우는 DPCM의 방법이 공간 중복성을 줄이는데 효과적이다.

### III. 제안한 물체지향 부호화 방법

얼굴 및 어깨 영상은 눈, 코, 입 등의 국부 움직임과

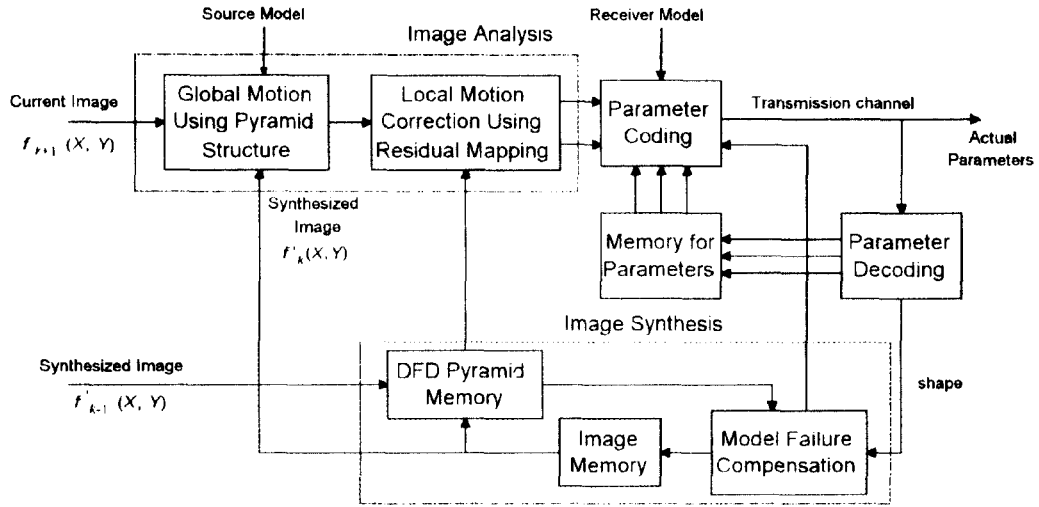


그림 2. 피라미드 구조 및 국부오차 영상 보정을 이용한 제안 방법의 전체 블럭선도.

Fig 2. Overall block diagram of the proposed method using pyramid structure and local residual image compensation.

그 이외의 전역 움직임이 동시에 존재하므로 각 영역의 특성에 맞는 움직임 추정이 필요하다. 움직임은 전역 및 국부 움직임을 나누어 추정하며 전역 움직임은 측정 화소값을 피라미드로부터 얻는다. 이 경우 국부 특성도 어느 정도 반영되며 보다 정확한 국부 움직임 추정을 위해 이전 오차 영상을 이용한다. MF 영역은 DFD 피라미드를 이용한다. 본 논문에서 제안한 방법의 전체 블럭선도를 그림 2에 나타내었다.

1. 영상내 물체의 움직임 특성

일반적으로 영상내의 물체의 움직임은 영상의 밝기값의 시간에 대한 변화 성분을 이용하여 나타낼 수 있다. 영상  $f(X, Y)$ 의 시간에 대한 미분은 chain rule에 의하여

$$\frac{df(X, Y)}{dt} = \frac{\partial f(X, Y)}{\partial X} u(X, Y) + \frac{\partial f(X, Y)}{\partial Y} v(X, Y) + \frac{\partial f(X, Y)}{\partial t} \tag{6}$$

과 같이 나타낼 수 있다. 여기서  $u(X, Y)$ 와  $v(X, Y)$ 은  $(X, Y)$ 에서의 수평, 수직 방향의 움직임 성분이다. 식

(6)에서 시간에 따라 밝기값 변화가 없고 오직 움직임에 의해서 변한다고 가정하여

$$\frac{df(X, Y)}{dt} = 0 \tag{7}$$

같이 움직임 constraint식을 설정하여 움직임을 구한다. 식 (7)에서 움직임  $u, v$ 의 평행이동 움직임을 고려하여 위치에 관계없이 일정하고,  $\hat{v}$ 가 0이면 추정될 움직임 성분  $\hat{u}$ 는

$$\hat{u} = - \frac{\sum_{i=1}^N \frac{\partial f(X_i, Y_i)}{\partial t} \frac{\partial f(X_i, Y_i)}{\partial X_i}}{\sum_{i=1}^N \left\{ \frac{\partial f(X_i, Y_i)}{\partial X_i} \right\}^2} \tag{8}$$

과 같다.<sup>(15)</sup> 여기서  $N$ 은 움직임 영역내에 포함된 화소의 수이다. 식 (8)에서 영상의  $X$ 방향의 그래디언트 값이 클 경우 영상내 움직임이 작으며 국부 움직임 특성을 보인다. 반대의 경우는 전역 움직임을 나타낸다.

식 (7)은 영상 밝기값의 변화가 없다고 가정한 경우이다. 그러나 실제의 동영상은 카메라 취득 환경이 일정하지 않아 전역적인 밝기값 변화가 존재하므로 식

(7)대신 움직임 수식을

$$\frac{df(X, Y)}{dt} = -C \quad (9)$$

과 같이 기존의 방법은 밝기값이 시간에 따라 일정하게 변한다고 하여  $C$ 를 하나의 파라미터로 하여 구하기도 하였다.<sup>(16)</sup>

식 (9)의 경우는 영상내 전역적인 특성만을 고려한 경우이다. 그러나 영상내의 국부 움직임 특성에 의해서 밝기값 변화가 상수 성분만으로 표현되지 않으며, 특히 6-파라미터에 적용할 경우 선형 변환 과정의 오차까지 고려해야 한다. 따라서 식 (9)대신 본 논문에서는

$$\frac{df(X, Y)}{dt} = -R(X, Y) \quad (10)$$

를 사용하며 물체의 위치에 따라 밝기값 변화를 달리 하여 국부 움직임을 나타낼 수 있도록 한다. 여기서 식 (10)은 영상이 시간에 따라 움직임 성분이외에 밝기값성분의 변화가 존재하고, 그 존재도 국부 움직임이 존재할 경우  $R(X, Y)$ 와 같이 그 변화가 화소의 위치에 따라 달라진다는 것을 의미한다.

