

영상처리 및 인식 기반의 MPEG-4 영상 압축 기술

文 柱 禧

現代電子産業株式會社
멀티미디어研究所 研究6室

I. 서 론

개인용 컴퓨터의 대중화, 디지털 전송 기술의 발전, 고화질 디스플레이 장치 실현, CD의 보편화 및 메모리 디바이스의 저가격화등에 따라 가전의 개념이 영상(image)이라는 정보 매체를 중심으로 급속히 변화되고 있다. 이는 소비자들에게 영상이 주는 정보 전달 효과가 가장 크다는 사실을 반영하는 것으로서, 최근의 관련 기술 발전에 따라 그동안 제공할 수 없었던 여러 영상 서비스에 대한 수요를 충족하는 방향으로 전자 업계가 재편되는 양상을 보이고 있다. 이러한 예로서 고화질 TV(HDTV), 주문형 비디오(VOD: Video On Demand), 영상전화기(Videophone), 영상회의 시스템(Video Conferencing System), 비디오 CD(Video Compact Disc), DBS(Direct Broadcasting Satellite), 멀티미디어 시스템(Multimedia System), 원격 강의(Remote Education), 영상 PDA(Personal Digital Assistant)등을 들 수 있다. 이러한 영상 관련 제품의 개발은 대부분 영상 부호화와 관련된 기술 표준인 ITU-T의 H.320, H.324 및 ISO의 MPEG-1, MPEG-2 등을 따르고 있다.

현재 ISO/IEC JTC1/SC29/WG11에서 활발히 작업이 이루어지고 있는 MPEG-4(Moving Picture Experts Group-Phase 4) 표준화가 있다^[1]. MPEG-4 표준화 작업은 10Kbps 정도의 아주 낮은 비트율로 부터 수 Mbps의 높은 비트율에 이르기까지 넓은 비트율에서 사용가능한 오디오-비디오 데이터 부호화 기술의 표준을 정하려는 것이다. MPEG-4는 자연 영상(natural image)은 물론 인공 영상(synthetic image) 데이터의 통신(communication), 액세스(access), 그리고 조작(manipulation)을 위한 새로운 방법들을 지원하기 위한 부호화 표준이 될 것이다. MPEG-4에서 지원될 수 있는 새로운 기능들의 대부분은, 단순히 블록 단위로 부호화를 수행하도록 했던 지금까지의 표준들(예를들어 H.261, H.263, MPEG-1, MPEG-2)과는 달리, 영상의 내용을 고려하는(content-based)

부호화 방식으로 부터 비롯된다^[2,3]. 따라서 물체의 인식 기술이 요구되며 인식된 물체의 처리 방법이 중요하게 대두된다. 1995년 11월과 1996년 1월에 걸쳐 각국의 여러 기관에서 제안된 알고리즘(algorithm)과 툴(tool)들의 시험이 있었으며, 시험 결과^[4,5]를 토대로 현재 MPEG-4 VM (Verification Model) 버전2.1^[6]이 나와있는 상태이다. VM의 가장 큰 특징은 영상 정보가 물체를 의미한다고 볼 수 있는 VOP(Video Object Plane) 단위로 표현될 수 있으며, VOP를 기준으로 부호화되고, VOP를 사용하여 영상을 재현할 때 여러가지 조작이 가능하다는 것이다. VOP는 움직임이 있는 하나 혹은 그 이상의 물체들을 의미한다. 따라서 VOP는 임의의 형태인 모양 정보로 표현되고 부호화가 이루어진다. 이것은 기존의 표준이 프레임(frame)이라는 단위로 부호화되고 조작되었던 것과 가장 큰 차이점이라 할 수 있다.

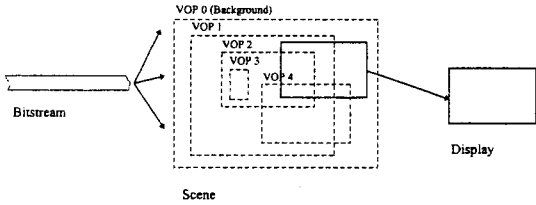
MPEG-4 표준화에서는 VM의 성능을 향상시키기 위한 여러가지 기술들이 공동실험(Core Experiment)으로서 제안되어 실험되고 있다. VM 및 공동실험되고 있는 부호화 기술들중 고압축율을 주로 고려하는 내용 중심의 기술들을 크게 나누어 보면, 임의의 모양 정보를 부호화하는 기술, VOP의 경계에 속한 영역의 신호 정보 부호화 기술, 전역 움직임 또는 다양한 크기/모양을 가진 영역을 고려하는 움직임 추정 기술들로 나눌 수 있다. 본 논문에서는 MPEG-4 VM의 개념에 대해서 살펴보고, VM과 공동실험으로 제안된 내용 중심의 물체 부호화 기술들중 기존의 표준들과 크게 차이를 보이는 모양정보 및 신호정보 부호화 기법에 대하여 차례로 살펴보기로 한다. 또한 앞으로의 MPEG-4 표준화 동향에 대하여 알아보기로 한다.

