

TMS320C80을 이용한 정지 영상 부호화에 관한 연구

김 상 기[†] · 정 진 현^{††}

요 약

DCT는 손실압축 중에서 가장 많이 사용되는 블록기반 변환코딩이다. DCT는 통계적으로 최적의 변환인 Karhunen-Loeve변환에 가깝다. 본 논문에서는 영상압축의 국제표준인 JPEG에 기반을 두어 TMS320C80을 이용하여 정지영상 부호화기를 만들었다. 이 정지영상 부호화기는 크게 변환기, 양자화기 그리고 인코더의 3가지 부분으로 나누어 구성되어 있다.

A Study On Still Image Codig With the TMS320C80

Sang-Ki Kim[†] · Chin-Hyun Chung^{††}

ABSTRACT

Discrete cosine transform (DCT) is the most popular block transform coding in lossy mode. DCT is close to statistically optimal transform - the Karhunen Loeve transform. In this paper, a module for still image encoder is implemented with TMS320C80 based on JPEG, which are international standards for image compression. The still image encoder consists of three parts - a transformer, a vector quantizer and an entropy encoder.

1. 서 론

영상을 디지털 신호로 처리하면 아날로그 신호에 비해서 잡음 처리에 유리하고 대역폭을 줄일 수 있다는 장점이 있지만 처리해야 할 정보량이 많아진다는 단점도 있다. 일반적으로 영상압축은 프레임내의 여유분(redundancy)을 줄이는 방법과 프레임간의 여유분을 줄이는 방법으로 나누어진다[2]. 본 논문에서는 정지영상의 경우에 프레임내 압축에 대해 다루고 있다.

일반적으로 영상압축 알고리즘은 변환(transform), 양자화(quantization), 부호화(encoding)의 세가지 단계

로 나눌 수 있다[1][2][3][4]. 첫 번째 변환(transform) 단계는 영상을 기저 함수(basis function)에 투영시킴으로 영상을 재구성하여 처리가 용이하도록 하는 것을 말한다. 영상은 주변의 화소값들 사이의 상관 관계가 높기 때문에 이미지를 주파수 성분에 따라 이산 여현 변환(discrete cosine transform)을 하게 되면 전 화면에 불규칙하게 퍼져있는 화소값들이 저주파함 쪽으로 집중되는 경향이 있다. 두 번째 양자화(quantization)는 이전의 단계에서 재구성된 영상 정보를 미리 정해놓은 기호에 대응시킨다. 이때 영상의 압축률이 결정되어진다. 압축률과 영상의 왜곡정도는 역함수관계를 가지게 되므로 어느 한 쪽에 편중되지 않도록 적절한 값을 취하는 것이 중요하다. 영상은 특성상 이산 여현 변환을 하면 저주파함에 집중되므로 저주파함에 큰 비중을

† 정 회 원 : (주)대동정밀 연구소 연구원

†† 정 회 원 : 광운대학교 제어계측공학과 교수

논문접수: 1997년 12월 19일, 심사완료: 1999년 1월 26일

두어 처리하고 고주파항에 작은 비중을 두어 이미지를 처리함으로써 데이터를 압축한다. 마지막 부호화(encoding) 과정은 제한된 수의 기호로 대응된 영상정보를 무손실 압축하는 과정이다. 기호의 발생 빈도를 이용한 엔트로피 부호화(entropy coding)을 주로 사용한다.

이러한 방법은 정지영상의 국제표준인 JPEG에서 사용하는 방법이다[5]. 본 논문에서는 JPEG규정에 의거한 정지영상의 압축을 TMS320C80을 이용하여 구현한다. TMS320C80은 32bit의 RISC구조를 가진 DSP(discrete signal processor)로서 한 개의 MP(master processor)와 4개의 PP(parallel processor)로 구성되어 있다[8]. TMS320C80은 멀티미디어 비디오처리를 위해 개발된 DSP로 이를 이용하여 영상을 입력받고 DCT 부호화 하여 영상을 압축하도록 시스템을 구성하였다.