### 2. 피라미드 구조에서 계층적 움직임 영역 검출

움직임 영역은 초기에 임계치를 3으로 설정하여 차영상과 비교하여 검출하고 이후 프레임은 이전 프레임에서 구한 배경영역에서 현재 프레임의 FD의 표준편차  $\sigma_{k+1}$ 를 구하여

$$T_{k+1}^{(1)} = c_0 \sigma_{k+1} \quad (11)$$

과 같은 임계치를 구한다. 여기서 첨자 (1)은 1계층을 나타내고 상수항  $c_0$ 는 보통 1에서 2사이의 값을 갖는다. 계층의 의미는 움직임 보상후의 영상을 몇번까지 다시 움직임 영역을 검출하고 움직임을 보상하는가를 나타낸다.  $3 \times 3$  창 of 평균 FD값을 임계치  $T_{k+1}^{(1)}$ 과 비교하여 움직임 영역을 구한다.

식 (11)에 의한 영역에서 전역 움직임을 추정하기 위해 피라미드를 이용한다. 피라미드는 2계층으로 되어 있고 저해상도의 영상은 4:1로 평균 샘플링된 것이다. 국부 및 전역 움직임 구분을 위한 임계치는 그

래디언트의 오차 특성으로 구한다. 영상내  $X$ 방향의 그라디언트 성분의 오차 성분은

$$\begin{aligned} GN_{k+1}(X+\Delta X, Y) &= (\tilde{f}_{k+1}(X+\Delta X, Y) - f_{k+1}(X+\Delta X, Y)) \\ &\quad - (\tilde{f}_{k+1}(X, Y) - f_{k+1}(X, Y)) \quad (12) \\ &= N_{k+1}(X, \Delta X, Y) - N_{k+1}(X, Y) \end{aligned}$$

과 같이 표현된다. 여기서  $f_{k+1}(X, Y)$ 는 입력 영상이고  $\tilde{f}_{k+1}(X, Y)$ 는 가산형 잡음  $N_{k+1}(X, Y)$ 가 포함된 영상이다. 식 (12)의 표준 편차를 기준으로 하여 임계치의 제곱을

$$T_{G, k+1}^2 = 2 \cdot \sigma_{N, k+1}^2 = \sigma_{G, k+1}^2 \quad (13)$$

과 같이 식 (5)에서의  $H$ 와  $FD$ 의 측정화소를 얻기 위한 해상도를 결정하도록 설정한다.<sup>(6)</sup> 여기서  $\sigma_{N, k+1}$ 는 잡음 성분의 표준 편차이고  $\sigma_{G, k+1}$ 는 FD의 표준 편차를 나타낸다. 눈, 코, 입 등 국부 움직임 특성을 보이는 곳은 그라디언트가 크며 식 (8)에서 작은 평행 이동 움직임을 보인다. 이 경우 측정화소를 원해상도에서 얻는다. 그밖의 영역은 전역 움직임을 나타낸다.<sup>(15)</sup> 이 경우 잡음에 민감하므로 이 영향이 작도록 피라미드의 저해상도 영상을 이용하고, 특히 식 (10)은 움직임이 작은 경우 만족하므로 움직임이 작은 저해상도 영상으로부터 측정화소를 얻는다.

움직임 파라미터에 의해 전역 보상되지 않은 영역 검출을 위한 임계치는

$$T_{k+1}^{(2)} = c_1 \frac{1}{N} \sum_{i=1}^N |f_{k+1}(X_i, Y_i) - \hat{f}_{k+1}(X_i, Y_i)| \quad (14)$$

로 설정한다. 여기서 첨자 (2)는 2계층을 나타내고  $\hat{f}_{k+1}(X_i, Y_i)$ 는 움직임 보상후의 영상이다. 상수항  $c_1$ 은 검출되는 오차 영역의 수를 조절하여 비트량을 조절하기 위한 것으로 실험상 0.5에서 1의 값으로 정하였다.

### 3. 피라미드 구조 및 국부오차 영상 보정을 이용한 움직임 파라미터 추정

#### 3.1 피라미드 구조를 이용한 전역적 움직임 추정

피라미드 구조를 이용하여 식 (5)의  $H, FD$  벡터를

구하기 위해 측정 화소값을 영상내 움직임 특성에 따라 원해상도 및 저해상도의 경우로 나눈다. 영상의 2×2 블럭 내의 그라디언트의 절대값의 평균값이 식 (13) 보다 작으면 원해상도의 각 화소를 측정화소로  $H_H$ 를

$$H_H = \begin{bmatrix} G_{X_1}X_1 & G_{X_1}Y_1 & G_{X_1} & G_{Y_1}X_1 & G_{Y_1}Y_1 & G_{Y_1} \\ G_{X_2}X_2 & G_{X_2}Y_2 & G_{X_2} & G_{Y_2}X_2 & G_{Y_2}Y_2 & G_{Y_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ G_{X_M}X_M & G_{X_M}Y_M & G_{X_M} & G_{Y_M}X_M & G_{Y_M}Y_M & G_{Y_M} \end{bmatrix} \quad (15)$$

과 같이 구한다. 여기서  $M$ 은 영상내의 원해상도로 측정되는 화소수를 나타내고 인덱스  $H$ 는 원해상도에서의 값을 의미한다.  $FD_H$  벡터는

$$FD_H = \begin{bmatrix} f_{k+1}(X_1, Y_1) - f_k(X_1, Y_1) \\ f_{k+1}(X_2, Y_2) - f_k(X_2, Y_2) \\ \vdots \\ f_{k+1}(X_M, Y_M) - f_k(X_M, Y_M) \end{bmatrix} \quad (16)$$

