

論文 2002-39SP-6-10

부호화된 영상의 크기를 변환하는 새로운 고속 트랜스코더

(A New Fast Transcoder for Resizing Coded Video Sequences)

朴一雨*, 趙鎔局*, 明鎭洙*, 蔡秉助*, 吳承竣*,
梁昌模**, 趙威德**

(IL-Woo Park, Yong-Gook Cho, Jin-su Myung, Byung-Jo Chae,
Seoung-Jun Oh, Chang-Mo Yang, and We-Duke Cho)

요약

최근 이동 무선망을 통한 발전된 비디오 응용 서비스에서 비디오 스트림의 크기를 효과적으로 줄이는 트랜스코더가 필요로 하게 되었다. 이러한 트랜스코더는 화소공간뿐만 아니라 DCT 공간상에서도 가능하다. 본 논문에서는 화소공간과 DCT 공간을 활용하여 시각적 화질을 유지하면서 영상 크기를 효과적으로 줄이는 새로운 고속 복합 트랜스코더를 제안하였다. 또한 부호화 모드 결정과 움직임 벡터 재산출 방법을 제안하였다. 대표적인 실험용 비디오 시퀀스에 적용하였을 때 기존의 공간영역 트랜스코딩 방법과 유사한 시각적 화질을 유지하면서도 계산량을 감소시킬 수 있었다.

Abstract

Recently, many advanced video application services over the mobile-wireless networks have required a transcoder which can efficiently reduce the size of compressed video bitstreams. The transcoder can be worked in either the spatial domain or the DCT domain. In this paper, we propose a new fast hybrid-type transcoder which can efficiently reduce the frame size with keeping the visual quality. The proposed scheme consists of two major processes: a transform domain process and a spatial domain process. We also propose a scheme for coding mode selection and motion vector refinement. Experimental results show that our approach can reduce the computational complexity more than any other conventional spatial-domain transcoder with keeping the visual quality.

Keywords : Video Coding, Trans Coder, Resolution Conversion

* 正會員, 光云大學校 電子工學部

(School of Electronics and Telecommunications Kwang
Woon University)

** 正會員, 電子部品研究院

(Korea Electronics Technology Institute)

※ 이 논문은 산업자원부가 지원하는 중기거점사업의
과제와 2002년도 광운대학교 교내 학술연구비 지원
에 의해 수행되었습니다.

接受日字: 2002年7月15日, 수정완료일: 2002年10月22日

I. 서론

향상된 네트워크 기술과 휴대용 통신 기기 등의 발달로 유·무선 정보 통신망을 이용한 동영상 서비스가 활성화되고 있다.^[1] 그러나 동영상 데이터는 방대한 대역폭을 요구하기 때문에 전송 및 저장 매체의 효율적인 사용이라는 측면에서 같은 데이터를 다양한 형태의 동영상 복호화를 위하여 중복적으로 생성하고 각각 저장한다는 것은 비효율적이다. 서비스를 요청하는 클

라이언트(Client)의 입장에서 볼 때, 제공되는 대역폭이나 계산능력 등이 모두 다르기 때문에 서비스의 한계를 가지게 되며 이러한 한계를 극복하기 위하여 여러 계층으로 분리된 부호화 방식이 사용되었다.^[2] 이 부호화 방식은 비디오 영상의 중요도에 따라 여러 계층으로 분리시킨 후 대역폭 및 복호기의 능력에 따라 계층의 일부나 전체를 받아 보도록 하는 방법이다. 그러나 계층으로 분리된 부호화 방식은 이러한 방식으로 부호화하지 않은 동영상 스트림은 서비스가 불가능하므로 최근에는 이와 달리 멀티미디어 서비스를 제공하는 서버와 서비스를 제공받는 클라이언트 사이에 트랜스코더(Transcoder)를 두어 전송 또는 저장 매체의 상태, 수신단의 특성과 화질에 대한 요구조건 등이 변경될 경우 능동적으로 비트율을 제어할 수 있도록 하는 방법이 많이 사용된다.^[3] <그림 1>은 트랜스코더의 물리적 위치를 나타낸 것이다. 이러한 능동적인 비트율 제어에 사용되는 대표적인 방법에는 양자화 폭(Quantization Step Size)을 조절하는 방법,^[4] 프레임율(Frame Rate)을 조절하는 방법,^[5] 그리고 비디오 영상의 크기(Frame Size)를 변환하는 방법^[6] 등이 있다.

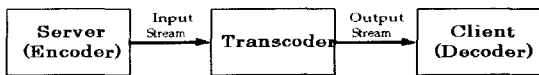


그림 1. 트랜스코더에 기반한 스트리밍 시스템
Fig. 1. A transcoding-based streaming system.

본 논문에서는 이러한 트랜스코딩 방법 중 비디오 크기를 감소시켜 표시 영역이 작은 휴대용 기기와 비교적 낮은 비트율을 요구하는 무선 통신 서비스, 대역폭이 고정되지 않은 네트워크 상황 등에 사용할 수 있는 트랜스코딩 방법을 제안한다. 기존의 비디오 크기를 감소시키는 트랜스코딩 방법은 입력 스트림을 완전히 복호화한 영상의 크기를 변환하여 다시 부호화하는 방법과 입력 스트림을 완전 복호화 후 움직임 벡터(MV: Motion Vector)를 이용하여 움직임 예측(ME: Motion Estimation)에 대한 부담을 줄이는 방법,^[11,12] 모든 과정을 DCT 공간에서 처리를 하는 방법^[7,9] 등으로 분류된다. 화소 공간에서의 처리를 이용하는 방법은 화질이 우수한 반면 많은 연산량을 필요로 하며, 모든 트랜스코딩 과정을 DCT 공간상에서 처리하는 트랜스코더는 연산량에는 이득이 있으나 화질의 열화 현상이 심각하게 발생한다. 본 논문에서는 입력 스트림을 복호

화 하여 얻어지는 부호화 정보를 가지고 DCT 공간과 화소공간에서 크기감소를 수행하여 많은 경우의 ME뿐만 아니라 IDCT, DCT 과정을 생략하여 트랜스코더의 부담을 줄이고자 한다. 또한 Docef와 그의 동료들이 제안한 양자화와 DCT가 결합된 양자화 DCT (QDCT: Quantization DCT)[8]를 사용하여 연산량을 감소시켰다.

