

# 임의 형태 영상 영역 부호화를 위한 변환 계수의 삭제 및 복원

김 지 홍

부산외국어대학교 이공대학 컴퓨터공학과

## Truncation and Recovery of Transform Coefficients for Coding Arbitrarily-Shaped Image Segments

Kim, Ji-Hong

*Dept. of Computer Engineering, Pusan University of Foreign Studies*

### 요 약

본 논문에서는 임의 형태 영상 영역에 대한 새로운 변환 부호화 방법을 제안한다. 제안된 방법에서는 영상 영역을 구성하는 각 열의 화소들을 변환 블록의 상단으로 이동시킨 후, 비어 있는 부분을 각 열의 화소 평균값으로 외삽한다. 그리고 수직 방향 1차원 변환을 실행한 후 외삽된 부분의 변환 계수들을 삭제한다. 수평 방향 1차원 변환은 각 행의 변환 계수들을 변환 블록의 좌측단으로 이동시킨 후, 수직 방향 1차원 변환에서와 동일한 과정을 실행함으로써 이루어진다. 복호화를 위해서는 먼저 삭제된 변환 계수들을 복원한 후에 수평 방향 1차원 역변환 과정을 수행하며, 역변환 계수들 중 복원된 부분의 계수들을 삭제한다. 그리고 수평 방향의 1차원 역변환 과정은 수직 방향 1차원 역변환 과정과 동일한 방식으로 수행된다. 역변환으로 만들어진 재생 영상은 변환 계수와 함께 전송된 형태 정보를 이용하여 원 위치로 이동된다.

모의 실험 결과는 제안된 방법의 압축 성능이 낮은 비트율에서 특히 우수함을 보여 준다.

### 1. 서 론

최근 멀티미디어의 적용과 매우 낮은 비트율에서의 영상 전송을 위해 임의 형태의 영상 영역(image segment)에 대한 변환 부호화 방식이 많은 주목을 받고 있다. 임의 형태 영상 영역 부호화 방식은 입력 영상에 대한 영상 분할 과정을 통해 추출된 임의 형태의 물체(object)에 적용되며, 주요 연구 결과는 다음과 같다. 첫번째 종류는 형태 적응 DCT(shape-adaptive DCT) 방식으로서, 변환 블록에 포함되는 임의 형태의 영상 영역에 대해서만

DCT를 실행하는 기법이다[1][2]. 이 방식에서는 2차원 DCT가 분리 가능(separable)하다는 특징을 이용하며, 영상 영역의 각 행 또는 열의 화소 수를 감안하여 다양한 크기의 1차원 DCT를 사용한다. 먼저 영상 영역 각 열의 화소를 변환 블록의 상단으로 이동시킨 후 각 열의 화소 수와 동일한 크기의 수직 방향 1차원 DCT를 적용하고, 이어서 수직 방향 1차원 DCT 계수들을 변환 블록의 좌측단으로 이동시킨 후 각 행의 화소 수와 동일한 크기의 수평 방향 1차원 DCT를 적용한다. 두번째 종류는 내삽(interpolation) 또는 외삽(extra-

polation)을 사용하여 영상 영역을 변환 블록 크기로 확장한 후 블록 DCT를 실시하는 방식이다. 예를 들면 MPEG-4 비디오 VM의 저역 통과 외삽법 (low pass extrapolation padding technique)에서는 Intra 블록의 명시도(luminance) 성분에 대해 변환 블록 내의 비어 있는 부분을 영상 영역의 평균값을 이용하여 외삽한 후 평균 필터를 적용하고, 이어서 변환 블록 크기의 DCT를 적용한다[3]. 이러한 두 종류의 방법은 각각의 장단점을 갖는다. 먼저 형태 적용 DCT 방법은 양자화를 고려하지 않을 경우 전송해야 할 변환 계수의 수가 영상 영역의 화소 수와 동일하다는 장점이 있는 반면, 변환 및 역변환 과정에서 다양한 크기의 변환이 요구되는 단점이 있다. 그리고 저역 통과 외삽법은 평균값의 외삽으로 인해 변환 계수 중 높은 주파수 성분 값이 억제되고 또한 고정된 크기의 변환을 사용할 수 있는 반면, 전송해야 할 변환 계수의 수가 영상 영역의 화소 수보다 증가하여 전송 비트율이 증가할 수도 있는 단점이 있다.