II. MPEG-4 VM의 부호화 방식

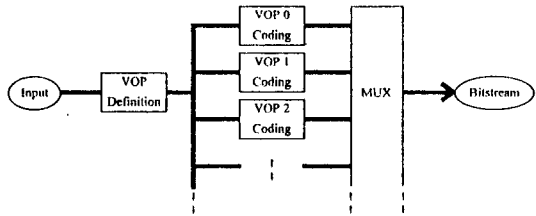
VOP는 사용자가 부호화된 비트열(bit stream)로부터 꺼낼 수 있고 조작할 수 있는 비트열의 구성요소로서 정의된다. 송신부에서는 VOP 자체의

재현을 위한 부호화 정보와 함께, 전송되는 VOP가 화면에 출력될 시간과 위치등의 합성(composition) 정보를 보낼 수 있다. 수신부에서는 VOP의 재현후에 사용자가 임의로 재현된 VOP의 합성정보를 변경하여 출력될 영상 화면을 구성할 수도 있다. 그림 1은 이와 같은 개념을 설명하고 있다. 비트열로부터 각 VOP를 재현해내고 VOP를 둘러싼 외곽 사각형에 의해 영상(scene)을 합성할 수 있다. 이때, 사용자가 임의로 합성정보를 변경할 수도 있어서 출력될 화면의 구성과 위치등을 변화시킬 수 있다. VOP는 영상에서 의미를 가지고 있는 하나 혹은 그 이상의 물체들로 구성될 수 있는데, 신호정보(Y, U, V)외에 모양정보(shape information)도 함께 필요하다. VOP 개념을 사용하는 VM은 기존의 부호화 기술들을 포함한다고 볼 수 있다. 즉, 연속영상(image sequence)에서 매 영상이 항상 하나의 동일한 직사각형 VOP로만 구성되고 출력될 시간 간격도 동일하다면, 이것은 기존의 프레임(frame) 단위 부호화 기술과 같게 된다. 한편, 출력 화면 구성을 위해 VOP들의 합성에 사용되는 모양정보는 VOP들의 투명도(transparency)를 나타내는 알파값(alpha value)들로 구성되는 알파평면(alpha plane)에 의해 표현된다. 투명도는 두 VOP가 합쳐질 때 뒤에 있는 VOP가 어느 정도 보일 수 있는가를 나타내는 것으로서, 알파값이 0이면 뒤에 있는 VOP가 완전히 보이는 것을 의미하며, 알파값이 255이면 뒤에 있는 VOP가 완전히 가려짐을 의미한다. 알파값이 가질 수 있는 범위에 따라 8비트 알파평면(gray scale alpha plane)과 1비트 알파평면(binary alpha plane)으로 나눌 수 있다. 8비트 알파평면은 알파값들이 0~255의 값(8 비트)으로 표현되며, 1비트 알파평면은 단순히 물체 밖을 0으로 물체 안을 255로 표현한다. 모양정보를 이용하여 VOP들을 합성하는 방법은 부록A에 나타나있다.

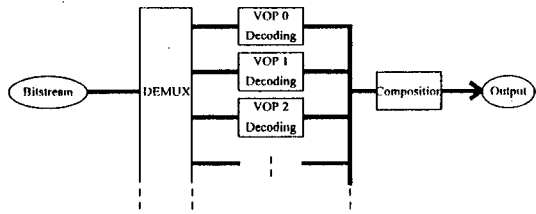
그림 2는 VOP를 사용하는 MPEG-4 VM의 송신기와 수신기 구조를 나타낸다^[6]. VOP 정의부에서는 기존의 영상신호를 VOP 단위로 분할하게 되는데, 자동적으로 원하는 물체를 분할하므로써 만들 수도 있으며 인위적으로 만들어 낼 수도 있



(그림 1) VOP 재현 및 화면 구성



(a) 송신기



(b) 수신기

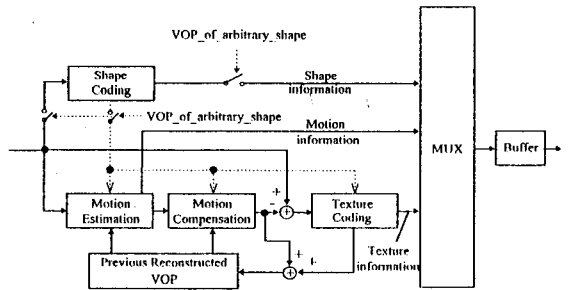
(그림 2) VM의 송신기와 수신기 구조

다. VOP를 분할해내는 방법은 MPEG-4의 표준화 대상은 아니다. 현재의 VM 구조에서는 모든 VOP들이 동일한 부호화 방법에 따라 부호화되며, VOP간에는 어떠한 정보의 교환도 없도록 규정되어 있다. 이는 수신부가 어떤 VOP를 복호화하는데 다른 VOP의 정보를 이용하지 않아도 되므로 여러가지 조작 기능상의 편리함을 얻을 수 있으나, 부호화 효율성은 상당히 저하된다.

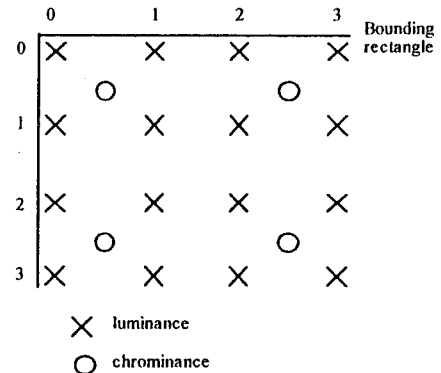
부호화가 시작되면 모든 VOP들에 대해서 그림 3과 같은 동일한 부호화 방법을 적용한다. VOP가 임의의 물체가 아니라 기존의 프레임으로 구성된다면 모양정보 부호화와 관련된 블록은 필요치 않

으므로 모든 스위치가 오프(off)된다. 임의의 모양을 가진 물체인 경우에도 움직임 보상과 신호정보 부호화는 기존의 방식과 마찬가지로 16x16 혹은 8x8과 같은 블록 단위로 수행된다. 그러나 일반적으로 VOP의 모양정보가 전송될 경우, 물체 경계에 있는 블록의 내부에는 일부만이 물체로 채워지므로 임의의 모양에 맞는 신호정보 부호화 방식을 사용해야 한다.