II. 비디오 크기 변환 트랜스코더 구조

<그림 2>는 가장 간단한 형태의 트랜스코더로 부호화와 복호화기가 직렬로 연결된 종속 트랜스코더(Cascaded Transcoder)의 구조를 보여준다. 이 모델은 입력 스트림을 완전 복호화한 다음 화소 공간 영역에서 영상의 크기를 줄인 후 다시 부호화하는 과정을 수행한다.

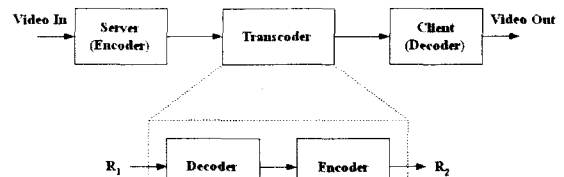


그림 2. 종속 트랜스코더의 구조
Fig. 2. The structure of a cascaded transcoder.

이러한 종속 트랜스코더는 입력된 비디오 스트림의 부호화 정보를 사용하지 않고 복호화한 영상을 가지고 다시 부호화 과정을 거치므로 화질 열화 현상을 줄일 수 있으나 많은 연산 시간을 요구한다.

그러므로 실시간 처리가 필요한 응용에서는 종속 트랜스코더는 적절하지 못하다. 연산 시간을 줄이기 위하여 모든 트랜스코딩 과정을 DCT 공간상에서 처리 방법은 움직임 추정 시에 MV를 정확히 예측하지 못하여 오류가 누적되고, 움직임 보상(MC: Motion Compensation)시에 사용되는 변환 행렬의 곱셈 과정에서 많은 연산량을 필요로 하는 단점을 가지고 있다.^[3] 본 논문에서는 DCT 공간과 정확한 MC가 필요한 부분에서는 화소 공간에서 처리하는 복합형 트랜스코더(Hybrid Transcoder)를 제안한다. <그림 3>은 본 논문에서 제안한 매체변환기의 블록 다이어그램이다. 원 영상에서 MV, 매크로블록(MB: Macro Block) 모드 정보, DCT 계수를 분리한 후 이들을 이용하여 DCT 공간상의 처리와 화소 공간상의 처리를 구분한다. 4개의 MB가 인

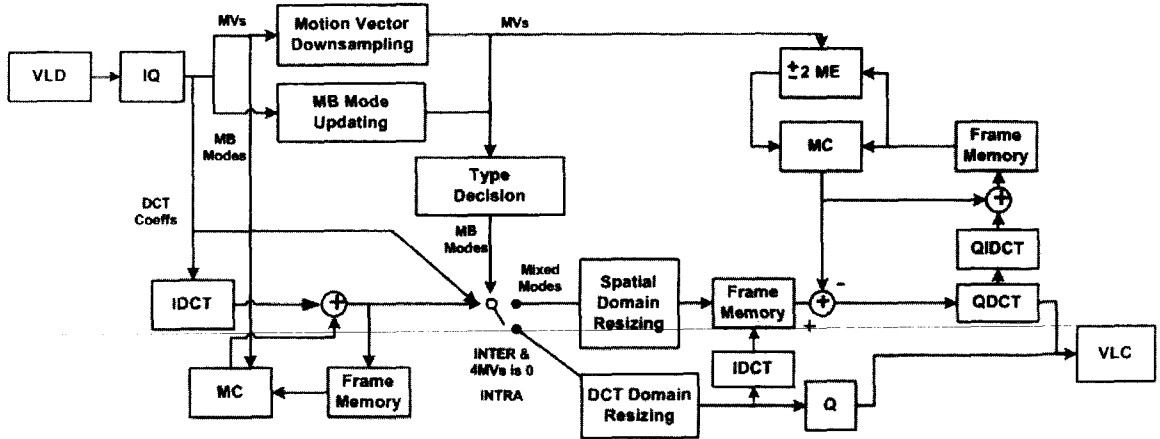


그림 3. 제안된 복합 트랜스코더 블록 다이어그램

Fig. 3. The block diagram of the proposed hybrid transcoder.

터(INTRA)모드이고 4개의 MV가 0 벡터인 경우와 4개의 MB 모두가 인트라(INTRA) 모드로 부호화된 경우는 3장에서 소개할 DCT 공간에서의 크기조절(DDR: DCT Domain Resizing)을 수행한다. 이 경우에는 ME를 수행하지 않아도 정확한 DCT 계수를 얻을 수 있으므로 바로 DDR이 가능하다. DDR 과정이 많을수록 재부호화 시에 가장 많은 계산량을 차지하는 ME과정을 생략할 수 있기 때문에 트랜스코더의 영상 크기 변환 속도가 증가된다. 그러므로 인트라 모드나 MV가 0 벡터인 MB가 많은 영상일수록 유리하다. 즉 화상회의와 같이 움직임이 적어 MV가 0인 경우가 많이 발생하거나 화면의 변화가 급격하여 인트라 MB가 많이 생기는 영상의 경우에는 ME 과정을 생략하고 DDR을 수행할 확률이 높다.

움직임이 없는 배경의 경우 MV가 대부분 영 벡터 또는 부호화 모드가 스킵(SKIP)인 경우가 많이 발생하므로 4개의 MB가 모두 스킵 모드인 경우는 스킵 모드로 판정하여 많은 부호화 과정을 생략한다. 4개의 MB

중 하나라도 인트라 모드로 부호화된 경우에는 인트라로 변환 후에 DDR을 수행한다. 마지막으로 4개의 MB 모드가 인트라 모드이고 네 개의 MV가 모두 0 벡터가 아닌 경우에는 4장에서 설명하는 움직임 벡터 재산출 방법을 사용하여 ME를 수행하고 DCT 한다. 이때에 DCT 과정 후에 따로 양자화를 수행하지 않고 5장에서 제시하는 QDCT를 이용하여 화소 공간 처리의 부담을 줄였다. <표 1>은 이러한 각 모드에 대하여 설명하고 있다.

III. DDR(DCT-Domain Resizing) 과정

입력된 비디오 프레임의 크기를 1/2로 줄여서 스트림의 크기를 줄이는 방법은 화소 공간에서는 여러 가지 방법으로 이루어 질 수 있으며 대부분 간단한 연산으로 가능하다.^[9] 이러한 비디오 프레임의 크기 감소는 DCT 공간에서도 마찬가지로 가능하고, 실시간 처리를 요하는 응용에 더 효율적이다. 대부분의 비디오 압축 표준에서 사용되는 블록의 크기는 8×8이고, 휘도 성분만을 고려 할 경우에 네 개의 블록이 하나의 MB으로 구성된다. 비디오 프레임 크기 감소는 MB 단위로 이루어진다. 즉 4개의 MB이 하나의 MB으로 재구성된다.