본 논문에서는 외삽을 적용하여 고정된 크기의 DCT를 수행하되, 변환 계수는 영상 영역의 화소 수만큼만 전송하는 새로운 방법을 제안한다. 제안된 방법은 복호화기로 전송된 변환 계수로부터 부호화기에서 외삽된 화소값을 계산할 수 있고, 따라서 전송되지 않고 삭제된 변환 계수들의 복원이 가능하다는 점을 이용한 것이다.

본 논문의 2장에서는 제안된 임의의 형태 영상 영역 부호화 기법에 대해 설명하며, 3장에서는 모의 실험 결과를 기술한다. 그리고 4장의 결론으로 논문을 맺는다.

## 2. 변환 계수의 삭제 및 복원을 이용한 DCT 부호화

### 2.1 순방향 DCT 과정

수직 방향 1차원 DCT 과정에서는 먼저 영상 영역을 구성하는 각 열의 화소들을  $N \times N$  크기 변환 블록의 상단으로 이동시킨 후, 비어있는 위치에 해당 열의 화소 평균값을 외삽한다.  $8 \times 8$  크기의 변환 블록에 대한 예를 그림 1에 나타내었다. 그림 1(a)는 검정색으로 표시된 영상 영역의 초기 형태를 나타내며, 그림 1(b)는 영상 영역의 각 열이 변환 블록의 상단으로 이동된 형태를 나

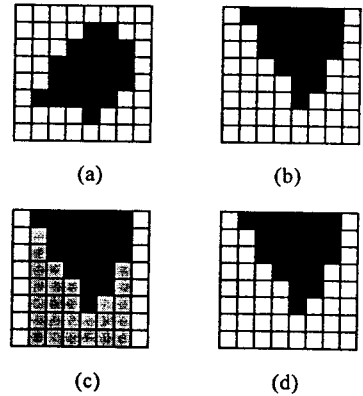


그림 1. 수직 방향 1차원 DCT 과정 (a) 영상 영역의 초기 형태 (b) 변환 블록의 상단으로 이동된 형태 (c) 평균값을 외삽한 형태 (d) 변환 후 부분 삭제된 형태

타낸다. 그리고 그림 1(c)는 회색으로 표시된 화소 평균값이 외삽된 형태를 나타낸다. 영상 영역이 변환 블록의 상단으로 이동되었을 때의 한 열을 벡터  $f$ 로 표기하고, 이 열의 영상 영역 화소 수를  $n$ 으로 가정할 때,

$$f = [f(0) \cdots f(n-1) \quad f(n) \cdots f(N-1)]^T$$

$$= [f(0) \cdots f(n-1) \quad 0 \cdots 0]^T \quad (1)$$

으로 나타낼 수 있다. 여기서  $f(x)$ 는 열의  $x$ 번째 화소를 의미한다. 따라서 외삽될 화소 평균값  $m$ 은

$$m = \frac{1}{n} \sum_{x=0}^{n-1} f(x) \quad (2)$$

이며, 화소 평균값이 외삽되면 벡터  $f$ 는

$$f = [f(0) \cdots f(n-1) \quad f(n) \cdots f(N-1)]^T$$

$$= [f(0) \cdots f(n-1) \quad m \cdots m]^T \quad (3)$$

이다. 그리고 이에 대한 1차원 DCT 변환 계수  $F$ 는

$$F = Gf \quad (4)$$

로 표기된다. 여기서  $G$ 는  $N \times N$  크기의 1차원 DCT 행렬을 나타내며, DCT-II의 경우

$$G(u, x) = \left(\frac{2}{N}\right)^{1/2} c(u) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \quad (5)$$

로서, 가중치  $c(u)$ 는  $u=0$ 이면  $1/\sqrt{2}$  이고, 그 외의 경우 1의 값을 갖는다[4]. 이상의 과정으로 각 열에 대한 1차원 DCT 변환이 완료되면, 생성된

변환 계수들 중 화소 평균값이 외삽된 부분에 위치하는 변환 계수들을 영의 값으로 삭제한다. 즉

$$F = [F(0) \dots F(n-1) F(n) \dots F(N-1)]^T \quad (6)$$

$$= [F(0) \dots F(n-1) 0 \dots 0]^T$$

이며, 그 예는 그림 1 (d)와 같다.