VOP를 둘러싼 외곽 사각형(bounding rectangle)의 휘도(luminance)와 색(chrominance) 신호는 그림 4와 같이 4:2:0 형식으로 구성되며, 프레임상의 좌표로 볼 때 휘도신호를 둘러싼 사각형의 왼쪽 위 모서리 좌표는 VOP를 포함하는 최소 사각형의 왼쪽 위 모서리 좌표 보다 크지 않은 짝



(그림 3) VOP의 부호화기 구조



(그림 4) VOP 외곽 사각형내의 휘도신호와 색신호 위치

수 위치로 만들어진다. 즉, 그림 4의 외곽 사각형에서 좌표 0은 프레임상에서의 좌표가 짝수인 위치이다. 색신호의 왼쪽 위 모서리 좌표는 휘도 좌표를 2로 나눈 위치가 된다. 이것은 임의의 모양을 가진 물체의 휘도와 색신호간의 위상차(phase difference)를 정확히 유지하기 위하여 필수적이다.

III. 모양정보 부호화(Shape Coding)

모양정보는 0~255사이의 알파값들을 가지고 있는 알파평면(alpha plane)에 의해 표현되는데, VM 버전 2.1에 현재 포함되어 있는 quadtree 부호화 방식^[6,7]과 polygon/spline 부호화 방식^[8]으로 나눌 수 있다.

1. Quadtree 부호화 방식

VOP 외곽 사각형의 알파평면을 먼저 16x16 블록들의 정수배가 되도록 확장한다. 이때, 확장된 부분들의 알파값들은 모두 0으로 채워진다. 확장된 알파평면은 16x16 블록인 알파블록(alpha block)들로 분할되고, 각 알파 블록들은 quadtree 형태로 부호화된다.

1비트 모양정보 부호화

1비트 모양정보 부호화를 위한 quadtree 구조는 그림 5에 나타나 있다. 부호화 과정은 다음과 같다:

단계 1: 비트율 제어를 위한 모양정보 수정(무손실 부호화시에 단계 1은 생략)

단계 2: 레벨 3에서의 부분블락(sub-block) 인덱스 계산

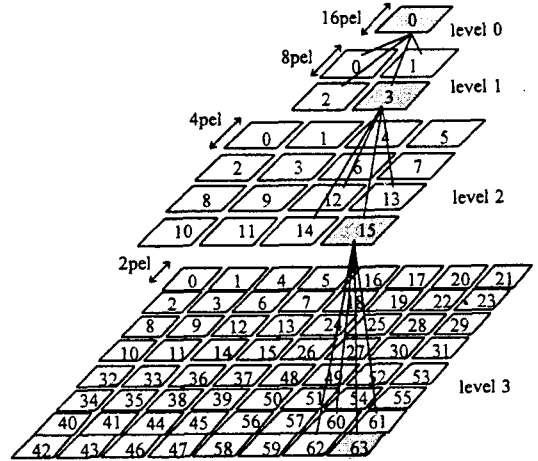
단계 3: 상위 레벨들로의 grouping

3.1 레벨 3에서 레벨 2로의 부분블락 grouping

3.2 레벨 2에서 레벨 1로의 부분블락 grouping

3.3 레벨 1에서 레벨 0로의 부분블락 grouping

단계 4: 인덱스 부호화

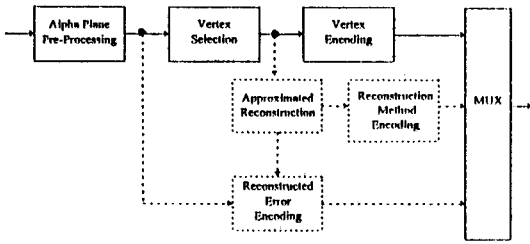


(그림 5) 1비트 모양정보 부호화를 위한 Quadtree 구조

4x4 혹은 2x2 부분블락 단위로 모양정보를 수정(rounding)한 뒤 수정오차가 어떤 문턱값보다 작으면 해당 부분블락을 모두 0 혹은 255의 알파값으로 대체하므로써 발생 비트율을 제어한다. 이것은 부분블락의 알파값들이 모두 0이거나 255이면 해당 부분블락의 인덱스를 상위 레벨의 인덱스가 표현해 줄 수 있도록 인덱스 계산을 수행하므로써 가능하다. 즉, 많은 오차를 허용할 수록 알파값들이 모두 0이거나 255인 부분블락들이 많아지게 되고 전송해야할 인덱스 수가 줄어들게 된다. 상위 레벨의 인덱스는 하위 레벨 인덱스의 함수로서 계산되는데, 인덱스의 통계적 분포가 특정한 값들로 집중되도록 하위 레벨의 인덱스 값들을 서로 교환(swapping)하는 단계를 거치게 된다. 각 레벨의 인덱스들은 가변길이 부호에 의해 부호화된다,

8비트 모양정보 부호화

8비트 모양정보 부호화 방법은 1비트 모양정보 부호화 과정과 유사하다. 단지, 그림 5에서 레벨 2부터 grouping이 시작되며, 레벨 2의 각 4x4 부분블락들은 벡터 양자화기에 의해 부호화된다. 벡터 양자화로 인하여 항상 손실 부호화가 이루어지게 된다.