과 같다. 만약 2×2 블럭내의 그라디언트 값의 평균이 식 (13)에 비해 작으면 그 영역은 전역특성을 나타낸다. 이 경우 움직임이 크므로 움직임 constraint를 만족하기 위해 피라미드의 저해상도에서 측정화소를 구한다. 2×2 블럭에 해당하는 저해상도의 한 화소값을 측정화소로 하여  $H_L$ 과  $FD_L$ 는

$$H_L = \begin{bmatrix} \frac{\bar{G}_{X_1}}{2} & \frac{\bar{G}_{Y_1}}{2} & \frac{1}{2} & \frac{\bar{X}_1}{2} & \frac{\bar{Y}_1}{2} & \frac{1}{2} \\ \frac{\bar{G}_{X_2}}{2} & \frac{\bar{G}_{Y_2}}{2} & \frac{1}{2} & \frac{\bar{X}_2}{2} & \frac{\bar{Y}_2}{2} & \frac{1}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\bar{G}_{X_N}}{2} & \frac{\bar{G}_{Y_N}}{2} & \frac{1}{2} & \frac{\bar{X}_N}{2} & \frac{\bar{Y}_N}{2} & \frac{1}{2} \end{bmatrix} \quad (17)$$

$$FD_L = \begin{bmatrix} \bar{f}_{k+1}(X_1, Y_1) - \bar{f}_k(X_1, Y_1) \\ \bar{f}_{k+1}(X_2, Y_2) - \bar{f}_k(X_2, Y_2) \\ \vdots \\ \bar{f}_{k+1}(X_N, Y_N) - \bar{f}_k(X_N, Y_N) \end{bmatrix} \quad (18)$$

과 같이 구한다. 여기서  $N$ 은 국부 영역의 측정화소 수이고 인덱스  $L$ 은 저해상도를 의미한다. 그리고 “-”는 평균값을 의미한다. 특히 이 영역은 잡음에 의

한 영향이 크므로 평균 필터링된 저해상도 영상을 이용하여 이를 줄인다. 식 (15)-(18)을 함께 고려하면 식 (5)의  $H, FD$  벡터는

$$H = \begin{bmatrix} \alpha H_H \\ \dots \\ \beta FD_L \end{bmatrix}, \quad FD = \begin{bmatrix} \alpha FD_H \\ \dots \\ \beta FD_L \end{bmatrix}, \quad (\alpha \leq \beta) \quad (19)$$

로 나타낼 수 있다. 여기서  $\alpha$ 와  $\beta$ 는 영역에 따른 가중치로서 LS(Least Squares) 기법에서 오차가 큰 영역에 작은 가중치를 할당하는 것과 비슷하다. 식 (19)로부터 국부 영역의 경우 가중치를 작게 한다. 이것은 outlier에 해당하는 관측화소가 주로 국부 움직임 영역에 존재하여 이를 제거하기 위한 것이다. 특히 국부 움직임은 경계 부분에서도 나타나므로 작은 가중치로 에지의 영향을 줄인다. 식 (19)를 식 (5)에 대입하여 피라미드에 의해 형태 변화 성분이 없는 전역 움직임 성분을 구할 수 있다. 저해상도의 경우 4분의 1로 샘플링된 영상으로부터 화소값을 이용하므로 원해상도보다 4분의 1로 계산량이 적은 장점도 있다.

### 3.2 국부오차 영상 보정을 이용한 국부 움직임 보정 파라미터 추정

식 (10)으로부터 전역 움직임 보상후의  $(X, Y)$ 에서 의  $k+1$ 번째의 DFD 성분은

$$DFD_{k+1}(X, Y) = (f_k(X+u(X, Y), Y+v(X, Y)) - f_{k+1}(X, Y)) + R_{k+1}(X, Y) \quad (20)$$

과 같다. 여기서  $u(X, Y), v(X, Y)$ 는 6-파라미터에 의한 전역 움직임 벡터이고  $R_{k+1}(X, Y)$ 는 전역 움직임으로 보상되지 않은 국부 움직임이나 움직임 constraint 방정식에서의 국부적인 밝기값 변화 성분, 그리고 6-파라미터 모델링 과정에서 생긴 오차를 모두 포함한다. 이 성분을 줄이기 위해 이전 오차를 이용한다. 오차 영상  $R_{O, k}(X, Y)$ 를 추정한 전역적 움직임 파라미터에 의하여

$$\hat{R}_{O, k}(X, Y) = R_{O, k}(X', Y') \quad (21)$$

과 같이 보상한다. 여기서 인덱스  $O$ 는 움직임 영역을 의미한다. 이를 국부 보상 파라미터인  $s_{O, k}, m_{O, k}$ 를 사

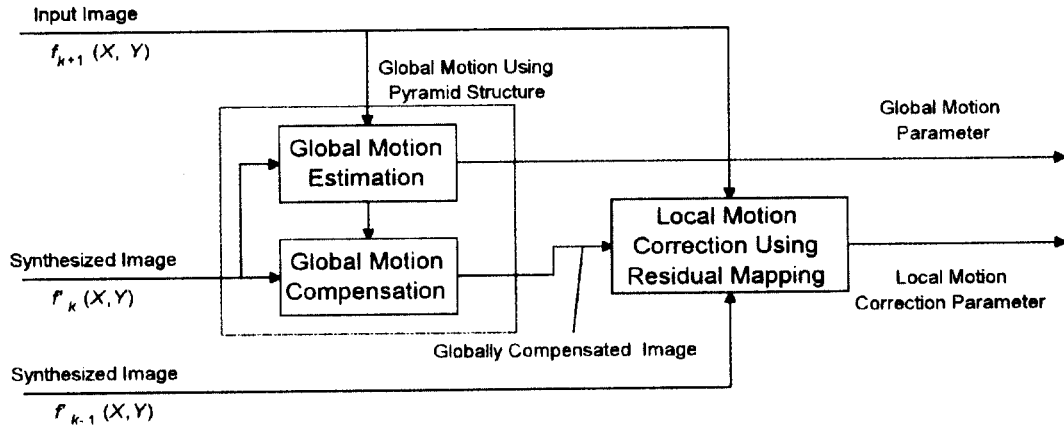


그림 3. 제안한 방법의 전역 및 국부 움직임 추정  
 Fig 3. Global and local motion estimation of the proposed method.