수평 방향 1 차원 DCT 과정은 수직 방향 1 차원 DCT 과정과 동일한 방식으로 수행된다. 그림 2 (b)와 같이 수직 방향 1 차원 DCT 계수들을 변환 블록의 좌측단으로 이동시킨 후, 비어 있는 부분에 대해 각 행의 평균값을 외삽한다. 그림 2 (c)에 평균값이 외삽된 변환 블록의 형태를 도시하였다. 그리고 수평 방향 1 차원 DCT를 수행한 후, 생성된 변환 계수들 중 평균값이 외삽된 부분에 위치하는 변환 계수들을 그림 2 (d)와 같이 영의 값으로 삭제한다.

수직 및 수평 방향 1 차원 DCT 변환에 의한 2 차원 DCT 과정이 완료되면, 삭제되지 않은 변환 계수들 즉 그림 2 (d)의 영상 영역에 위치하는 변환 계수들을 지그재그(zigzag) 스캐닝하여 영상 영역의 초기 형태 정보와 함께 전송한다.

이상에서 기술한 임의의 형태 영상 영역에 대한 DCT 부호화 방법은 변환 블록에서 비어 있는 부분을 영상 영역의 열 또는 행의 평균값으로 외삽한 후 DCT 부호화를 수행하고, 이어서 생성된 변환 계수들 중 외삽된 부분에 위치하는 계수들을 삭제하는 방식을 취한다. 따라서 평균값의 외삽으로 인해 블록 크기의 DCT 변환을 할 수 있을 뿐 아니라, 변환시 압축 효율이 향상될 수 있다. 그리고 변환 계수의 삭제로 인하여 변환 계수에 대한 양자화를 고려하지 않을 경우, 전송될 변환 계수의 수는 변환 블록 내 영상 영역의 화소 수와 동일하게 된다. 이러한 점은 영상 영역의 화소 수가 매우 적더라도 영이 아닌 변환 계수가 최대

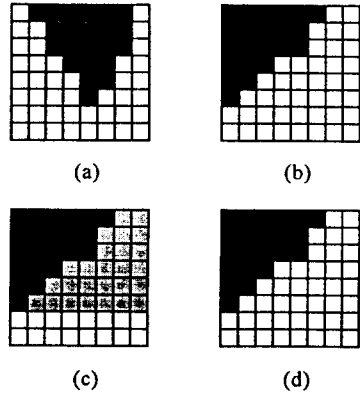


그림 2. 수평 방향 1 차원 DCT 과정 (a) 입력 블록 (b) 변환 블록의 좌측단으로 이동된 형태 (c) 평균값을 외삽한 형태 (d) 변환 후 부분 삭제된 형태

$N \times N$  개까지 발생할 수 있다는 점을 감안할 때 매우 유리한 장점이 될 수 있다.

### 2.2 역방향 DCT 과정

역방향 DCT 과정에서는 수평 및 수직 방향의 순방향 DCT 과정에서 삭제된 변환 계수들에 대한 복원 과정이 선행된다. 수직 방향 1 차원 역방향 DCT 과정에 앞서 수직 방향 1 차원 DCT 과정에서 영으로 삭제된 변환 계수들  $F(u), u = n, n+1, \dots, N-1$ 의 복원 과정은 다음과 같다.

먼저 1 차원 역방향 DCT 과정은

$$f = G^{-1}F \quad (7)$$

$$= HF$$

로 나타낼 수 있으며, 이를 행렬 형태로 표기하면 다음과 같다.

$$\begin{bmatrix} f(0) \\ \vdots \\ f(n-1) \\ f(n) \\ \vdots \\ f(N-1) \end{bmatrix} = \begin{bmatrix} H(0,0) & \dots & H(0,n-1) & H(0,n) & \dots & H(0,N-1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ H(n-1,0) & \dots & H(n-1,n-1) & H(n-1,n) & \dots & H(n-1,N-1) \\ H(n,0) & \dots & H(n,n-1) & H(n,n) & \dots & H(n,N-1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ H(N-1,0) & \dots & H(N-1,n-1) & H(N-1,n) & \dots & H(N-1,N-1) \end{bmatrix} \begin{bmatrix} F(0) \\ \vdots \\ F(n-1) \\ F(n) \\ \vdots \\ F(N-1) \end{bmatrix} \quad (8)$$

식 (8)의 행렬식을 점선으로 구분된 부분으로 분해하면

$$\begin{bmatrix} \mathbf{f}_{\text{org}} \\ \mathbf{f}_{\text{pad}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\text{LT}} & \mathbf{H}_{\text{RT}} \\ \mathbf{H}_{\text{LB}} & \mathbf{H}_{\text{RB}} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{\text{rec}} \\ \mathbf{F}_{\text{del}} \end{bmatrix} \quad (9)$$