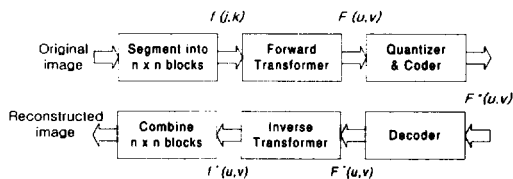
2. 정지 영상 부호화

2.1 부호화 알고리즘

(그림 1)에서 실제 영상압축의 과정을 간단히 나타내었다. 영상처리에 있어서 특히 DCT와 같은 경우에 영상의 임의의 $n \times n$ 의 크기로 분할하게 되는데 이는 계산상에 있어서 실제 영상의 크기를 그대로 DCT하기는 현실적으로 불가능하기 때문이다. 이렇게 영상을 분할함으로 인해서 양자화 크기가 클 경우 복원된 영상에서 블록화 현상(block artifact)이 보이기도 한다[1][4].

변환 과정은 영상을 기저함수(basis function)에 투영시키는 과정인데 본 논문에서는 주파수 변환의 일종인 이산 여현 변환(DCT)을 이용하였다. 영상을 주파수 공간으로 재구성하면 전 화면에 불규칙하게 퍼져있는 화소값들이 저주파항 쪽으로 집중되는 경향이 있다. 이는 영상이 특성상 주변의 화소값들 사이의 상관관계가 높기 때문이다. 양자화(quantization)는 제한된 수의 기호로 이전의 단계에서 재구성된 영상정보를 비슷한 여러 개의 값을 하나의 값으로 사상하는 과정이다. 이 과정에서 영상의 왜곡-비트율(R-D)관계가 결정되어 진다[2]. 압축률과 영상의 왜곡정도는 역함수관계를 가지게 되므로 어느 한 쪽에 편중되지 않도록 적절한 값을 취하는 것이 중요하다. 영상은 DCT를 하게 되면 저주파항에 집중하므로 양자화에 있어서 저주파항의 손실이 적도록 양자화 크기를 작게 하고 고주파

항은 양자화 크기를 크게 한다. 마지막 부호화 과정은 제한된 수의 기호로 대응된 영상정보를 무손실 압축하는 과정이다. 기호의 발생빈도를 이용한 엔트로피 코딩을 사용한다.



(그림 1) 변환 코딩의 구성도
(Fig. 1) Transform coding block diagram

2.2 이산 여현 변환(DCT)

DCT는 블록단위의 변환부호화 중에서 가장 널리 사용되는 방법이다. DCT는 변환하였을 때 성능이 이상적인 변환인 KLT(Karhunen Loeve transform)에 가장 가깝다 [1]. 이차원에서 이산함수 $f(j, k)$, $j, k=0, 1, \dots, N-1$ 의 DCT는 식(1)으로 정의되고, 역 DCT는 식(2)과 같다

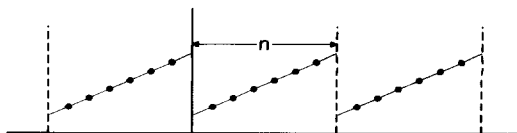
$$F(u, v) = \frac{4C(u)C(v)}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k) \cos \left[\frac{(2j+1)u\pi}{2N} \right] \cos \left[\frac{(2k+1)v\pi}{2N} \right] \quad (u, v=0, 1, \dots, N-1) \quad (1)$$

$$f(j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u, v) \cos \left[\frac{(2j+1)u\pi}{2N} \right] \cos \left[\frac{(2k+1)v\pi}{2N} \right] \quad (j, k=0, 1, \dots, N-1) \quad (2)$$

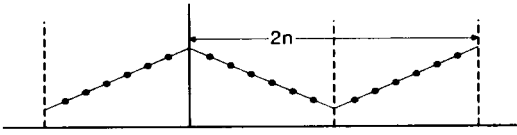
여기서

$$C(w) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } w=0 \\ 1 & \text{for } w=1, 2, \dots, N-1 \end{cases}$$

이다. 일반적으로 $N=8$ 의 8×8 의 블록을 사용하여 계산하는데 이는 행렬의 곱을 계산할 때 효율이 가장 좋기 때문이다[1][5].