크기를 감소할 DCT 공간의 매크로블록 행렬을 M이라 하고, 감소된 후의 DCT 공간에서의 8×8 블록의 행렬을 B이라 하면 두 행렬은 식 (1)의 관계를 갖는다.

$$B = TMT^T \quad (1)$$

식 (1)에서 행렬 T는 DCT 공간에서 16×16 행렬을 8×8

표 1. MB 부호화 모드 결정

Table 1. MB coding mode decision.

4개의 MB 모드	부호화 방법
모두 INTRA	DDR
모든 MB가 0 벡터일 경우	
모든 MB의 MV가 같은 방향	
하나라도 INTRA 포함	INTRA+DDR
모든 MB가 SKIP	SKIP
모두 INTER이며 서로 다른 MV일 경우	ME(탐색영역:±2)

행렬로 변환시키는 행렬이다. T의 크기는 8×16이다. M 행렬에 대해 원하는 동작을 하나씩 추가시킴으로써 T 변환행렬을 정의한다.

$$B_F = FMF^t \tag{2}$$

식 (2)의 F는 크기감소 후의 앨리어싱(aliasing)을 방지하기 위한 저대역 필터(Low-pass filter) 행렬이다. 즉, 이 행렬의 크기는 16×16이 되고, 식 (3)과 같이 정의된다.

$$F = I_2 \otimes \begin{bmatrix} I_4 & 0 \\ 0 & 0 \end{bmatrix} \tag{3}$$

IN 은 N×N 단위 행렬이고 ⊗는 크로네커 곱(Kronecker product)을 나타낸다. F 는 각각의 블록 좌상단의 계수만 남겨두고 나머지 고주파성분을 모두 0으로 만드는 저대역 필터 역할을 수행하여, 블록화 현상(Block Artifact)을 감소시키는 역할을 함과 동시에 앨리어싱을 방지하여 크기감소 후 화질이 왜곡되는 것을 방지한다.

다음 단계로 저대역 필터가 수행된 행렬 BF를 IDCT 한다. IDCT 된 행렬 BC는 식 (4)와 같이 정리된다.

$$B_C = \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix} B_F \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} \tag{4}$$

C 행렬은 DCT 변환 행렬로서 크기는 8×8이다. C 행렬은 식 (5)로 정의되는 c(i, j) 성분들을 가진다.

$$c(i, j) = \frac{1}{2} k(i) \cos \frac{(2j+1)i\pi}{16} \tag{5}$$

k(i)는 식 (6)과 같다.

$$k(i) = \begin{cases} \frac{1}{\sqrt{2}}, & i=0 \\ 1, & otherwise \end{cases} \tag{6}$$

식 (4)에서 보는 바와 같이 B_c는 화소 공간의 16×16 매크로블록이 되고, 화소 공간의 8×8 블록 B_R은 행렬 R을 적용하여 다음과 같이 얻을 수 있다.

$$B_R = RB_C R^t \tag{7}$$

R 행렬의 크기는 8×16이고 다음과 같이 정의된다.

$$R = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{bmatrix} \tag{8}$$

마지막으로 B_R을 DCT 변환하면 8×8 DCT 공간의 행렬 B가 생성된다.

$$B = CB_R C^t \tag{9}$$

위 식 (1)~(9)을 정리하면 T 변환행렬을 식 (10)과 같이 정의할 수 있다.

$$T = CR \begin{bmatrix} C^t & 0 \\ 0 & C^t \end{bmatrix} F \tag{10}$$

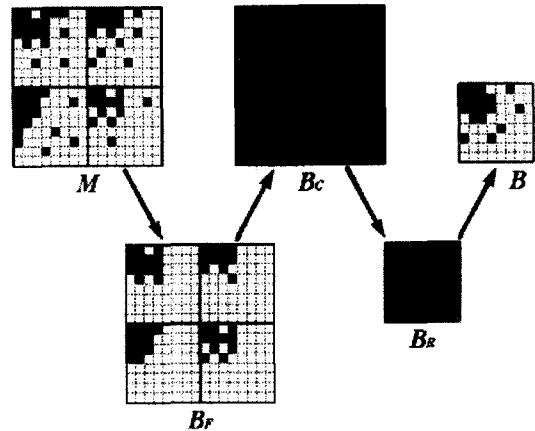


그림 4. DCT 공간에서의 매크로블록의 크기 감소 과정
Fig. 4. The process flow for MB resizing in the DCT-domain.

<그림 4>는 크기 감소 과정을 보여준다. 실선으로 표시된 블록은 화소 공간을 나타내며, 점선으로 표시된 블록은 DCT 공간을 나타낸다. DCT 공간 블록에서의 검은 점은 0이 아닌 DCT 계수값을 표시하고, 흰 점은 0인 DCT 계수값을 표시한다. 그리고 화소 공간 블록에서의 회색 점은 화소값을 뜻한다. 변환행렬 T는 미리 계산하여 완성된 하나의 행렬로 만들어 놓고 식 (1)을 바로 적용할 수 있다.

IV. 움직임 벡터 재사용과 부호화 모드

DCT 공간상에서의 이미지 크기 변환이나 MC을 실제 비디오스트림에 적용시킬 경우 인터 프레임에서 고려해야 할 사항은 다음과 같다. 첫째, 현재 인터 프레임의 DCT 계수는 참조 프레임에 입력된 MV와 예측 오차로 인해 MC가 되는 것이므로 두 정보 중 하나라도 오류가 생기면 오류가 누적되는 문제가 발생한다. 둘째, 인터 프레임에도 인트라 모드와 인터 모드 MB가 존재

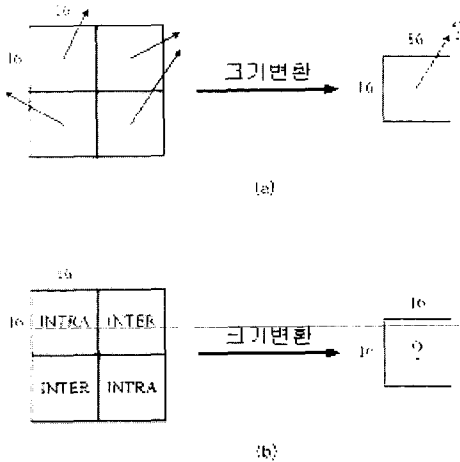


그림 5. 매크로블록의 움직임 벡터 할당 방법
Fig. 5. A scheme for motion vector assignment in frame size reduction.