와 같이 표기할 수 있다. 여기서  $\mathbf{f}_{\text{org}}$ 와  $\mathbf{f}_{\text{pad}}$ 는 순방향 DCT 과정에서 변환 블록의 상단으로 이동된 한 열의 화소값들과 그 열의 화소 평균값을 각각의 원소로 갖는 부분 벡터들이다. 그리고  $\mathbf{F}_{\text{rec}}$ 와  $\mathbf{F}_{\text{del}}$ 은 복호화에 수신된 변환 계수들과 부호화기에서 삭제된 변환 계수값을 원소로 갖는 부분 벡터를 각각 의미한다.  $\mathbf{H}_{\text{LT}}$ ,  $\mathbf{H}_{\text{RT}}$ ,  $\mathbf{H}_{\text{LB}}$  및  $\mathbf{H}_{\text{RB}}$ 는 변환 행렬  $\mathbf{H}$ 의 부분 행렬들이다.

따라서 식 (9)의 역방향 변환을 수행하기 위해서는 부호화기에서 삭제된 변환 계수들로 구성되어 있는 부분 벡터  $\mathbf{F}_{\text{del}}$ 의 복원이 필수적이다. 먼저 식 (9)로부터

$$\mathbf{f}_{\text{pad}} = \mathbf{H}_{\text{LB}} \cdot \mathbf{F}_{\text{rec}} + \mathbf{H}_{\text{RB}} \cdot \mathbf{F}_{\text{del}} \quad (10)$$

의 관계가 성립하므로, 부호화 과정에서 삭제된 변환 계수  $\mathbf{F}_{\text{del}}$ 은

$$\mathbf{F}_{\text{del}} = \mathbf{H}_{\text{RB}}^{-1} [\mathbf{f}_{\text{pad}} - \mathbf{H}_{\text{LB}} \cdot \mathbf{F}_{\text{rec}}] \quad (11)$$

과 같다. 식 (11)에서  $\mathbf{H}_{\text{RB}}^{-1}$ ,  $\mathbf{H}_{\text{LB}}$ 와  $\mathbf{F}_{\text{rec}}$ 는 알려진 값이므로,  $\mathbf{F}_{\text{del}}$ 을 구하기 위해서는  $\mathbf{f}_{\text{pad}}$ 의 계산이 요구되며 그 과정은 다음과 같다.

식 (3)으로부터  $\mathbf{f}_{\text{pad}}$ 를 구성하는 각 원소의 값은 열의 화소 평균값인  $m$ 이므로, 이는  $\mathbf{F}_{\text{rec}}$ 에 포함된 DC 성분 값인  $F(0)$ 으로부터 식 (4)와 (5)를 이용하여 계산될 수 있다. 즉

이므로,

$$\begin{aligned} \mathbf{f}_{\text{pad}} &= [f(n) \ f(n+1) \ \dots \ f(N-2) \ f(N-1)]^T \\ &= [m \ m \ \dots \ m \ m]^T \\ &= \frac{1}{\sqrt{N}} [F(0) \ F(0) \ \dots \ F(0) \ F(0)]^T \end{aligned} \quad (13)$$

이다.

이상의 과정을 통해 복원된 변환 계수  $\mathbf{F}_{\text{del}}$ 과 수신된 변환 계수  $\mathbf{F}_{\text{rec}}$ 는 조합되어 식 (7)에 의해 역방향 DCT 변환을 수행하게 된다. 그리고 역방향 DCT 변환에 의해 생성된 열 벡터  $\mathbf{f}$ 에서 부분 벡터  $\mathbf{f}_{\text{pad}}$ 의 원소들은 영으로 삭제된다. 그 이유는  $\mathbf{f}_{\text{pad}}$ 의 원소들은 부호화기에서 인위적으로 의삽되었기 때문이다. 이러한 과정은 그림 1에 도시된 수직 방향 1차원 DCT 과정의 역순과 동일하며, 수평 방향 1차원 역방향 DCT 과정도 이와 동일한 방식을 취한다.

수평 및 수직 방향 1차원 역방향 DCT 변환이 완료되면, 수신된 초기 형태 정보를 이용하여 역변환 계수들을 원위치로 이동한다.