용하여

$$\hat{R}_{O, k+1}(X, Y) = s_{OA} \hat{R}_{O, k}(X, Y) + m_{OA} \quad (22)$$

과 같이 현재 보상하려는 오차 성분으로 근사화한다. 식 (22)의 움직임 보상된 오차영상과 현재의 국부 오차 성분  $R_{O, k+1}(X, Y)$ 과의 차가 최소가 되도록 하기 위하여 목적함수

$$MSE(s_{OA}, m_{OA}) = \sum_{(X, Y) \in O_{MF}} [R_{O, k+1}(X, Y) - \hat{R}_{O, k+1}(X, Y)]^2 \quad (23)$$

를 정의한다. 여기서  $O_{MF}$ 은 움직임 영역안의 화소의 집합이다. 식 (23)을 통해 국부 움직임을 보상할 수 있다. 그림 3은 제안한 움직임 추정 기법의 블록선도이다.

#### 4. DFD 피라미드의 프랙탈 매핑을 이용한 MF 영역 보상

MF 영역 보정을 위해 피라미드의 각 계층을 움직임 보상후 DFD 피라미드를 저장해 놓고 이 값을 이용한다. 그림 4는 제안한 방법의 전체 블록도이다. 먼저 물체 단위의 MF 영역을 얻고 각 오차 성분이 이전 오차 성분과 오차가 적도록 affine 파라미터를 구한다. 두 오차영상의 밝기값 정보의 크기가 다르므로 영역내 각 화소에 대하여 일정크기의 창을 두어 탐색 범위내에서 MSE가 최소인 움직임 벡터를 구하고

$$MSE(B_{MF}) = \sum_{(X, Y) \in O_{MF}} \{ (d_{MF, X} - (b_1 X + b_2 Y + b_3))^2 + (d_{MF, Y} - (b_4 X + b_5 Y + b_6))^2 \} \quad (24)$$

를 최소화하여 affine 파라미터  $b_i (1 \leq i \leq 6)$ 를 구한다. 여기서  $B_{MF}$ 는  $b_i$ 를 원소로 하는 파라미터 벡터이고  $O_{MF}$ 는 MF 영역에 속하는 화소 집합을,  $d_{MF, X}$ 와  $d_{MF, Y}$ 는 이전 오차에서 현재의 MF로의 화소단위의 수평, 수직 위치 변화를 나타낸다.

추정된  $B_{MF}$  파라미터인  $b_i (1 \leq i \leq 6)$ 로 물체 단위로 DFD 피라미드를 보정한다. 보정후의 오차 성분은 수치변환 파라미터를 이용해 보정한다. 보상된 DFD 피라미드를 이용하여 이전 오차성분  $\hat{R}_{DF, k}(X, Y)$ 에 대해

$$MSE(s_{MF}, m_{MF}) = \sum_{(X, Y) \in O_{MF}} [R_{MF, k+1}(X, Y) - (s_{MF} \hat{R}_{MF, k}(X, Y) + m_{MF})]^2 \quad (25)$$

를 최소화하는 수치변환 파라미터인, MF 영역의 크기 및 밝기값 변화 성분  $s_{MF}$ 와  $m_{MF}$ 를 블록 단위로 구한다. 식 (25)의 파라미터를  $8 \times 8$  블록 각각에 대해 보정된 DFD의 원해상도 영상을 이용하여 구하고 또  $8 \times 8$  블록을  $4 \times 4$  블록 4개로 나누어 각각 같은 파라미터를 저해상도의 DFD 영상을 이용하여 구한다. 두 경우에 대해서 MSE가 작은 파라미터로 MF 영역

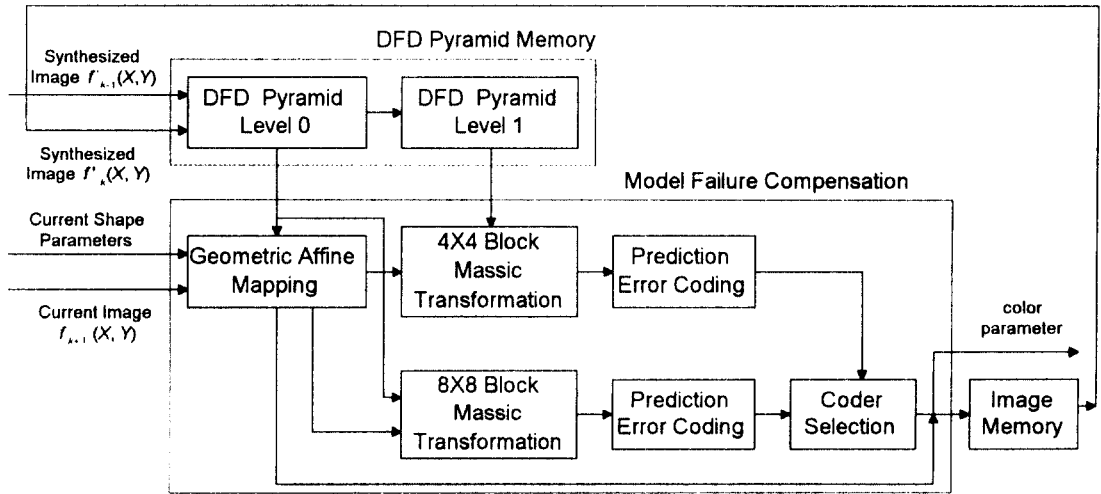


그림 4. DFD 피라미드를 이용한 제안한 MF 영역 보상  
 Fig 4. Proposed MF region compensation using the DFD pyramid.

의 DFD를 근사화한다. 근사화후 적응 DCT/DPCM에 의해 MF 영역을 보상한다. 기존의 프랙탈 변환인 self-VQ (self-vector quantization)<sup>(17)</sup>와 달리 피라미드의 정보량을 추가로 이용한다.