(그림 2) DFT의 주기성
(Fig. 2) DFT Periodicity



(그림 3) DCT의 주기성
(Fig. 3) DCT Periodicity

DCT는 이산 푸리에 변환(discrete fourier transform)에서와 같이 허수 스펙트럼 성분이 발생하지 않기 때문에 좀더 높은 압축효과를 얻을 수 있다. DFT는 유한 주기 수열의 이산 푸리에 급수 표현이고 (그림 2)에서 보여 주듯이 절대적인 주기성(implicit periodicity)이 있다. 이러한 주기성은 주파수 공간에서 표현함으로 보여진다. 이 경우에 반복되는 원래 수열은 각각의 주기 사이에 불연속성을 발생하기도 하는데 불연속성은 변환의 효율을 적지 않게 감소시킬 수 있는 허수의 고주파 성분들의 결과이다. 비록 이 허수 성분들이 원 수열에 실제 부분이 아니라 할지라도 이 성분들은 주기적인 수열에서 선명한 경계를 복원시키는데 필요하다. 이러한 경계의 불연속성을 제거하기 위해서 원 수열을 수직 축에 대해서 대칭을 이루도록 하여 2배의 길이(원 수열의 길이를 n 으로 할 경우 $2n$ 의 수열을 만들 수 있는데 이렇게 확장된 수열은 경계의 불연속성도 없고, DFT를 하였을 때 허수 스펙트럼 성분도 없다(그림 3 참고). 이렇게 n 개의 수열을 $2n$ 으로 확장한 수열의 DFT는 원래의 n 개 수열을 DCT한 것과 동일하다. DCT를 사용할 때의 이점을 보면 크게 두 가지로 볼 수 있는데 첫째는 허수 스펙트럼 성분의 발생이 없어서 부호화 효과를 높일 수 있고, 블록화 현상을 많이 줄일 수 있다는 것이고 둘째는 실수 계산만 필요하다는 것이다. 이러한 이유에서 DCT가 영상처리에서 가장 많이 사용된다[3].

2.3 양자화와 부호화

식(2)의 DCT를 이용하여 얻어진 데이터를 양자화시키면 실제 많은 에너지를 가지고 있는 DC성분과 AC의 저주파 성분을 제외한 부분의 항들은 0이 될 수 있다. (그림 4)는 JPEG 권고안-(165)으로부터 얻은 예이다. 그림에서 낮은 주파수 값은 작은 수로 나누어줌으로 인해 실제 값과 거의 비슷한 값으로 사상시키고 높은 주파수는 큰 수로 나누어주는 것을 알 수 있다. DCT된 신호를 $F(u, v)$ 라고 하고 (그림 4)의 값

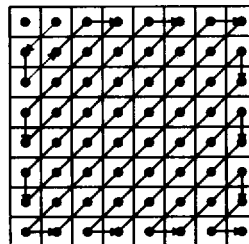
들을 $Q(u, v)$ 라고 하면 식(3)을 통하여 양자화기를 거친 신호 $F^*(u, v)$ 을 얻을 수 있다[5].

$$F^*(u, v) = Round\left(\frac{F(u, v)}{Q(u, v)}\right) \approx \left\lfloor \frac{F(u, v)}{Q(u, v)} + 0.5 \right\rfloor \quad (3)$$

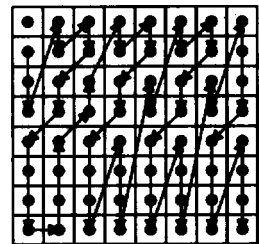
이러한 과정을 거친 신호들은 다시 기호의 발생빈도를 이용한 엔트로피 부호화를 하게 된다. 고주파항이 대부분이 0이 되므로 저주파에서 고주파순으로 (그림 5), (그림 6)과 같은 지그재그 스캔이나 대체 스캔을 하여 나열하면 더 나은 부호화 효과를 얻을 수 있다. 그림의 화살표 방향으로 2차원 배열의 영상 데이터를 1차원 배열로 바꾸어 준다. 이때 0이 아닌 값이 나올 때까지만 데이터를 읽어 들이고 어떤 값부터 0이고 그 이후 모든 값이 0이면 마지막 0이 아닌 값까지만을 데이터로 이용하고 나머지는 사용하지 않는다 복원할 때는 블록의 크기를 알고 있기 때문에 데이터의 나머지 항은 0으로 채워 $n \times n$ 길이의 데이터로 바꾼 뒤 역 스캔을 한다.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	66	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(그림 4) 양자화 표의 예
(Fig. 4) Sample quantization table



(그림 5) 지그재그 스캔
(Fig. 5) Zigzag scan

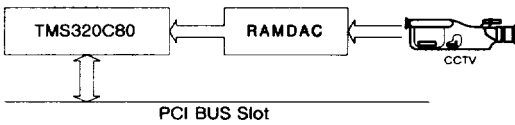


(그림 6) 대체 스캔
(Fig. 6) Alternative Scan

2.4 시스템의 구성

(그림 7)은 본 논문의 전체적인 DCT 인코더의 구

성도이다. 비디오용 RAMDAC을 이용하여 CCTV카메라로부터 얻은 영상을 TMS320C80을 이용하여 DCT 부호화를 하고, 코딩된 데이터는 역 DCT코딩을 하여 RAMDAC으로 다시 보낸다. 그러면 RAMDAC을 통하여 모니터에 영상이 보여진다. 이때 카메라로부터 얻어지는 영상은 640×480크기의 흑백영상이고 NTSC 비디오형식이다.