### 3. 모의 실험

제안된 임의의 형태 영상 영역 부호화 기법을 그림 3에 나타난 시험 영상들에 대해 모의 실험하였다. "Girl"은 256×256 크기의 정지 영상이며, "Mother and Daughter"는 QCIF 크기로서 비디오 열

$$\begin{aligned} F(0) &= \left(\frac{2}{N}\right)^{1/2} \left(\frac{1}{\sqrt{2}}\right) [1 \ 1 \ \dots \ 1] [f(0) \ f(1) \ \dots \ f(N-2) \ f(N-1)]^T \\ &= \left(\frac{1}{N}\right)^{1/2} \{f(0) + \dots + f(n-1) + f(n) + \dots + f(N-1)\} \\ &= \left(\frac{1}{N}\right)^{1/2} \{f(0) + \dots + f(n-1) + m + \dots + m\} \\ &= \left(\frac{1}{N}\right)^{1/2} mN \\ &= m\sqrt{N} \end{aligned} \quad (12)$$

(sequence) 중의 한 프레임이다. 모의 실험에서는 제안된 기법의 효율성을 검증하기 위해 가장 기본적인 부호화 구조를 선택하였다. 즉 입력 영상을 영상 분할한 후 물체부에 대해 DCT 변환 및 양자화 과정을 실행하고, 양자화된 변환 계수에 대해 허프만 부호화를 하였다. 영상 분할된 시험 영상은 그림 4와 같다. DCT 변환은  $8 \times 8$  크기의 DCT-II 를 사용하였으며, 양자화를 위해 MPEG-4 비디오 VM 에서 권고된 양자화 행렬에 대해 다양한 양자화 파라미터를 적용하였다[3]. 그리고 허프만 부호화 과정에서는 MPEG-4 비디오 VM 의 Intra 블록을 위한 허프만 표를 적용하였으며, 실험 결과의 정량적 성능 평가 척도로서 PSNR 을 사용하였다.

모의 실험에서는 제안된 방법의 성능을 MPEG-4 비디오 VM 에 권고된 저역 통과 외삽법과 비교하였다. 저역 통과 외삽법에 적용된 양자화 및 허프만 부호화는 제안된 방법에서와 동일하다. 그림 5는 영상 분할된 시험 영상에서 물체부의 가장자리에 걸쳐 있는 경계 블록들에 대해, 임의의 형태 영상 영역 부호화 기법이 주로 적용되는 낮은 비트율에서의 모의 실험 결과를 나타낸다. 그림 5의 모의 실험 결과에 형태 정보를 전송하기 위해 소요되는 비트량은 포함되지 않았다. 그림 5에서 볼 때 제안된 방법은 대략 0.3 ~ 1 bpp 의 범위에서 저역 통과 외삽법에 비해 약 1 dB 정도 성능이 향상됨을 알 수 있다. 모의 실험에서는 제안된 방법이 비교적 높은 비트율에서는 저역 통과 외삽법을 포함한 기존 방법들에 비해 성능 향상 정도가 그리 크지 않은 것으로 나타났다. 그 이유는 높은 비트율에서는 양자화 스텝 크기가 작으므로 영으로 양자화되는 높은 주파수의 변환 계수가 적어지고, 그만큼 제안된 방법에서 삭제된 변환 계수의 복원으로 인한 성능 향상 효과가 크지 않기 때문으로 판단된다. 그러나 임의의 형태 영상 영역 부호화 기법은 주로 매우 낮은 비트율의 부호화에 적용되므로, 제안된 방법의 실용적 가치는 기존 방법들에 비해 매우 뛰어남을 알 수 있다.

#### 4. 결 론

본 논문에서 제안한 임의의 형태 영상 영역에 대한 DCT 부호화 기법은 변환 입력 벡터의 비어

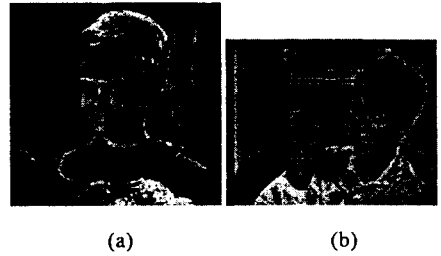


그림 3. 시험 영상들 (a) Girl (256x256) (b) Mother and Daughter (QCIF)