5. 파라미터 부호화 및 움직임 벡터 보간

부호화에서 전송할 파라미터로는 형태, 움직임, MF 보상 성분이다. MF 영역의 affine 파라미터는 움직임 파라미터와 같은 VQ<sup>(18)</sup>를 사용한다. 6-파라미터에서  $(a_1, a_2)$ 와  $(a_4, a_5)$ 는 2차원 크기의 벡터로 구성하여 부호화하고  $a_3$ 와  $a_6$ 는 스칼라 양자기를 사용한다.

움직임 파라미터에 의한 움직임 벡터는 정수 단위로 나타나지 않는다. 또한 파라미터 양자화에 의해 복원영상의 화질이 나빠진다. 이를 보정하기 위해 밝기값 성분의 보간이 필요하다. 이를 위해 이웃한 4개의 화소값으로부터 복원 영상은

$$\hat{f}_{k+1}(X, Y) = f_k(X, Y)(1-h)(1-v) + f_k(X+1, Y)h(1-v) + f_k(X, Y+1)(1-h)v + f_k(X+1, Y+1)hv \quad (26)$$

과 같이 구한다. 여기서  $f_{k+1}(X, Y)$ 는  $k+1$ 번째의 움직임 파라미터에 의하여 복원된 영상이고  $h, v$ 는 이웃한 화소의 수평, 수직 위치에서의 비율이다.

형태 성분의 부호화를 위하여  $2 \times 2$  블록을 기본 단위로 한 패키지형 경계 부호화를 사용했다.<sup>(19)</sup> 패키지형 경계 부호화 방법은 23개의 부호화 구조를 정의하였고 화소당 0.8비트가 사용되었다.

IV. 실험 결과

모의실험을 위해 사용한 영상은  $352 \times 288$  크기의 Clair 영상 30 프레임과 Miss America, Salesman 각 150 프레임으로 8비트로 양자화되었다. 그림 5는 실험에 사용한 영상을 나타낸다.



(a)



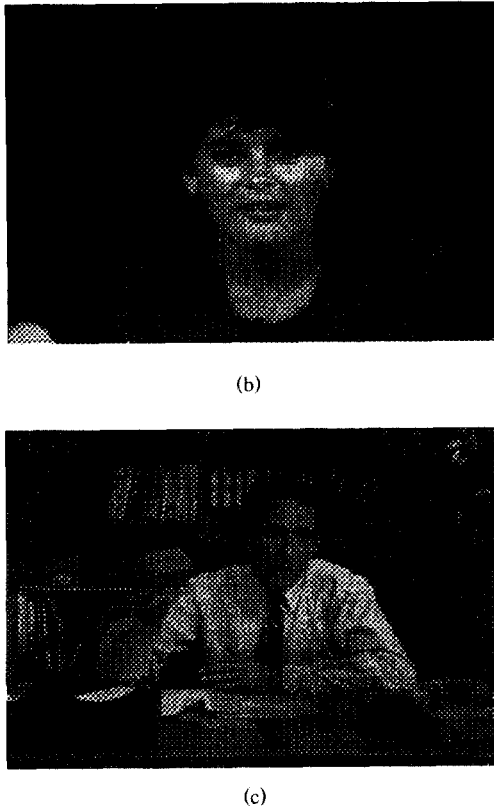


그림 5. 실험 영상

(a) Clair (b) Miss America (c) Salesman

Fig 5. Test sequences.

(a) Clair. (b) Miss America. (c) Salesman.

그림 6의 왼쪽은 피라미드를 사용하기 위해 식 (11)의 임계치를 사용한 후 식 (13)에 따라 그래디언트로 국부 및 전역 움직임 영역으로 나눈 결과이다. 눈, 코, 입 등 국부 움직임을 갖는 흰 영역은 그래디언트 값이 크게 나왔고 그외의 흰색 부분은 전역 움직임을 나타낸다. 그림 6의 오른쪽 부분은 인접한 두 프레임간에 계층적으로 검출된 물체 영역 및 MF 영역을 나타낸다. 영상에서 밝기값이 가장 흐린 부분이 식 (11)을 사용한 영역이고 회색인 부분이 식 (14)를 사용하여 검출한 영역이다. 하얗게 표시된 부분은 MF 영역이다.

표 1은 6-및 8-파라미터를 이용한 기존의 움직임 추정 방법<sup>4)</sup>과 피라미드 및 국부오차 보정을 이용한 제안한 방법을 비교하였다. 성능 비교를 위한 PSNR은

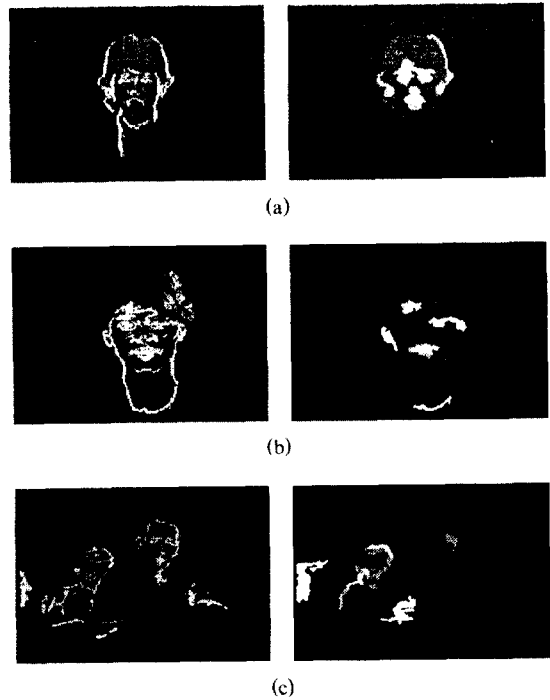


그림 6. 초기 검출된 움직임 영역 및 계층적 구조를 이용하여 검출된 움직임 영역

(a) Clair (b) Miss America (c) Salesman

Fig 6. Initial moving regions and moving regions detected using hierarchical structure.