하기 때문에 이 점도 고려해 주어야 한다.

첫 번째 문제가 발생하는 이유를 <그림 5(a)>에서 보여준다. 네 개의 MB를 하나의 MB로 축소시킬 때 각각의 다른 MV를 가지고 있다면, 크기감소 후에는 대표되는 하나의 MV를 산출하기가 쉽지 않다. 이 문제를 해결하기 위해서 다양한 MV 재산출 방법이 이용된다. 가장 간단한 MV 재산출은 4개의 MV의 평균값이나 각 MV의 중요도를 다르게 하여 새로운 MV를 구하는 방법 등이 있다. 식 (11)은 각 블록의 에너지로 가중치를 주도록 한 AWW (Adaptive Worst Weighting) 방법이다.^[10]

$$\overline{mv} = \frac{1}{2} \frac{\sum_{i=1}^4 \overline{mv}_i A_i}{\sum_{i=1}^4 A_i} \quad (11)$$

여기서 \overline{mv} 는 새로 산출된 MV이며 \overline{mv}_i 는 원 영상에서 각 블록의 MV를 나타내고 A_i 는 각 블록의 에너지 값을 나타낸다. 각 블록의 에너지 값은 간단히 전체 DCT 계수의 절대값의 합을 더한 것을 사용한다. <그림 9>와 <그림 10>은 CIF(352x288)로 부호화된 영상을 중속 트랜스코더를 사용하여 완전 재부호화한 영상의 MV와 원영상의 4개 MV의 평균값과 AWW 방법으로 구한 MV의 거리 오차를 구한 것이다. 이때 재부호화될 CIF 영상은 ME 과정에서 ±16 탐색 영역으로 부호화 하였으므로 앞의 방법 등으로 MV를 예측하면 탐색영역이 반이 되므로 중속 트랜스코더의 탐색영

역은 ±8로 설정하여 비교하였다. <그림 9>와 <그림 10>의 결과에서 알 수 있듯이 MV의 평균값과 AWW를 이용한 방법의 결과가 유사하다. 하지만 AWW는 각 블록의 에너지를 구하는 과정이 들어가므로 평균값보다 많은 계산량이 필요하다. 그러나 이들 예측된 움직임 벡터만을 사용하여 새로운 영상을 부호화할 경우 정확한 최소 예측 에러 값을 찾을 수 없어 출력 비트량 증가와 화질 저하를 일으킨다. 따라서 좀더 정확한 MV의 ME 과정이 필요하다. 본 논문에서는 대표적인 실험 영상들을 실험해 본 결과 MV의 예측 정확도가 2 이하인 경우가 90%이상 차지하고 있으므로 <그림 6>에서와 같이 ±2.5 탐색 영역을 재탐색 하여 좀더 정확한 MV를 구하여 화질의 열화를 막도록 하였다. 그러나 4개의 MB의 MV가 모두 같은 방향을 가리키고 있거나 영 벡터일 경우는 ME 후에 MV를 다시 산출하지 않고 DCT 공간상에서 직접적으로 DDR을 수행한다. 이로 인해 불필요한 ME 과정을 생략하여 계산량의 많은 이득을 얻을 수 있다. 일반적인 영상의 경우 부호화 되지 않는 스킵 모드인 경우도 또한 높은 빈도로 발생한다. 이러한 경우 4개의 MB 모두가 스킵 모드이면 ME를 수행하지 않고 스킵모드로 부호화 한다.

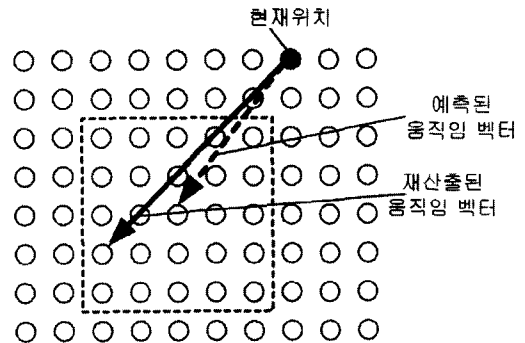


그림 6. MV 재산출
Fig. 6. MV refinement.

두 번째 문제는 <그림 5(b)>의 경우에서 알 수 있다. 각 MB 모드가 인트라와 인터가 섞여 있을 경우 모두 인트라로 부호화 하도록 한다. 이러한 경우 인터 모드는 복호화 하여 화소공간으로 변환 후에 DCT를 수행하여 인트라로 변환한 다음 DDR을 수행한다. 이 영향으로 비트량과 계산량이 증가될 수 있으나 트랜스코딩 때문에 발생하는 오류의 전파를 막아주고 MB 모드 결정을 위한 ME를 생략할 수 있다. <그림 7>은 이러한

부호화 모드의 결정 과정을 간략하게 나타낸 순서도이다.

V. QDCT(Quantization DCT)의 정수 구현

양자화 DCT는 양자화기와 DCT가 결합된 형태를 가진다. 양자화 DCT 구현 시 고려 할 점은 두 가지이다. 첫째는 선형 양자화기를 DCT 과정에 넣는 방법이고 둘째는 정수형 구현 시 발생하는 문제에 대한 해결 방법이다.^[8]

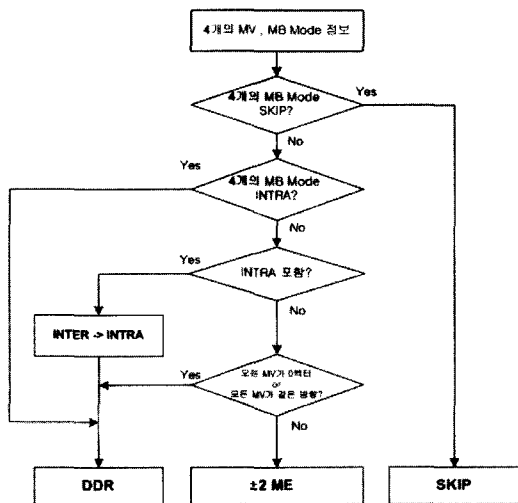


그림 7. 매크로블록 부호화 순서도
Fig. 7. The flow chart for MB mode decision.