〈그림 6〉 Polygon/Spline 부호화 방식의 구조

2. Polygon/Spline 부호화 방식

본 방식은 1비트 모양정보를 부호화하는 방법이며, 일반적으로 그림 6의 구조를 가지게 된다. 데이터를 줄이기 위해 항상 손실 부호화가 이루어지며, 알파평면 전처리(alpha plane pre-processing), 정점 선택(vertex selection), 정점 부호화(vertex encoding), 근사 재현(approximated reconstruction), 재현 방법 부호화(reconstruction method encoding), 재현 오차 부호화(reconstruction error encoding) 과정들로 구성된다.

가. 알파평면 전처리: CCIR601 형식에서 (Q) CIF 형식으로의 down-conversion과 같은 해상도 변환, 8비트 알파평면의 1비트화(binari-zation), 1비트 마스크(binary mask)로부터의 윤곽선(contour) 추출 등으로 구성된다. 윤곽선 추출은 8-neighborhood connectivity 관계를 사용한다.

나. 정점 선택: 이 방법은 표준화 대상이 아니나 현재 MPEG-4 실험에서는 반복 선택 기법(iterated refinement technique)을 사용하고 있다. 이 방법의 진행과정은 다음과 같다.

- (1) 한쌍의 윤곽선 화소간의 거리가 가장 긴 한쌍의 윤곽선 화소를 찾는다. 이 두 점을 연결한 선을 주축(main axis)으로 한다.
- (2) 주축 양쪽에서 주축과 가장 거리가 먼 윤곽선 화소를 각각 하나씩 찾는다. 이 두 점과 주축을 구성하는 두 점, 총 4개의 정점을 초기 정점으로 한다.
- (3) 정확도 척도(accuracy criterion)인 d_{max} (두 정점들을 연결하는 선분(line seg-

ment)와 두 정점 사이의 원윤곽선(original contour)과의 허용되는 최대 오차거리를 만족할 때까지 각 선분에서의 최대 오차를 갖는 점들을 정점으로 삼입

이러한 polygon 근사화에서는 d_{max} 로 재현 모양정보의 질(quality)과 발생 데이터량을 조절할 수 있다. 한편, 물체 내부와 외부에서 d_{max} 를 달리하는 방법도 고려할 수 있다.

다. 정점 부호화: 선택된 정점은 모양정보 재현을 위해 수신단에 전송되어야 한다. 현재 고려되는 방법으로는 크게 3가지 방식이 있다.

- (1) 절대 위치 전송(absolute addressing : default): 가장 단순한 방식이다. 소요되는 비트수는 영상의 크기에 따라서 결정된다. 예를들어 QCIF 영상(176 pixels/line x 144 lines/frame)인 경우, 정점당 16 비트가 소요된다.
- (2) 상대 위치 전송(relative addressing): 정점들의 위치를 이전 정점 위치와의 상대적 위치를 구하고 이를 전송하는 방식이다. 초기 정점(initial vertex)은 이전 정점을 영상의 중심점(image center point)으로 가상한다. 비트열 발생을 위해서는 syntax adaptive arithmetic coding(SAAC)혹은 Huffman coding을 이용하여 전송한다.
- (3) 물체 중심 방법: 이 방법도 상대 위치 전송 방법을 기본으로 한다. 단, 물체의 특성을 고려하게 된다. 우선 선택된 모든 정점에서 상대적 위치를 계산한다. 상대위치의 거리가 너무 멀리 떨어져 있으면 두 정점 사이에 추가의 정점을 둔다. 이렇게 하면 상대적 위치가 너무 멀리 떨어져 있으므로써 발생할 수 있는 가변길이 부호의 길어짐을 막을 수 있다. 이렇게 하여 구해진 정점의 개수와 초기 정점을 부호화한 뒤, 가까이 있는 정점들을 차례로 부호화하게 된다. 이때, 상대위치는 그림 7과 같은 8 구획(octant-based) 표현을 이용하여 부호화한다. 그림에서 R(x)와 R(y)는 각각 상대위치의 x-성분, y-성분을 나타낸다. 상대위치가 어

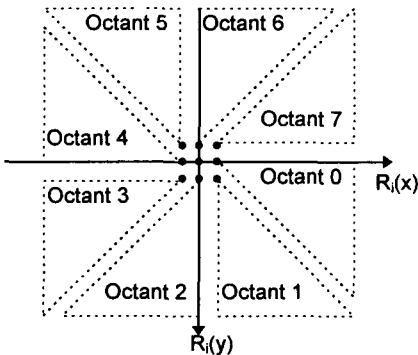
면 구획에 들어가는지의 정보를 전송한다. 그리고 각 성분의 크기를 전송한다. 각 정보의 전송은 SAAC 또는 Huffman coding 을 이용한다.

라. 근사 재현 : 정점이 주어졌을 때 윤곽선을 어떻게 재현하는가에 대한 것이다. 여기에 관한 방법으로는 다음과 같은 3가지 방법이 고려되고 있다.

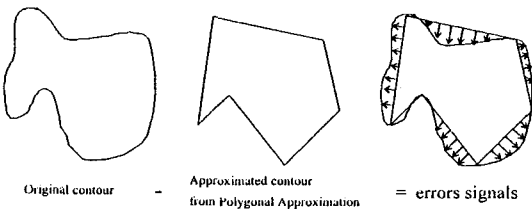
- (1) polygon만 사용하는 경우
- (2) spline만 사용하는 경우
- (3) polygon과 spline을 동시에 사용하는 경우 : 이 경우 polygon을 사용하는가, spline을 사용하는가 하는 flag 전송이 요구된다. 이 결정을 정점당 수행하는 경우와 mask 전체에서 하는 경우의 두가지 모두 고려되고 있다.

마. 재현 오차 부호화 : 그림 8에서와 같이 재현 윤곽선과 원윤곽선(original contour)의 오차를 1차원 신호로 만들고 이를 부호화한다.