(a) Clair. (b) Miss America. (c) Salesman.

$$PSNR = 10 \log_{10} \frac{255^2}{(1/(M \times N)) \sum \sum [f(X, Y) - f'(X, Y)]^2} \quad (27)$$

과 같이 정의된다. 여기서  $M \times N$ 는 영상의 크기이며  $f(X, Y)$ 는 원영상을 나타내고  $f'(X, Y)$ 는 부호화후의 복원영상을 나타낸다. 기존의 움직임 파라미터 추정 방법은 Musmann 등의 방법<sup>4)</sup>을 사용하였고 제한한 방법은 이를 피라미드로 확장하였다. 명확한 성능 비교를 위해 움직임 영역 내의 PSNR을 나타냈고 전체 영상의 PSNR 결과도 나타내었다. 움직임 영역 검출을 위한 임계치는 3으로 하여 프레임간의 움직임을 구하였다. 기존의 방법의 경우 8-파라미터가 6-파라미터보다 성능이 우수하였고 저해상도 영상의 경우는 국부 특성을 제대로 반영하지 못하여 원해상도보다 화질이 떨어졌다. 특히 저해상도의 8-파라미터는

표 1. 움직임 파라미터 추정 기법의 성능 비교(PSNR : dB)

Table 1. Performance comparison of motion parameter estimation methods (PSNR : dB).

실험 영상 방법	Clair		Miss America		Salesman	
	움직임 영역	전체영상	움직임 영역	전체영상	움직임 영역	전체영상
8 파라미터(원해상도)	33.50	40.82	35.58	38.29	30.72	34.13
8 파라미터(저해상도)	31.82	40.27	35.43	38.21	30.59	34.03
6 파라미터(원해상도)	33.13	40.72	35.52	38.26	30.60	34.05
6 파라미터(저해상도)	32.50	40.57	35.37	38.17	30.47	33.95
제안한 방법 1	33.03	40.66	35.50	38.24	30.57	34.03
제안한 방법 2	34.07	40.99	36.11	38.47	30.91	34.18

원해상도에서의 6-파라미터보다 성능이 저하되었다. 제안한 방법 1은 피라미드를 이용한 결과로서 기존의 6-파라미터 방법에 비해 주관적 화질 차이는 거의 없으며 PSNR은 거의 같다. 특히 기존의 방법에 비해 계산량이 줄었다. 피라미드 구조를 이용하여 전역 움직임을 구하고 이전 오차영상을 이용하여 국부 움직임을 보정한 제안한 방법 2의 경우는 기존의 6 및 8-파라미터에 비해서 모두 계산량이 적고 PSNR도 높았다.

그림 8은 표 1의 6-파라미터의 경우에 대해 PSNR 성능을 그래프로 나타내었다. 피라미드 및 국부오차

성분을 이용한 제안한 방법이 기존의 방법에 비해 우수하였고 계산량도 적었다. 저해상도를 이용한 경우가 가장 계산량이 적지만 국부 오차의 증가로 PSNR이 낮았다.

표 2는 추정화소 수를 비교하여 움직임 추정의 계산량을 비교하였다. 저해상도인 경우가 가장 계산량이 적지만 국부오차 증가하여 움직임 추정에 적합하지 않다. 국부 움직임을 고려한 제안한 방법은 기존의 방법에 비해 계산량이 적고 영상의 국부 움직임 및 전역 움직임 특성을 모두 만족하였고 PSNR도 높았다.

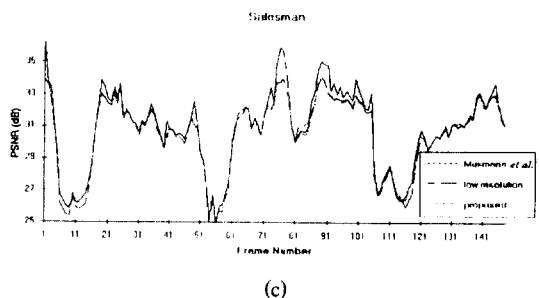
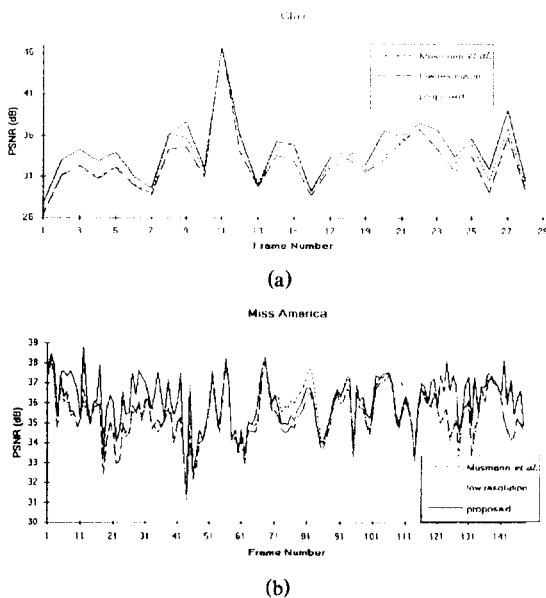


그림 7. 세가지 움직임 추정 방법에 의한 복원 영상의 PSNR 비교

(a) Clair (b) Miss America (c) Salesman

Fig 7. PSNR comparison of images reconstructed by three motion estimation methods.

(a) Clair. (b) Miss America. (c) Salesman.