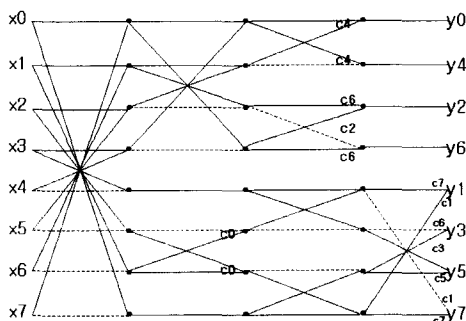


그림 8. Chen 알고리즘의 흐름도
Fig. 8. The flow of Chen's algorithm.

DCT의 정수 구현에 있어 가장 중요한 점은 DCT 각 경로에 단 한번의 곱셈이 있어야 한다는 것이다. 그 이유는 곱셈 연산에 n 비트 이동 연산이 적용되기 때문

이다. 즉, 특정 경로에 두 번의 곱셈이 존재하면 그 경로에 대한 결과는 n 비트가 아닌 2n 비트만큼 이동 연산이 적용이 되어 오버플로우가 발생하게 된다. 위의 조건을 만족시키기 위해서 DCT 알고리즘 중 각 경로에 한번의 곱셈을 수행하는 Chen 알고리즘을 사용하였다. Chen 알고리즘의 흐름도는 <그림 8>과 같다.

본 논문에서는 코사인 계수를 11 비트 이동하여 정수형 계수를 사용하였다. <표 2>는 정수형 계수로 표현된 코사인 계수표이다. <표 3>에서 보여주듯이 11 비트 이동 계수를 사용할 경우 화질의 열화가 일어나지 않음을 알 수 있다. 11 비트 이동된 계수는 다음과 같이 정의된다.

$$C_i' = \lfloor C_i \times 2^{11} + 0.5 \rfloor \quad (12)$$

식 (12)를 이용하여 11 비트만큼 좌측으로 이동하여 정수로 만들어진 코사인 값을 사용해서 1차원 DCT를 행과 열에 대해서 수행한 후, 22 비트 우측 이동 연산으로 2차원 QDCT 과정을 수행하게 된다.

표 2. Chen 알고리즘 사용 시 필요한 DCT 계수

Table 2. DCT coefficients in Chen's algorithm.

QP	c1	c2	c3	c4	c5	c6	c7
1	710	669	602	512	402	277	141
2	502	473	426	362	284	196	100
3	410	386	348	296	232	160	82
4	355	334	301	256	201	139	71
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
28	134	126	114	97	76	52	27
29	132	124	112	95	75	51	26
30	130	122	110	93	73	51	26
31	128	120	108	92	72	50	25

표 3. DCT 모델간의 성능 비교

Table 3. PSNR values in different DCT models.

	평균 PSNR (dB)
DCT + 양자화 모델	33.55
양자화 DCT (11bit)	33.55
양자화 DCT (10bit)	33.52

VI. 실험결과

제안된 복합 트랜스코더를 움직임이 많은 비디오 영상인 Foreman 영상과 움직임이 작은 Akiyo 영상, 순간적인 움직임이 있는 Hall Monitor 영상에 적용시켰다. 각 실험 영상의 화면 크기는 352x288 크기의 CIF이다. 이 크기를 H.263 부호화 표준을 사용하여 256Kbps로 부호화된 비트 스트림과 양자화 파라미터(QP: Quantization Parameter)를 10으로 고정하여 가변 비트율을

갖는 비트 스트림 100 프레임을 176x144 크기의 QCIF로 트랜스코딩 하였다. 트랜스코딩된 비트 스트림은 I와 P 프레임으로만 변환이 되도록 하였다. 실험은 정제 범위에 따른 MV의 예측 정확도와 PSNR의 비교, 트랜스코딩 시 QP 변화에 따른 비트량 변화, 부호화 시 가장 많은 연산량을 차지하는 ME과정에 대한 제안된 방식과 종속 트랜스코더와의 계산량을 비교하고 마지막으로 시각적 화질에 대하여 비교하였다.

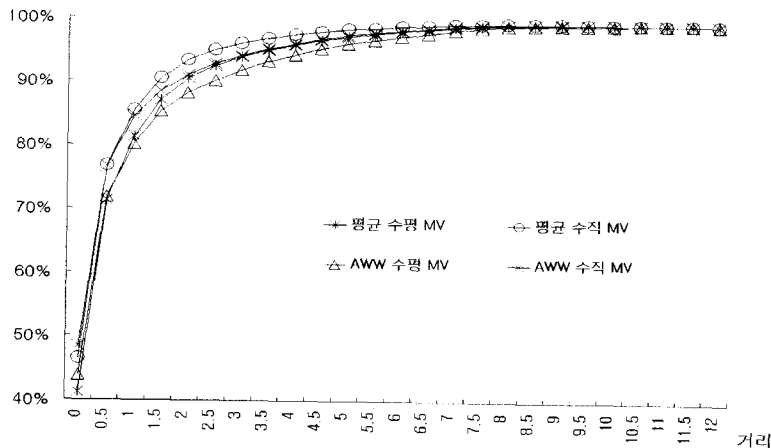


그림 9. Foreman 영상에 대한 4개의 MV의 평균과 AWW로 정제된 MV값의 정확도

Fig. 9. MV refinement accuracies using the average of four MV's and AWW, respectively, in Foreman sequence.

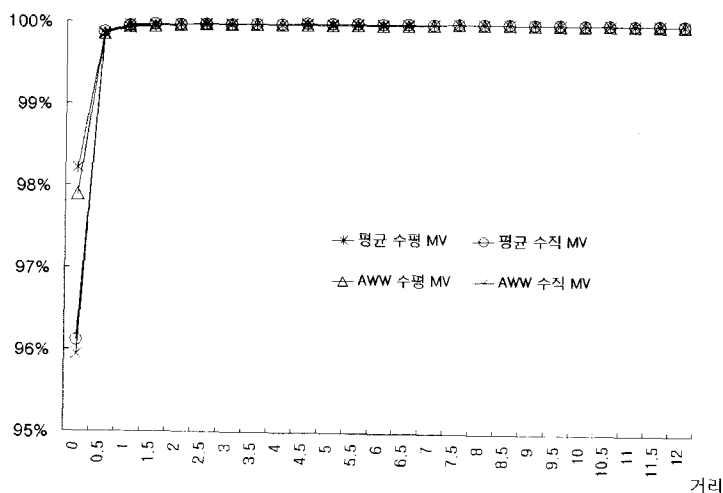


그림 10. Akiyo 영상에 대한 4개의 MV의 평균과 AWW로 예측된 MV의 예측 정확도

Fig. 10. MV refinement accuracies using the average of four MV's and AWW, respectively, in Akiyo sequence.