그림의 각 선분에서 화살표들이 1차원 오차 신호를 구성한다. 이 1차원 오차신호에는 1차원 DCT(discretecosine transform) 이나 DST(dis-



<그림 7> 8구획을 사용하는 상대 위치 표현 방법



<그림 8> 윤곽선 재현 오차 부호화

crete sine transform)등을 적용하고 계수를 양자

화하여 수신단에 전송할 수 있다.

바. 윤곽선 예측 부호화 방식

시간축상에서 연속되는 같은 물체의 모양정보 간에는 많은 유사성이 존재하며, 이를 이동보상예측(motion compensated prediction)에 의해 제거함으로써 부호화 이득을 얻을 수 있다. 이러한 목적으로 제안된 예측 부호화 방식들은 크게 두 가지 방식으로 나누어 진다.

(1) 정점포 추가(vertex list update) : 이전 윤곽선 부호화에 사용된 정점들을 가지고 추정된 움직임 정보를 이용하여 예측한다. 예측된 정점들이 주어진 dmax를 만족하는가를 판단하여 이를 만족하지 못하는 경우는 제거한다. 나머지 유지정점들을 초기 정점들로 하여 다시 주어진 정확도를 만족할 때까지 새로운 정점들을 삽입하게 된다. 전송이 필요한 정보로는 제거/유지를 나타내는 flag, 새로 삽입되는 정점들의 위치정보 등이다.

(2) 모양 불일치 방법(Shape mismatch method) : 정점들을 예측하는 앞의 방식대신 영역(또는 그의 윤곽선)을 이동보상예측한다. 예측된 영역과 원영역간의 오차를 구하고 오차가 있는 영역의 원윤곽선을 부호화하여 전송한다.

이 밖에 예측 부호화에서는 모양정보를 위한 이동 모형(motion model) 또는 움직임 추정 방식(motion estimation)의 연구, intra 부호화/예측 모드 결정 등의 연구가 필요하다.

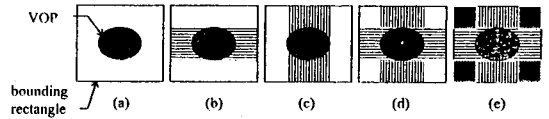
IV. 경계 영역 신호정보 부호화(Boundary Texture Coding)

VOP는 임의의 형태를 가지므로 경계 영역의 신호정보 부호화 방식이 필요하다. 본 논문에서는 경계 영역의 신호정보를 DCT를 근간으로 하여 부호화하는 방식들에 대해 살펴본다^[6, 9, 10, 11]. 물체의

경계를 구분하여 부호화하기 때문에, 일반적으로 경계 영역에서의 순수한 물체 신호정보는 정방형 블락들로만 분할될 수 없다. 즉, 경계 영역에서는 정방형 블락의 일부에만 물체 신호가 채워지게 된다. 따라서 이와같은 형태의 블락들에 대한 효율적 부호화 기법들이 연구되고 있다.

1. 반복 패딩 기법(Repetitive Padding Technique)

현재 VM 버전 2.1에서 사용되고 있는 방법으로 서 물체 신호정보가 채워지지 않은 정방형 블락의 일부 영역을 그림 9와 같이 물체 경계의 신호들을 사용하여 채운 뒤 DCT를 수행하게 된다^[6]. 그림 9의 (a)와 같은 VOP의 패딩과정을 설명하면 다음과 같다:



(그림 9) 반복 패딩 과정

- (1) (b)와 같이 수평방향으로 스캔하여 VOP 왼쪽과 오른쪽에 있는 물체 외부 화소는 각각 왼쪽경계정보와 오른쪽 경계정보로 패딩하고, VOP 사이에 있는 외부 화소는 그 화소의 왼쪽과 오른쪽 경계 정보의 평균값으로 패딩한다.
- (2) (c)와 같이 수직방향으로 스캔하여 VOP 위쪽과 아래쪽에 있는 물체 외부 화소는 각각 위쪽경계정보와 아래쪽 경계정보로 패딩하고, VOP 사이에 있는 외부 화소는 그 화소의 위쪽과 아래쪽 경계 정보의 평균값으로 패딩한다.
- (3) (d)와 같이 단계 (1)과 (2)가 공통적으로 적용되었던 화소는 단계 (1)과 (2)의 패딩 값들 평균으로 대체한다.
- (4) (e)와 같이 단계 (1), (2), (3) 이후에도 패딩되지 않은 화소는 단계 (3)까지의 패딩된 영상을 기준으로 단계 (1), (2), (3) 과정을 순서대로 반복하여 패딩한다. 이때, 단계 (1), (2)의 과정중, 평균값 대신 왼쪽 화소값과 위쪽 화소값이 각각 사용된다.

수신단에서는 역DCT를 수행한 후, 모양정보를 사용하여 물체 내부의 신호만을 추출하게 된다. 반복 패딩 기법은 물체의 경계 신호만을 사용하게 되므로, 블락에 포함된 물체 신호의 전반적 특성 대

신 경계 신호 특성에 의존하는 문제가 발생한다. 특히, 이동 보상 오차의 반복 패딩에서는 경계를 따라 오차값이 큰 값을 갖는 경우가 많으므로 반복 패딩에 의하여 채우게 될 경우 의미없는 정보량이 추가되어 DCT 효율의 저하를 가져온다.