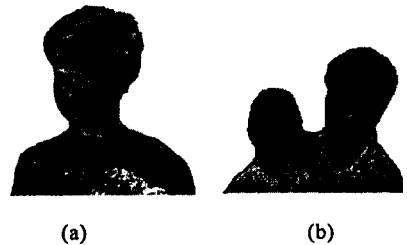


그림 4. 영상 분할된 시험 영상들 (a) Girl (b) Mother and Daughter

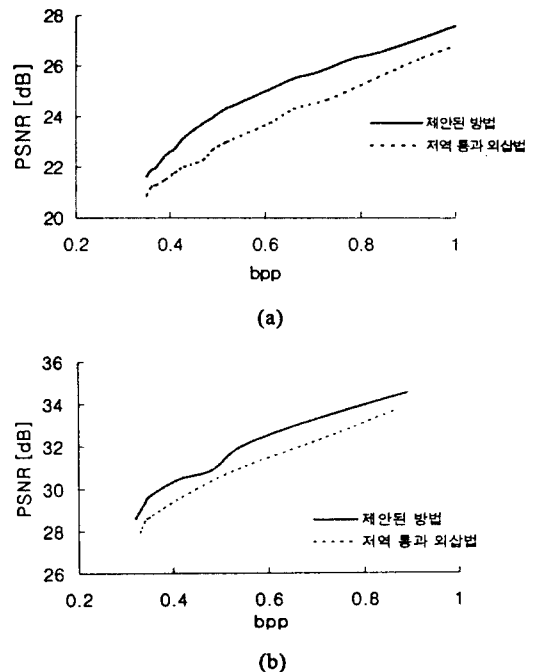


그림 5. 경계 블록에 대한 실험 결과 (a) Girl (b) Mother and Daughter

있는 부분을 화소 평균값으로 외삽한 후 DCT 변환을 수행하고, 이어서 외삽된 부분에 위치하는 변환 계수들을 영으로 삭제하는 방식으로 이루어진다. 그리고 역방향 DCT 과정에서는 삭제된 변환 계수의 복원 과정이 선행되며, 또한 역변환으로 추출된 화소들 중 순방향 DCT 과정에서 외삽된 화소들을 영으로 삭제한다.

제안된 방법은 기존 형태 적응 DCT 방법과 지역 통과 외삽법이 갖는 장점들을 동시에 가졌다고 할 수 있다. 즉 블록 크기로 고정된 DCT 변환을 사용할 수 있으며, 변환 계수의 수는 영상 영역의 화소 수와 동일하므로 전송 비트율에서 큰 이익을 가질 수 있는 장점이 있다. 그리고 DCT 부호화 과정에서 영상 영역을 변환 블록의 상단 및 좌측단으로 이동시키는 과정이 포함되어 있으나, 이 과정은 삭제된 변환 계수의 복원 과정을 단순화하기 위한 목적으로 수행된 것이다. 따라서 본 논문은 화소 평균값으로 외삽하는 방식이 일반적 경향임을 고려할 때, DCT 변환 후 변환 계수들 부분 삭제하고 그리고 역방향 DCT 과정에서 삭제된 변환 계수들을 복원하는 점에 핵심이 있다고 할 수 있다.

모의 실험에서는 MPEG-4 비디오 VM에 권고된 지역 통과 외삽법에 비해 제안된 방법의 부호화 성능이 우수함을 알 수 있었다.

## 참 고 문 헌

- Arbitrarily-Shaped Regions," *SPIE* Vol. 2501, 1995
- [6] T. Sikora, S. Bauer, and B. Makai, "Efficiency of Shape-Adaptive 2-D Transforms for Coding of Arbitrarily Shaped Image Segments," *IEEE Trans. on Circuits and Syst. For Video Tech.*, Vol. 5, No. 3, June 1995
- [7] R. Koenen, F. Pereira, and L. Chiariglione, "MPEG-4: Context and Objectives," *Sig. Proc.: Image Comm.* 9, 1997
- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, 1992
- [1] T. Sikora and B. Makai, "Shape-Adaptive DCT for Generic Coding of Video," *IEEE Trans. on Circuits and Syst. on Video Tech.*, Vol. 5, No. 1, Feb. 1995
- [2] M. Bi, S. H. Ong, and Y. H. Ang, "Comment on Shape-Adaptive DCT for Generic Coding of Video," *IEEE Trans. on Circuits and Syst. For Video Tech.*, Vol. 6, No. 6, Dec. 1996
- [3] ISO/IEC JTC1/SC29/WG11 Ad hoc group on MPEG-4 video VM editing, *MPEG-4 Video Verification Model Version 5.0*, Nov. 1996
- [4] K. R. Rao and P. Yip, *Discrete Cosine Transform : Algorithms, Advantage, Applications*, Academic Press, Inc., 1990
- [5] J. G. Apostolopoulos and J. S. Lim, "Coding of