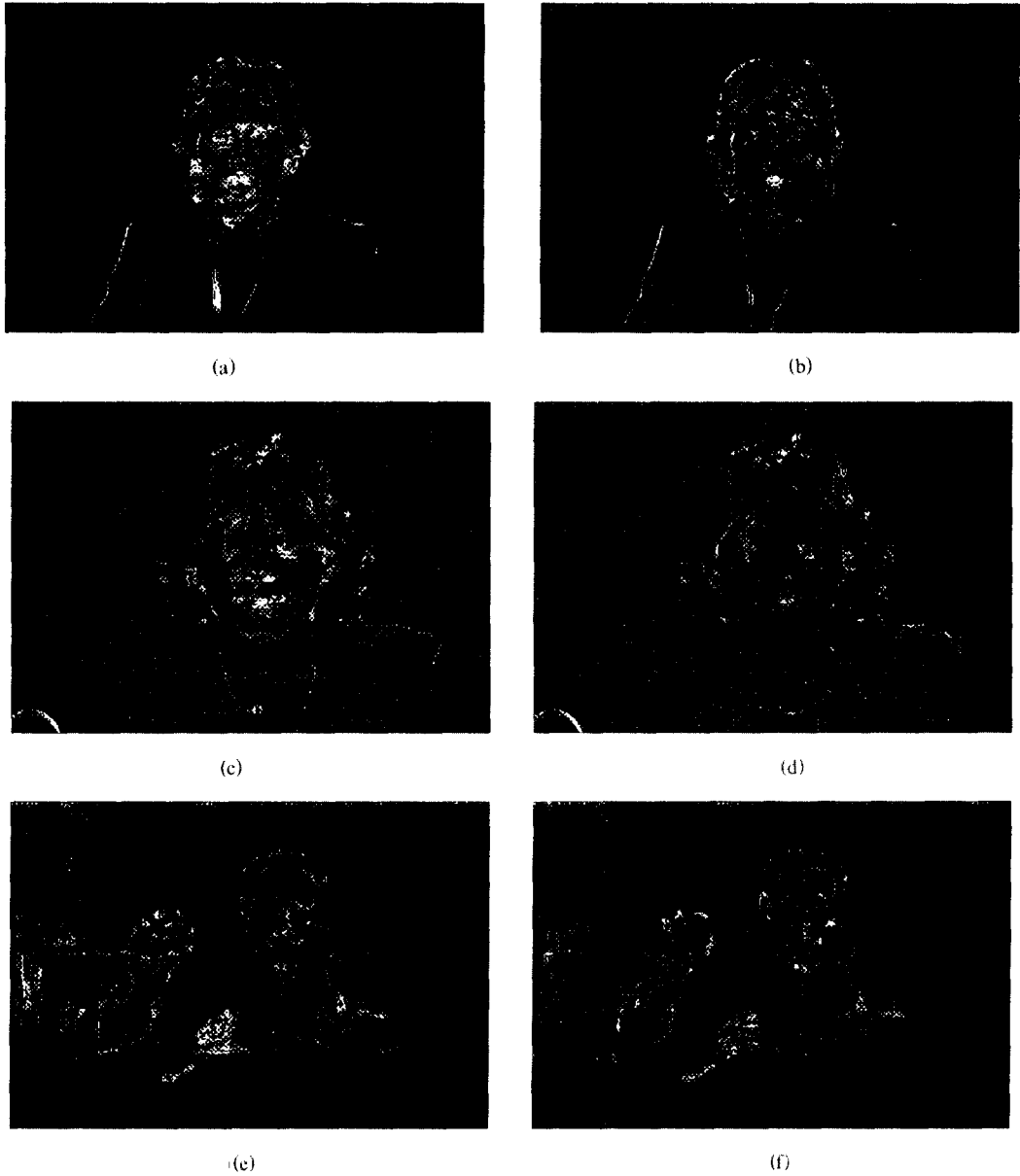


그림 8. 오차영상

(a) Clair (Schiller와 Hötter) (b) Clair (제안) (c) Miss America (Schiller와 Hötter) (d) Miss America (제안) (e) Salesman (Schiller와 Hötter) (f) Salesman (제안)

Fig 8. Error images.

(a) Clair (Schiller and Hötter). (b) Clair (proposed). (c) Miss America (Schiller and Hötter). (d) Miss America (proposed). (e) Salesman (Schiller and Hötter). (f) Salesman (proposed).

표 2. 움직임 추정을 위한 측정 화소수 비교  
Table 2. Comparison of the number of observation points for motion parameter estimation.

실험영상	Clair	Miss America	Salesman
Musmann 등의 방법 <sup>[4]</sup>	4806(100%)	24330(100%)	29317(100%)
제안한 방법	2175(45%)	8835(36%)	15291(52%)
저해상도를 이용한 방법	1201(25%)	6082(25%)	7329(25%)

표 3은 기존 및 제안한 방법의 부호화 성능을 나타낸다. 기존의 방법은 움직임을 Musmann 등의 방법을 사용하고 MF 영역은 Schiller와 Hötter의 방법<sup>[5]</sup>을 사용하였다. 제안한 방법은 피라미드 구조를 이용하고 MF 영역은 DFD 피라미드를 사용하였다. 15 프레임마다 원영상을 삽입하여 inter 프레임 부호화만을 고려하였다. 제안한 방법이 압축율 및 PSNR이 높았다. 기존의 방법은 국부오차 누적을 고려하지 못한데 반해 제안한 방법은 이전 오차를 이용한 국부 오차의 보정으로 오차의 누적을 효과적으로 줄였다.

표 3. 부호화 성능비교  
Table 3. Performance comparison of each coding method.

실험영상	Clair		Miss America		Salesman	
	CR	PSNR	CR	PSNR	CR	PSNR
Schiller와 Hötter의 방법 <sup>[5]</sup>	99	38.37	148	35.04	71	33.54
제안한 방법	151	38.63	169	35.21	94	34.93

그림 8은 복원화질을 비교하기 위해 부호화 후의 복원영상의 오차에 대해 절대값의 15배로 표시하였다. 기존의 방법은 눈, 코, 입 등의 MF 영역과 그밖의 국부오차 성분이 크다. 그러나 피라미드 구조를 이용한 제안한 방법의 경우는 이들 영역의 오차 성분이 적었다.

### V. 결 과

본 논문에서는 피라미드 구조를 이용하여 전역 움직임을 구하고 국부 움직임을 이전오차 영상을 이용하는 방법을 제안하였다. 그리고 MF 영역에 대해서는 오차 영상의 특성을 다양하게 이용하기 위하여

DFD 피라미드를 이용하여 보상하였다. 모의실험 결과로부터 기존의 방법에 비해 압축율, PSNR 및 계산량 측면에서 우수한 결과를 보였다.