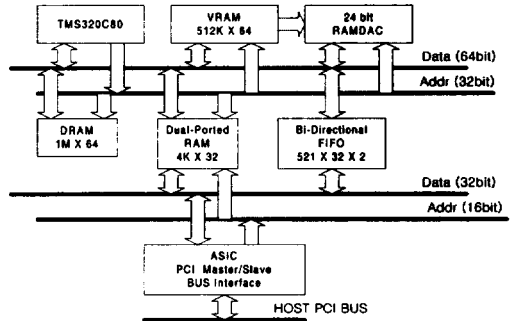


(그림 7) 시스템 구성도
(Fig. 7) System block diagram

TMS320C80은 32bit의 RISC구조를 가지고 있는 master processor(MP)와 32bit의 ADSP(advanced DSP) 구조를 가진 4개의 parallel processor(PP), transfer controller (TC), video controller(VC), 50K byte의 SRAM으로 구성되어 있는 DSP 칩이다[8]. 이러한 특성 때문에 흔히 MVP(multimedia video processor)라고 불린다. 이 MVP를 사용함에 있어서 다른 DSP와 비교되는 가장 큰 장점은 바로 내부가 5개의 프로세서로 구성이 되어 있다는 점이다. 다른 DSP의 경우 프로세서들의 병렬처리를 위해서는 복잡한 하드웨어 및 소프트웨어가 필요하다. 그러나 MVP를 이용할 경우 32비트의 데이터버스로 내부가 연결되어 있어 병렬처리를 위한 하드웨어 구성 작업이 필요하지 않고, 마스터 프로세서가 정의되어 있어 마스터 프로세서의 제어로 병렬처리가 가능하기 때문에 소프트웨어 역시 많은 도움을 받을 수 있다.

또한 4G 바이트의 메모리 영역이 있고, 메모리 맵(memory-mapped)I/O로 구성되어 있다. 넓은 메모리 영역과 메모리 맵 I/O는 영상을 처리를 위한 버퍼를 충분히 제공할 수 있고, 여러 가지 외부 소자들을 연결하기에 편리하게 해준다.

(그림 8)은 인코더의 구성도이다. NTSC 영상을 얻기 위한 RAMDAC과 4M 바이트의 비디오 RAM, 주 프로세서인 TMS320C80, 640×480크기의 이미지를 처리하기 위한 8M byte의 메모리, 코딩된 데이터 전송을 위한 양방향 FIFO 그리고 PCI 버스와 통신을 위한 PCI Bus interface 등으로 구성되어 있다[7].



(그림 8) DCT 부호화기의 구성도
(Fig. 8) DCT encoder block diagram

2.5 영상의 평가

영상을 복원한 후에 평가하는 방법은 크게 2가지가 있다. 주관적인 평가 방법과 객관적이 평가 방법으로 나누어지는데, 주관적인 평가라 하는 것은 복원된 영상을 여러 사람에게 보여주고 사람들 각각의 주관적인 평가를 토대로 영상을 비교하는 것이고 객관적 평가라 하는 것은 원래영상과 복원된 영상과의 차이를 수식적으로 계산하는 PSNR을 이용하여 비교하는 방법이다.

$$PSNR = 10 \log_{10} \left[\frac{255^2}{\frac{1}{NM} \sum_{m=0}^M \sum_{n=0}^N [(U(m, n) - U_E(m, n))^2]} \right] \tag{4}$$

여기서 M, N은 영상의 크기이고, $U(m, n)$ 은 원래 영상이며 $U_E(m, n)$ 은 복원된 영상이다. PSNR의 단위는 dB이다. 이 두 가지 평가 방법은 어느 한 쪽에 치우쳐서는 안 된다. 예를 들어 영상의 어느 작은 한 부분만 손실되었을 경우 사람의 눈에는 그 손실된 부분으로 인하여 전체적인 영상을 나쁘게 볼 수도 있지만 PSNR을 구해보면 큰 값을 가질 수도 있기 때문이다. 반대로 DCT에서처럼 저주파 성분에 민감한 사람의 눈에는 영상이 좋게 보일 지도 모르지만 PSNR을 구해보면 낮은 값을 가지는 경우도 있다[2][3].