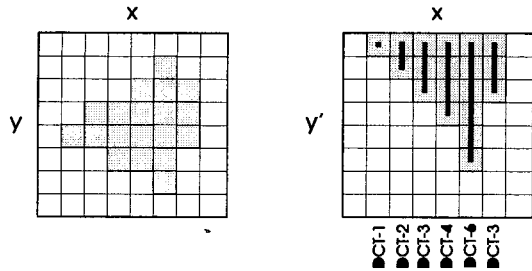
2. SADCT(Shape Adaptive DCT)^[9]

물체내부 영상신호만을 이용하여 부호화한다. 부호화해야할 VOP 내의 영상정보가 그림 10의 (a)와 같이 구성되어 있는 블락을 예로써 설명하면 다음과 같다:

- (1) (a)와 같은 영상정보를 먼저 위쪽으로 이동시켜 (b)와 같이 정렬한 후 VOP 내의 영상신호만을 이용하여 수직방향으로 1차원 DCT를 한 후 위쪽부터 아래로 저주파 성분부터 고주파성분까지 (c)와 같이 정렬한다. 그림에서 DCT-N은 N-point DCT를 의미한다.
- (2) 1차원 수직 DCT 이후의 결과인 (d)를 (e)와 같이 왼쪽으로 이동시켜 정렬한 후 VOP 내의 영상신호 즉, 1차원 수직 DCT 계수들만을 이용하여 수평방향으로 1차원 DCT를 한 후 왼쪽부터 오른쪽으로 저주파 성분부터 고주파성분까지 (f)와 같이 정렬한다.
- (3) 단계 (1), (2)를 모두 거친 2차원 SADCT 계수 (f)는 실제 데이터가 있는 영역만을 (g)와 같이 지그재그(zigzag) 스캔하여 가변길이 부호화(VLC)한다.

SADCT는 임의의 모양을 가진 물체를 의미하는 VOP 개념에 일치하며, VOP내의 신호정보만을 고려하여 부호화하므로써 실제의 신호 중복성만을 효율적으로 제거할 수 있다. 또한, 변환 계수를 전송할 때 물체 내부만을 지그재그 스캔하므로써 0(zero) 값의 연속길이를 줄일 수 있어서 유리하다.

현재 가장 주목받고 있는 기술중의 하나이지만, 매번 적용되는 DCT의 구조가 1차원 DCT를 위한 화소수 N에 따라 변해야 한다는 다소 복잡한 과정이 요구된다. 따라서 더욱 간단한 방법들이 제안되고 있다.



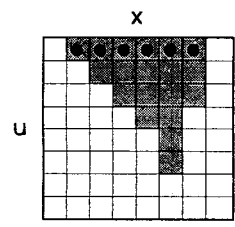
1	2	6	7	15	16		
3	5	8	14	17			
4	9	13	18				
10	12						
11							
19							

(g)

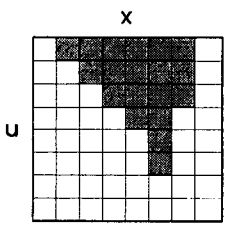
〈그림 10〉 SADCT 부호화 예

(a) Original Segment

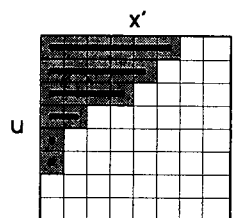
(b) Ordering of Pels and Vertical SADCT Used



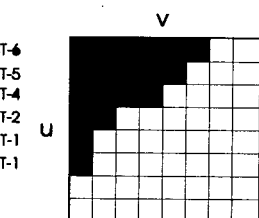
(c) Location of Pels after Vertical SADCT



(d) Location of Pels Prior to Horizontal SADCT



(e) Ordering of Pels and Horizontal SADCT Used



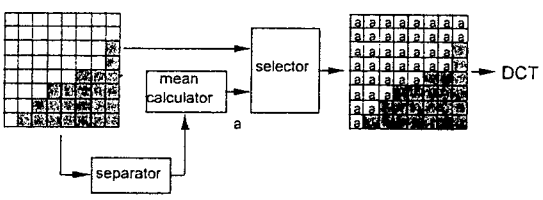
(f) Location of 2-D SADCT Coefficients

3. 평균값 대체 기법 (Mean Replacement Technique)^[10]

그림 11과 같이 VOP 내의 신호정보 평균값을 이용하여, VOP 밖의 영역을 채워 DCT 하는 방법으로서, 다음과 같은 단계를 거친다:

- (1) VOP 내의 정보만을 이용하여 평균값을 구한다.
- (2) VOP 밖의 영역을 계산된 평균값으로 채운다.
- (3) 기존의 블럭 DCT 회로를 그대로 이용하여 DCT를 수행한다.
- (4) 양자화, 역양자화, 역DCT 과정후, VOP 모양 정보를 이용하여 VOP 내부에서만 신호 정보를 재현한다.

이동 보상 예측오차를 부호화할 때에는, 평균값 대체 기법의 변형으로, 예측오차의 평균값이 0 (zero)에 가까운 확률이 높은 특성을 이용해 평균값 계산 과정이 없는 0(zero) 값 패딩 기법(zero