표 4. 각 트랜스코더를 통하여 크기 조절된 시퀀스의 PSNR 비교

Table 4. PSNR values of size-reduced sequences after applying different transcoding schemes.

실험 영상	256Kbps로 부호화된 CIF 영상 (CIF)	중속 트랜스코더	제안된 Hybrid 트랜스코더 재탐색 영역별 평균 PSNR(dB)				OL-DCT
			±0.5	±1.5	±2.5	±3.5	
Akiyo	40.81	34.48	34.03	34.05	34.05	34.05	32.56
Foreman	32.52	32.61	32.31	32.43	32.46	32.50	23.91
Hall Monitor	36.07	32.84	32.77	32.77	32.78	32.79	28.32

<그림 9> <그림 10>은 예측 MV에 대한 예측 정확도를 나타낸 것이다. AWW와 평균을 이용한 예측 MV를 구하는 방법 모두 Akiyo 영상과 같은 움직임이 적은 영상에서는 예측 MV와의 차이가 ±1.5사이에서 98%의 정확도를 나타내었으며, Foreman 영상과 같은 움직임이 많은 영상에서는 ±2.5사이에서 90%의 정확도를 나타내었다.

<표 4>에서는 각 트랜스코더에 실험영상들을 적용하였을 경우의 평균 PSNR을 나타낸다. 여기서 개방형 루프 DCT(Open Loop DCT : OL-DCT) 트랜스코더는 예측된 MV만을 가지고 원 영상의 부호화된 데이터들 DCT 공간상에서 직접 트랜스코딩 하는 트랜스코더이다.

256Kbps로 부호화된 CIF 입력 영상을 제외한 각 트랜스코더의 PSNR은 CIF로 부호화된 영상을 복호화한 후 QCIF로 줄인 영상과 비교 하였다. OL-DCT의 경우 정확한 MV를 예측하지 못하여 프레임이 증가함에 따라 오류가 누적되어 PSNR이 급격하게 떨어지는 것을 볼 수 있다. <그림 11>은 Foreman 실험 영상에 대한 트랜스코딩 시 QP 변화에 따른 PSNR을 보여 준다. 256Kbps로 부호화된 CIF영상에 QP 10과 15를 사용하여 트랜스코딩 하였다. 결과에서 알 수 있듯이 제안된 복합 트랜스코더와 중속 트랜스코더의 PSNR 변화가 유사한 것을 알 수 있으며 QP 5 변화 시켰을 때에 평균 2dB의 PSNR 감소가 나타났다.

<그림 12>와 <그림 13>은 영상의 크기 감소에 따른 비트 이득을 나타내고 있다. <그림 12>는 Foreman 256Kbps CIF영상을 QP 10과 15로 트랜스코딩 하였을 경우의 비트량을 나타낸다. <그림 13>은 QP 10으로 고정하여 부호화한 영상을 QP 10과 15로 트랜스코딩 하였을 때의 비트량을 나타낸다. <표 5>에서 알 수 있듯이 영상의 크기를 줄였을 경우 2-8배의 비트 이득을 얻을 수 있었다.

표 5. 총 부호화 비트량

Table 5. Total bits of size-reduced video sequence.

실험영상	CIF 부호화 비트량		QCIF 부호화 비트량	
	256Kbps	848K	QP 10	QP 15
Akiyo	256Kbps	848K	104K	64K
	QP 10	296K	96K	56K
Foreman	256Kbps	840K	368K	224K
	QP 10	1104K	384K	232K
Hall	256Kbps	832K	168K	105K
	QP 10	547K	160K	104K

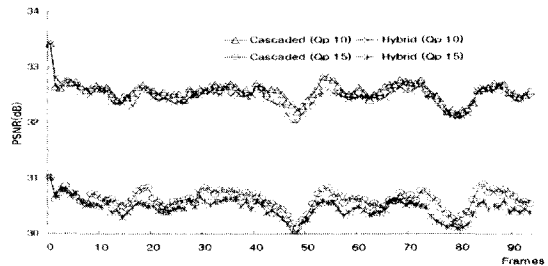


그림 11. 256Kbps CIF Foreman 영상의 크기 변환 시 QP=10과 15에 대한 PSNR 값

Fig. 11. PSNR value of each frame after applying the proposed scheme at QP=10 and 15 in 256 Kbps CIF Foreman sequence.

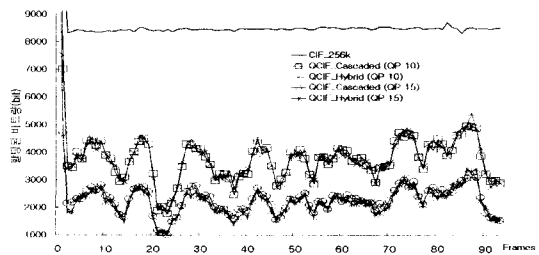


그림 12. 256Kbps CIF Foreman 영상의 크기 변환 시 QP=10과 15에 대한 비트량

Fig. 12. the bit amount of each frame after applying the proposed scheme at QP=10 and 15 in 256 Kbps CIF Foreman sequence.

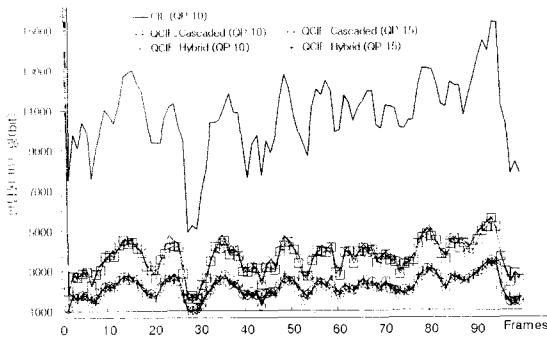


그림 13. QP 10으로 고정하여 부호화한 CIF 영상의 화면 크기를 줄였을 경우의 비트량
 Fig. 13. The bit amount of each frame after transcoding CIF Foreman sequence coded at QP=10.