3. 결과 및 결론

3.1 실험 결과 및 결론

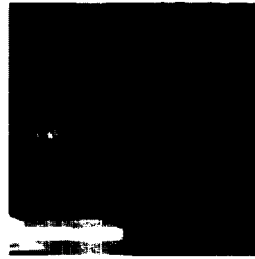
실험은 (그림 7)과 (그림 8)에서와 같이 CCTV와 RAMDAC을 이용하여 영상을 받고 TMS320C80을 이용하여 DCT 부호화와 복호화를 한 후 비디오 메모리

에 영상을 전송하여 영상이 RAMDAC을 통하여 모니터에 나타나도록 하였다. 화면(640×480)에 원 이미지와 코딩 후 이미지를 모두 나타내어 비교하기 위해서 이미지의 크기는 320×240으로 하였다. 실험 중 알고리즘상의 문제가 없음에도 불구하고 자주 다운되는 경우가 발생하였는데, 이는 영상의 버퍼로 사용한 메모리의 양이 커짐으로 발생한 문제였다. 영상 데이터가 알고리즘의 어떤 부분을 침범한 것으로 예상되어 지고, 불필요하다고 생각되는 버퍼링을 줄임으로 해결을 되었다. 메모리를 많이 사용했기 때문에 메모리 문제는 고려하지 않았었다.

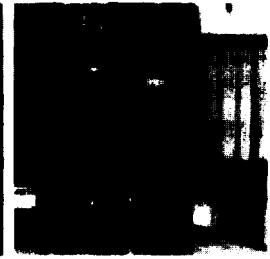
DCT 계산을 정수 연산하였을 경우 영상의 블록화가 사람의 눈에 인식될 정도로 보였으나 TMS320C80 내부에 있는 IEEE-754 부동 소수점 하드웨어 장치(floating point hardware unit)를 이용하여 실수 연산을 할 경우 블록화 현상은 상당히 감소한 것을 볼 수 있었다. 320×240 크기의 영상을 입력받아 DCT 코딩을 하고 다시 복원하여 이를 화면에 보이는데 걸리는 시간은 18초 정도이다. 이러한 계산상의 지연 시간 때문에 영상을 연속적으로 입력받는 경우 영상이 일정 시간 간격으로 끊어져서 보이게 된다.

(그림 9)와 (그림 10)은 DCT 부호화 전의 원래영상이고 (그림 11)과 (그림 12)는 DCT 부호화와 양자화를 한 후 다시 복원한 영상이다. 앞의 (그림 4)에서 제시한 양자화표를 이용하여 부호화 한 후 영상을 복원할 때 Cup영상의 경우 PSNR값은 약 38.77 dB이고 0.65bpp로 압축이 되었고 Cooler영상은 36.49dB이고 1.07bpp로 압축된다. 영상을 보면 전체적으로는 원래영상과 차이가 없이 보이나 자세히 보면 영상의 높은 공간 주파수를 가진 부분이나 복잡한 경계조건을 가진 부분을 보면 양자화로 인하여 경계정보가 상실되었음을 볼 수 있다. (그림 13)과 (그림 14)는 DCT 부호화를 하기 전과 후의 영상의 차이 값인데, (그림 13)의 Cup영상의 경우 글자 부분과 컵의 모양의 경계성분이 잘 나타나 있고 (그림 14)의 Cooler의 경계부분의 성분이 잘 나타나있다. 또한 영상의 전체적인 블록화현상도 확실히 볼 수 있다.

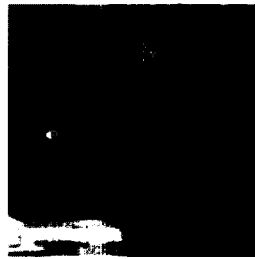
전체적으로 Cup영상이 Cooler영상보다 높은 PSNR 값을 가지는데, 그 이유는 영상의 경계를 보면 Cooler영상의 경우 Cup영상보다 복잡하고 많은 경계성분을 가지고 있기 때문이다. 이러한 경계(고주파)성분들은 양자화 시에 대부분이 상실되기 쉽기 때문에 Cooler영상의 PSNR값이 더 작다.