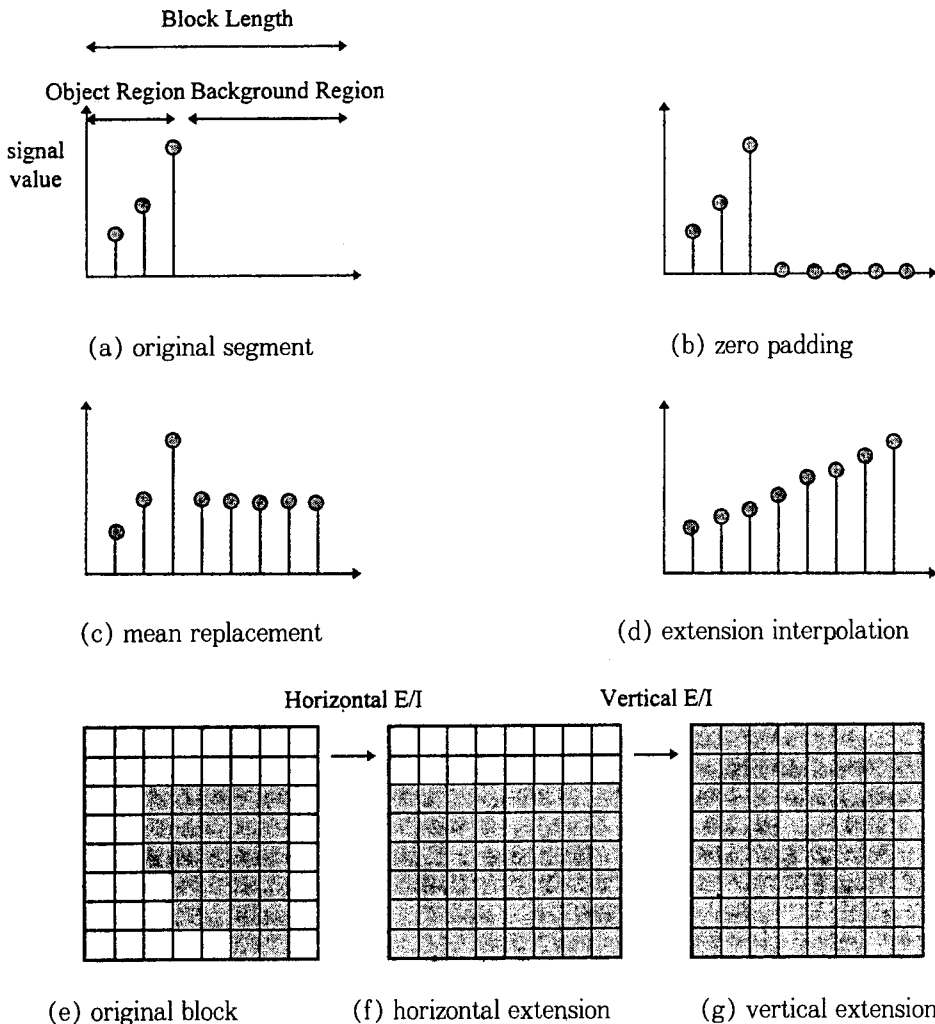


〈그림 11〉 평균값 대체 기법

padding technique)도 아울러 제안되고 있다. 본 방법은 알고리즘이 간단하고 기존의 DCT 회로를 그대로 사용할 수 있다는 장점으로 인해 MPEG-4에서 SADCT와 함께 유력하게 거론되고 있는 실정이다. 하지만, VOP 내의 경계부분에서 고주파 성분이 발생하므로써, 부호화 효율이 다소 떨어지는 단점이 있다.

4. 확장/보간 기법(Extention Interpolation Technique)^[11]

SADCT를 제외하면 대부분 VOP 밖을 새로운 정보로 채워 넣는 과정을 거치는데, 그림 12의 (e)와 같은 VOP 내의 1차원 공간 주파수가 (a)와 같다고하면 0값 패딩 기법은 (b)와 같이, 평균값 대체 기법은 (c)와 같이 나타나고, 따라서 VOP 경계에서 VOP 내에는 존재하지 않았던 고주파의 공간주파수 성분이 발생하게 된다. 이를 극복하기 위해 VOP 내의 공간주파수 성분만을 이용하여 (d)와 같이 확장/보간하는 방법이 제안되었으며, 다음의 과정을 거친다:



<그림 12> 확장/보간 기법

- (1) (e)와 같은 영상인 경우 먼저 영상내의 정보를 수평방향으로 확장하여 보간한다.
- (2) (f)와 같이 수평방향으로 확장/보간된 영상을 수직방향으로 확장하여 보간한다.
- (3) (g)와 같이 확장/보간법에 의해 모두 채워진 블록을 2차원 DCT 한다.
- (4) 양자화, 역양자화, 역DCT 이후의 정보를 이용하여 확장/보간의 역과정을 거쳐 VOP 내부 영역만을 재현해 낸다.

V. MPEG-4 표준화 동향

MPEG-4는 현재, 영상/음향 부호화를 위해 제안된 알고리즘과 툴들의 1995년 11월 시험 결과를 토대로 VM 버전 2.1이 나와있는 상태이며, MPEG-4의 시스템 역할을 담당하고 각종 기능들을 지원하게 될 MSDL(MPEG-4 System and De-

scription Language)의 WD(Working Draft) 버전 1.1이 만들어졌다. 한편, 자연영상뿐 아니라 합성영상(synthetic image) 부호화(SNHC: Synthetic and Natural Hybrid Coding)를 위한 방식들이 제안 요청된 상태이다. 앞으로의 MPEG-4 작업 일정은 표 1과 같다. 1996년 8월에 SNHC에 대한 제안을 받아 9월에 시험하게 된다. VM을 기반으로 하는 비디오 및 오디오 부호화 기법과 MSDL을 계속 향상시켜 나가고, 96년 9월의 SNHC시험 결과를 토대로 합성영상의 부호화 구조를 결정하여 1996년 11월에 MPEG-4 WD가 나오게 된다. WD는 계속 발전되어 1997년 11월 CD(Committee Draft), 1998년 11월에 IS(International Standard)로 확정될 예정이다. 한편, 1996년 11월에는 지금의 VM과는 다른 새로운 알고리즘과 툴들을 제안 요청하게 되며, 1997년 7월에 제안된 새로운 기술들을 이용하여 VM 기반 기술의 성능을 최종 시험하게 될 것이다.