기존의 방법의 경우 잡음에 의한 영향이 크고 또한 영역내에 여러 움직임이 존재하여 국부 움직임 특성이 존재할 경우 전역적 움직임 파라미터에 의한 움직임 표현에 한계가 있다. 계층적 구조로 국부 오차 영역을 다시 움직임 보상한다 하더라도 이전의 보상이 정확하지 않으므로 오차가 증가하였다. 이것이 프레임마다 진행되어 부호화할 때 오차의 누적으로 정확한 움직임 추정이 어렵다. 이에 반해 제한한 방법은 움직임을 전역 움직임과 국부 움직임으로 나누어 구하였고 전역 움직임 추정시에는 국부 움직임 특성을 어느 정도 반영하며 피라미드 구조를 이용하여 적은 계산량에서도 비교적 정확한 전역 움직임을 구하였다. 이 경우 가중치를 두어 전역 움직임에 큰 영향을 주는 국부 움직임 특성은 제거하였다. 전역 보상된 영상을 이용하여 국부 움직임을 보정하기 위하여 이전오차의 정합을 이용하여 비교적 정확한 보정 파라미터를 구할 수 있었다. 특히 이러한 국부 움직임 파라미터는 기존의 방법에서 고려하지 못한 프레임간의 누적 오차를 효율적으로 줄였다. 또한 기존의 방법의 경우 누적 오차 성분으로 MF 영역에 비트량이 가중되어 이 부분의 부호화 효율이 낮았지만 제안한 방법은 DFD 피라미드에서 많은 정보량을 이용하기 때문에 이를 개선시켰다.

추후 과제는 파라미터 양자화에 의한 복원화질의 성능을 분석하고 보다 효율적인 피라미드 구조의 구현을 위한 연구이다.

### 참 고 문 헌

1. A. K. Jain, "Image data compression: A review," *Proc. IEEE*, vol. 69, pp. 349-389, Mar. 1981.
2. ISO/IEC JTC/SC2/WG11 "MPEG video simulation model three," MPEG90/041, July 1990.
3. M. Kunt, A. Ikonopoulos, and M. Kocher, "Second-generation image coding techniques," *Proc. IEEE*, vol. 73, no. 4, pp. 549-574, Apr. 1985.
4. H. G. Musmann, M. Hötter, and J. Ostermann, "Object-oriented analysis-synthesis coding of mov-

- ing images," *Signal Processing: Image Commun.*, vol. 1, no. 2, pp. 117-138, Oct. 1989.
5. H. Schiller and M. Hötter, "Investigations on colour coding in an object-oriented analysis-synthesis coder," *Signal Processing: Image Commun.*, vol. 5, no. 4, pp. 319-326, Oct. 1993.
  6. M. Hötter and R. Thoma, "Image segmentation based on object oriented mapping parameter estimation," *Signal Processing*, vol. 15, no. 3, pp. 315-334, Oct. 1988.
  7. R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-29, no. 6, pp. 1147-1152, Dec. 1981.
  8. 김준식, 박래홍, "물체지향분석 및 합성부호화에서 가산 투영을 이용한 영상분석 기법," 전자공학회 논문지-B, 제31권-B권, 제8호 pp. 87-98, 1994년 8월.
  9. J.-S. Kim and R.-H. Park, "Mapping parameter estimation using integral projections and segmented moving objects in object-oriented analysis-synthesis coding," *Optical Engineering*, vol. 35, no. 1, pp. 156-165, Jan. 1996.
  10. R. Thoma and M. Bieling, "Motion compensating interpolation considering covered and uncovered background," *Signal Processing: Image Commun.*, vol. 1, no. 2, pp. 191-212, Oct. 1989.
  11. 조대성, 박래홍, "블럭기반 움직임 벡터필드 보정을 이용한 객체지향 부호화기," 신호처리 합동 학술대회 논문집, 제8권 제1호, pp. 971-975, 1995년 9월.
  12. 조대성, 박래홍, "블럭기반 움직임 벡터와 오차영상 보상을 이용한 물체 지향 부호화기," 전자공학회 논문지-B, 제33권-B권, 제3호, pp. 96-108, 1996년 3월.
  13. D.-S. Cho and R.-H. Park, "An object-oriented coder using block-based motion vectors and residual image compensation," *IEEE Trans. Circuits and Systems for Video Tech.*, revised.
  14. J.-H. Moon and J.-K. Kim, "On the accuracy and convergence of 2-D motion models using minimum MSE motion estimation," *Signal Processing: Image Commun.*, vol. 6, no. 4, pp. 319-333, Aug. 1994.
  15. C. Cafforio and F. Rocca, "The differential method for image motion estimation," in *Image Sequence Processing and Dynamic Scene Analysis*, T. S. Huang, ed., Springer, Berlin, pp. 104-124, 1983.
  16. J. M. Odobez and P. Bouthemy, "Robust multiresolution of parametric motion models," *J. Visual Commun. Image Representation*, vol. 6, no. 4, pp. 348-365, Dec. 1995.
  17. Y. Fisher, D. Rogovin, and T. P. Shen, "Fractal (self-VQ) encoding of video sequences," in *Proc. SPIE Visual Commun. Image Processing '94*, vol. 2308, pp. 1359-1370, Chicago, Illinois, Sep. 1994.
  18. A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston: Kluwer Academic Pub., 1992.
  19. J. J. Chae, S. B. Chae, W. Y. Choi, and R.-H. Park, "Effective contour coding technique using the  $2 \times 2$  block," in *Proc. Int. Symp. Information Theory and Its Applications 1990*, pp. 27-30, Waikiki, Hawaii, Nov. 1990.



趙大星(Dae-Sung Cho) 정회원  
 1971년 9월 5日生.  
 1994년 2월: 서강대학교 전자공학과 졸업(공학사).  
 1996년 2월: 서강대학교 대학원 전자공학과 졸업(공학석사).  
 1996년~현재: 삼성 종합기술원 근무.

※주관심분야: 영상부호화, 컴퓨터 비전, 영상 통신 등임.

朴來弘(Rae-Hong Park) 정회원  
 현재: 서강대학교 전자공학과 교수.  
 한국통신학회 논문지 제20권 제10호 참조