(그림 9) 원래영상
(Fig. 9) Original Image
(Cup)



(그림 10) 원래영상
(Fig. 10) Original Image
(Cooler)



(그림 11) 복원된 영상
(Fig. 11) Reconstructed
image (Cup)



(그림 12) 복원된 영상
(Fig. 12) Reconstructed
Image (Cooler)



(그림 13) 차 영상
(Fig. 13) Difference Image
(Cup)



(그림 14) 차 영상
(Fig. 14) Difference Image
(Cooler)

앞에서 제시한 알고리즘을 TMS320C80을 이용하여 구현할 때 모든 값을 정수로 계산하면 영상을 복원할 때는 약간의 블록화가 보이게 된다. 물론 양자화를 하지 않으면 블록화 현상은 발생하지 않지만 영상을 압축하기 위해서 양자화는 필수과정이다. 정수계산과 실수계산시의 차이는 고주파 성분에서 나타난다. 정수계산에서는 계산 시에 발생하는 모든 소수점이하의 값들은 반올림되기 때문에 이로 인한 오차는 작지만 고주파성분들 중에서 양자화 시에 0값으로 되는 경우가 발

생시키고 블록화 현상은 심하게 나타난다.

3.2 추후 과제

부호화와 디코딩의 계산량이 많기 때문에 DCT계산 시간이 오래 걸리지만 알고리즘을 분할하여 PP(parallel processor)를 사용하고 불필요한 버퍼링을 줄여 코드의 최적화를 이루면 빠른 계산이 가능하리라 예상된다. 예를 들면 영상의 블록화, 이산 여현 변환, 양자화 및 스캔, 엔트로피 부호화의 과정을 각각의 프로세서로 분할하여 처리함으로써 시간을 단축시킬 수 있다. 물론 DCT의 빠른 알고리즘을 적용함으로써 시간을 줄일 수 있다.

DCT를 이용한 영상의 압축에서 가장 큰 문제점인 영상의 블록화 현상은 영상의 압축율을 높이면 어느 정도 개선이 되기는 하지만 영상정보량이 증가한다는 단점을 가지고 있다. 근래 들어 이러한 단점을 개선하고자 웨이브렛(wavelet) 변환을 이용한 영상 압축법과 프랙탈(fractal) 방법 등의 연구가 활발히 진행 중이다.

참 고 문 헌

- [1] K. R. Rao, P. Yip, "Discrete Cosine Transform Algorithms, Advantages, Applications," Academic Press, 1990.
- [2] A. Murat Tekalp, "Digital Video Processing," Prentice Hall, 1995.
- [3] Majid Rabbani, Paul W. Jones, "Digital Image Compression Techniques," SPIE Optical Engineering Press, 1991.
- [4] Khalid Sayood, "Introduction to Data Compression," Morgan Kaufmann Publishers, 1996.
- [5] K. R. Rao, J. J. Hwang, "Techniques and Standards for Image · Video and Audio Coding," Prentice Hall, 1996.
- [6] Phillip E. Mattison, "Practical Digital Video with Programming Examples in C," John Wiley & Sons, 1994.
- [7] Prabhat K. Andleigh, Kiran Thakrar, "Multi-media System Design," Prentice Hall, 1996.
- [8] "TMS320C80(MVP) User's Guide," Texas Instruments, 1995.
- [9] "User's Guide for the GRIFFIN 100 (TMS320C80 PCI-Bus Board)," Ariel Corporation, 1996.

김 상 기

e-mail : violet@shinbiro.com

1996년 광운대학교 제어계측공학과
졸업(학사)

1998년 광운대학교 대학원 제어계
측공학과 졸업(공학석사)

1998년~현재 (주)대통령정밀 연구소
연구원

관심분야 : 영상처리, MPEG, 자동화시스템

정 진 현

e-mail : chung@daisy.kwangwoon.ac.kr

1981년 연세대학교 전기공학과(공
학사)

1983년 연세대학교 대학원 전기공
학과(공학석사)

1990년 Rensselaer Polytechnic In-
stitute(ph.D)

1991년~현재 광운대학교 제어계측공학과 부교수, 신
기술연구소 연구원

관심분야 : fuzzy control, neural network, MPEG II,
stochastic control, CIM