〈표 1〉 MPEG-4 표준화 작업 일정

07/96	MSDL WD 1.2 비디오 VM v. 3.0 오디오 VM v. 2.0 SNHC 실험 방법 설계
08/96	SNHC 제안 완료
09/96	MSDL WD 1.3 비디오 VM v. 4.0 오디오 VM v. 3.0 SNHC 제안 기법 평가
11/96	7/97 시험을 위한 새로운 알고리즘 및 툴 제안 요청 MPEG-4 WD-1 : Part 1-MSDL Part 2-자연 및 합성 (구조) 비디오 툴 및 알고리즘 Part 3-자연 및 합성 (구조) 오디오 툴 및 알고리즘
03/97	MPEG-4 WD-2
07/97	MPEG-4 WD-3 새로운 제안 기술과 VM의 공식 시험
11/97	MPEG-4 CD
11/98	MPEG-4 IS

VI. 결 론

MPEG-4는 물체를 의미하는 VOP를 기반으로 영상정보를 부호화하며 수신단에서는 사용자가 임의로 VOP를 조작하여 출력될 화면을 합성할 수도 있는 영상 부호화 표준을 제공할 것이다. 물체를 기반으로 영상정보를 압축하기 위하여 물체의 모양정보 부호화가 필요하며, 물체의 경계에서 형성될 블락들에 대해서는 특별한 신호정보 부호화 기술이 요구되고 있다. 본 논문에서는 현재 이를 목적으로 제안되어 있는 기술들에 대하여 간단히 살펴보았다. 그러나 이외에도 물체를 인식해 내는 기법이 필수적으로 요구되고 있으며, 임의의 모양을 가진 물체의 움직임 정보를 추정하고 보상하는 등의 특별한 기술들이 필요하다. 현재 이러한 부분들에 대한 각종 틀들이 제안되어 실험되고 있으며, 많은 새로운 영상 부호화 관련 연구 분야를 제공하고 있는 상태이다.

부록 A. VOP 합성 방법

MPEG-4 수신기에 의해 재현된 VOP들은 출력될 화면을 구성하기 위하여 VOP 합성과정을 거치게 된다. 이때, 매번 2개의 VOP들끼리 반복적으로 합성되어진다. 예를 들어, VOP N이 VOP M 위에 합쳐져 VOP P를 새로 합성해 내게 된다면, 합성된 VOP P의 Y,U,V(Pyuv) 및 알파값(Palpha)들은 다음과 같이 계산된다 :

$$Pyuv = ((255 - \text{Nalpha}) * Myuv + (\text{Nalpha} * Nyuv)) / 255$$

$$Palpha = 255$$

만일, VOP P에 합쳐질 또다른 VOP가 있다면, 위의 과정을 다시 수행하면 된다.

참 고 문 헌

[1] "MPEG-4 Project Description," ISO/IEC

JTC1/SC29/WG11 DOC. NO. 96/N1177, Jan. 1996.

[2] H.G. Musmann, M. Hoetter, and J. Ostermann, "Object-Oriented Analysis-Synthesis Coding of Moving Images," Signal Processing: Image Communication, Vol. 1, No. 2, pp. 117-138, Oct. 1989.

[3] K. Aizawa, H. Harashima, and T. Saito, "Model-Based Synthetic Image Coding System," Picture Coding Symposium '87, Stockholm Sweden, 1987.

[4] "Report on the Ad Hoc Group on the Evaluation of Tools for Non-tested Functionalities of Video Submissions," ISO/IEC JTC1/SC29/WG11 DOC. NO. 95/N1064, Nov. 1995.

[5] "Report on the Ad Hoc Group on the Evaluation of Tools and Algorithms for MPEG-4 in January 1996," ISO/IEC JTC1/SC29/WG11 DOC. NO. 96/0677, Jan. 1996.

[6] "MPEG-4 Video Verification Model Version 2.1," ISO/IEC JTC1/SC29/WG11 DOC. NO. 96/XXXX, May 1996.

[7] M. Etoh, H. Arakawa, S. Kadono, T.K. Tan, and C.S. Boon, "Technical Description of MPEG-4 Proposal," ISO/IEC JTC1/SC29/WG11 DOC. NO. 95/0393, Nov. 1995.

[8] "Core Experiments on MPEG-4 Video Shape Coding," ISO/IEC JTC1/SC29/WG11 DOC. NO. 96/0889, May 1996.

[9] T. Sikora and B. Makai, "Shape-Adaptive DCT for Generic Coding of Video," IEEE Trans. on Circ. and Syst. for Video Technol., Vol. 4, No. 1, pp. 59-62, Feb. 1995.

[10] Toshiba, "Technical Description for MPEG-4 First Round of Test," ISO/IEC JTC1/SC29/WG11 DOC. NO. 95/0354,

Nov. 1995.
 [11] C.S. Park, et al, "Daewoo Proposal for Re-
 gion Texture Coding," ISO/IEC JTC1/

SC29/WG11 DOC. NO. 95/0555, Jan.
 1996.

저자 소개



文 柱 禧

1962年 10月 9日生

1985年 2月 서강대학교 전자공학과 졸업(공학사)

1987年 2月 한국과학기술원 전기및전자공학과 졸업(공학석사)

1992年 8月 한국과학기술원 전기및전자공학과 졸업(공학박사)

1992年 9月~1994年 2月 한국과학기술원 정보전자연구소 선임연구원(Post-Doc.)

1994年 3月~현재 현대전자산업주식회사 책임연구원

주관심분야: 영상처리, 영상부호화, 정보이론