ME는 부호화 시 가장 많은 연산시간을 차지하고 있으므로 ME의 수행 회수와 계산량은 트랜스코더의 성능에 가장 큰 영향을 준다. 종속 트랜스코더는 완전 부호화 후 완전 부호화를 수행한다. 그러나 제안된 복합 트랜스코더의 경우 DDR, 스킵모드 그리고 인트라 부호화 등을 활용하여 많은 부분의 ME 과정을 생략된다. 또한 ME의 탐색 영역이 ± 2 이므로 종속 트랜스코더에서 ± 8 ME를 수행할 경우 보다 연산량이 1/16로 줄어든다. <표 6>에서 알 수 있듯이 Akiyo 영상과 같이 움직임이 작은 영상은 4개의 MV가 0 벡터인 경우와 스킵 모드가 95%를 차지하고 있다. 이때에는 제안된 트랜스코더에서는 ME 과정을 하지 않으므로 5%의 ME 과정만이 필요하게 된다. 또한 DDR이 수행되는 회수가 많아지면 DCT 공간상에서 직접적으로 처리하여 트랜스코딩 시 DCT 과정을 줄일 수 있다.

표 6. 0 벡터와 스킵된 매크로블록의 비율
 Table 6. Percentage of zero vector and skipped MB's.

	0 벡터와 스킵된 MB의 비율	INTRA 부호화 비율	ME 수행 비율
Akiyo	95.43%	0.00%	4.57%
Foreman	41.35%	8.78%	49.87%
Hall	94.12%	0.08%	5.8%

<그림 14>와 <그림 15>는 100 프레임 중 임의의 프레임에 대한 시각적인 화질을 보여준다. 평균 PSNR은 0.1dB~0.8dB 정도 떨어지나 시각적 화질은 종속 트랜

스코더와 유사함을 알 수 있다. 특히 DDR 과정 중 저대역 필터로 인하여 블록화 현상을 감소시킬 수 있었다.



(b)

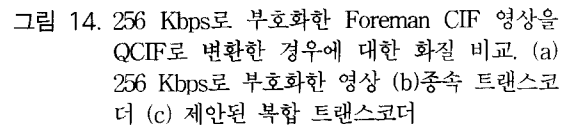
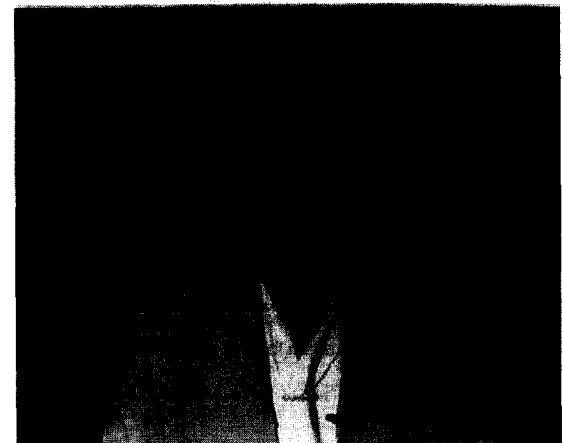


그림 14. 256 Kbps로 부호화한 Foreman CIF 영상을 QCIF로 변환한 경우에 대한 화질 비교. (a) 256 Kbps로 부호화한 영상 (b) 종속 트랜스코더 (c) 제안된 복합 트랜스코더
 Fig. 14. Transcoded Foreman sequences: (a) original (256 Kbps CIF), (b) cascaded transcoder and (c) proposed hybrid transcoder.



(a)

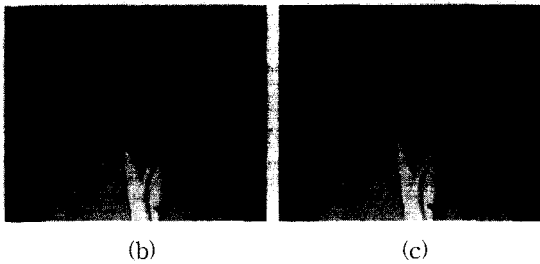


그림 15. 256 Kbps로 부호화한 Akiyo CIF 영상을 QCIF로 변환한 경우에 대한 화질 비교. (a) 256 Kbps로 부호화한 영상 (b)종속 트랜스코더 (c) 제안된 복합 트랜스코더

Fig. 15. Transcoded Akiyo sequences: (a) original (256 Kbps CIF), (b) cascaded transcoder and (c) proposed hybrid transcoder.

V. 결 론

본 논문에서는 가변적인 대역폭을 갖는 네트워크에 비디오 데이터를 효율적으로 전송하기 위한 트랜스코딩 방법 중 비디오 크기 감소 방법을 제안하였다. 효율적인 비디오 크기 감소를 위해서 DCT 공간과 화소 공간에서의 크기 감소를 움직임 벡터와 매크로블록의 부호화모드를 이용하여 적응적으로 선택하였다. 이로 인하여 기존의 화소 공간을 이용한 트랜스코더에 비하여 움직임 검출 과정의 연산량을 크게 줄일 수 있었다. DCT 공간을 활용한 트랜스코더에서의 화질 저하 문제를 화소 공간에서 움직임 벡터 재검출 방법을 적용하여 종속 트랜스코더와 유사한 PSNR을 얻을 수 있었으며, 인트라 모드 및 영 벡터를 가지는 매크로블록의 경우는 DCT 공간에서 처리하여 화소 공간으로 변환시키는 추가적인 부호화 과정을 제거하면서 화질을 보존시킬 수 있었다. 제안한 복합형 트랜스코더를 종속 트랜스코더와 비교하였을 때 평균 PSNR 값은 0.1dB~0.5dB 감소하였으나, 시각적 화질은 유사하였다. 특히 움직임이 적은 영상에 대해서는 계산량을 크게 줄일 수 있었다. 또한 화소 공간에서의 크기 감소 시 수행되는 DCT 과정을 양자화 DCT 방법을 적용하여 제안한 트랜스코더의 성능을 향상시켰다.

참 고 문 헌

[1] N. Chaddha and A. Gupta, "A Frame-work for Live Multicast of Video Streams over the

Internet", in *Proc. IEEE International Conference on Image Processing*, Lausanne, Switzerland Vol. 1, pp. 1~4, 1996.

- [2] ISO/IEC, International Standard 11172, Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s, Aug. 1993.
- [3] Pedro A. A. Assuncqo and Mohammed Ghanbari, "A Frequency-Domain Video Transcoder for Dynamic Bit-Rate Reduction of MPEG-2 Bit Streams" *IEEE Trans. CSVT*, Vol. 8, No. 8, December 1998.
- [4] Oliver Werner, "Requantization for Trascoding of MPEG-2 Intraframes", *IEEE Trans. on Image Processing*, Vol. 8, No. 2, Feb. 1999.
- [5] Tamer Shanableh and Mohammed Ghanbari, "Heterogeneous Video Transcoding to Lower Spatio-Temporal Resolutions and Different Encoding Formats", *IEEE Trans. on Multimedia*, Vol. 2, No. 2, pp. 101~110, June 2000.
- [6] Neri Merhav and Vasudev Bhaskaran, "Fast Algorithms for DCT-Domain Image Down-Sampling and for Inverse Motion Compensation", *IEEE Trans. on CSVT*, Vol. 7, No. 3, pp. 468~476, June 1997.
- [7] Changhoon Yim and Michael A. Isnardi, "An Efficient Method for DCT-Domain Image Resizing with Mixed Field/Frame-Mode Macroblocks", *IEEE Trans. on CSVT*, Vol. 9, No. 5, pp. 696~700, Aug. 1999.
- [8] Alen Docef, F. Kossentini, K. Nguyen-Phi, I. Ragab Ismaeil, "The Quantized DCT and Its Application to DCT-Based Video Coding", *IEEE Trans. on Image Processing*, Vol. 11, No. 3, March 2002.
- [9] S-F. Chang and D.G. Messerschmitt "Manipulation and compositing of MC-DCT compressed Video", *IEEE J. Select. Areas Commun.*, Vol. 13, pp. 1~11, Jan. 1995.
- [10] Bo Shen, Ishwar K. Sethi, Bhaskaran Vasudev,

“Adaptive Motion-Vector Resampling for compressed Video Downscaling”, *IEEE Trans. on CSVT*, Vol. 9, No. 6, pp. 929~936, Sep. 1999.

[11] Jeongnam Youn, Min-Ting Sun and Chia-Wen Lin, “Motion Vector Refinement for High-Performance Transcoding”, *IEEE Trans. on*

Multimedia, Vol. 1, No. 1, pp. 30~40, Mar 1999.

[12] Niklas and Charilaos Christopoulos, “Transcoder Architectures for Video Coding”, *IEEE Trans. on Consumer Electronics*, Vol. 44, No. 1, pp. 88~98, Feb. 1998.

저 자 소 개



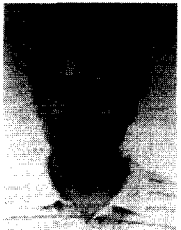
朴 一 雨(正會員)
2001년 2월 : 광운대학교 전자공학부 졸업(학사). 2001년 3월~현재 : 광운대학교 대학원 전자공학과 석사과정. 2001년 6월~현재 : 인티스 (주) 정보통신연구소 연구원. <주관심분야 : 비디오 트랜스코딩, 무선 멀티

미디어 시스템>



蔡 乘 助(正會員)
1997년 2월 : 광운대학교 전자공학과 졸업(학사). 1999년 2월 : 광운대학교 대학원 전자공학과 졸업(석사). 1999년 3월~현재 : 광운대학교 대학원 전자공학과 박사과정. 1999년 3월~

현재 : 바로비전 (주) 뉴미디어 연구소 <주관심분야 : 비디오 처리 및 압축, 무선 멀티미디어 시스템, Embedded System, Transcoder>



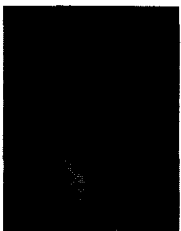
趙 鎔 局(正會員)
2001년 2월 : 광운대학교 전자공학부 졸업(학사). 2001년 3월~현재 : 광운대학교 대학원 전자공학과 석사과정. 2001년 6월~현재 : 인티스 (주) 정보통신연구소 연구원. <주관심분야 : 영상 후처리 기법, 동영상 포맷 변

환, 동영상 크기 변환>



吳 承 竣(正會員)
1980년 2월 : 서울대학교 전자공학과 졸업(학사). 1982년 2월 : 서울대학교 전자공학과 대학원 졸업(석사). 1986년 7월~1986년 8월 : NSF Supercomputer Center 초청 학생연구

원. 1987년 5월~1988년 5월 : Northeast Parallel Architecture Center 학생연구원. 1988년 5월 : 미국 Syracuse University 전기 및 컴퓨터공학과 졸업(박사). 1982년 3월~1992년 8월 : 한국전자통신연구원 근무(멀티미디어연구실 실장). 1992년 9월~현재 : 광운대학교 전자공학부 및 정보통신연구원 교수 (멀티미디어연구실). 2000년 3월~현재 : (주)인티스 정보통신연구소 연구소장. <주관심분야 : 비디오처리, 비디오 및 영상압축, 멀티미디어시스템>



明 鎭 洙(正會員)
1999년 2월 : 광운대학교 전자공학과 졸업(학사). 2001년 2월 : 광운대학교 대학원 전자공학과 졸업(석사). 2002년 3월~현재 : 광운대학교 대학원 전자공학과 박사과정. 2002년 3월~

현재 : (주)인티스 정보통신연구소 연구원. <주관심분야 : 실시간 비디오 처리 및 압축, Embedded System, Transcoder>



梁 昌 模(正會員)

1998. 2 : 한국항공대학교 항공전자 공학과(학사). 2000. 2 : 광주과학기술원 정보통신공학과(석사). 2002. 2 : 광주과학기술원 정보통신공학과 (박사수료). 2002. 3~현재 : 전자부 품연구원 연구원. <주관심분야 :

영상 부호화, 디지털 신호처리, 디지털 영상 워터마킹>



趙 威 德(正會員)

1981. 2 : 서강대학교 전자공학과(학사). 1983. 2 : 한국과학기술원 전기 및 전자공학과(석사). 1987. 2 : 한국 과학기술원 전기 및 전자공학과(박사). 1983. 3~1990. 3 : 금성전기(현

LG정보통신) 기술연구소 연구실장. 1990. 4~1991. 10. : 생산기술연구원 HDTV사업단 개발팀장. 1995. 11~1996. 5. : 미국 TCSI 공동개발 연구원. 1995. 2~1995. 10. : 영국 TTP-COM 공동개발 연구원. 1991. 11~현재 : 전자부품연구원 인터넷미디어연구센터장. <주관심분야 : Post-PC (차세대 Smart PDA), Interactive DTV 방송기술, 고품질 홈서버/게이트웨이기술, 디지털 방송/이동통신 연계 융합플랫폼기술, 무선인터넷응용기술>