
초저전송율 동영상 부호기를 위한 퍼지 양자화 및 울 제어에 관한 연구

양근호*

Fuzzy Quantization and Rate Control for Very Low Bit-rate Video Coder

Keun-Ho Yang*

요 약

본 논문에서는 H.263 부호기에서 적정 화질과 일정한 전송율을 유지하는 양자화 변수 추정을 위한 퍼지 제어기를 제안하였다. 퍼지화 방법으로는 Mamdani 방법을 이용하며, 역퍼지화는 무게중심법을 이용하였다. 움직임 벡터에 따른 시간적 특성을 추정할 때 공간영역에서 분산과 엔트로피는 시각적 특성과 상관관계가 있다. 따라서 퍼지 입력변수로 영상의 분산값, 엔트로피 및 움직임 벡터를 이용하였다. 퍼지 소속함수를 유도하고, 시각적 특성에 적합한 퍼지감도를 결정하였으며, FAM 뱅크를 설계하였다. 실제 동영상 압축에 퍼지 규칙을 적용하였으며, 퍼지 양자화를 이용하여 효율적인 울제어와 최적의 비트할당 및 주관적 화질 특성이 유지되는 결과를 얻었다.

ABSTRACT

In this paper, we proposed a fuzzy controller for the evaluation of the quantization parameters in the H.263 coder to optimize the subjective quality of each coded frame, keeping the transmission rate constant. We adopted the Mamdani method for fuzzification and the centroid method for defuzzification. The energy and entropy are correlated to features of the HVS in spatial domain, while motion vectors are used to estimate the temporal characteristics of the signal. And then, the fuzzy inputs adapted the variance and the entropy in spatial domain, and the motion vector in temporal domain. We induced the fuzzy membership function and decided the fuzzy relevance to be compatible in visual characteristics. And then, we designed FAM banks. The fuzzy technology has been applied to a practical video compression. This results is obtained an effective rate control technique, an optimum bit allocation and a high subjective quality using fuzzy quantization.

키워드

Fuzzy Relevance, Mamdani, Centroid, FAM banks

1. 서 론

퍼지시스템은 주어진 대상에 대한 전문가의 경험적 지식이나 공학적인 상식을 토대로 체계적인 퍼지 규칙과 소속함수를 생성하고, 추론기법에 의해, 주어진 입력에 대한 적절한 출력을 구하여 대

상을 제어한다. 수학적 모델링이 복잡하거나 비선형적 요소가 많아서 제어 대상에 적용하기 어려운 경우 퍼지시스템은 각각의 변수로부터 퍼지함수를 추출하고 합성하는 기술을 제공한다.

영상 압축에서 제어의 개념은 주로 양자화기의 선택과 적절한 울제어(rate control)에 있다[1,2].

*한려대학교 멀티미디어정보통신공학과

접수일자 : 2003. 10. 10

또한, 시각적으로 유연한 화질을 제공해야 한다는 변수가 있다. 파형 분석에 의한 처리는 수리적인 것이며 화질 요구사항은 감각적인 것이기 때문에 불일치가 존재한다. 즉 시각적으로 만족스런 화질을 강조하면 수리적인 채널용량을 만족할 수 없으며 반대로 마찬가지이다. 따라서 이러한 불일치 문제를 해소하면서 최적의 부호기를 설계하는 연구가 요구된다.

본 논문에서 동영상 부호기에서 비트율을 유지하면서 만족스런 화질을 제공하는 퍼지 양자화 기법 및 울제어 방법을 제시하였다. 부호기의 모형은 영상전환기에 응용될 수 있도록 설계된 H.263이며, 30[Kbps] 이하에서 실험하였다. 양자화 변수를 제어하며, 매크로블록 단위의 울제어 기법을 유지하면서 퍼지함수를 적용하였다. 주어진 비트를 최대한 사용하고 프레임 내 국부 특성에 따라 최적으로 비트를 할당하였다. 주로 움직임이 큰 부분에서 오차가 크게 발생하고 시각적 감도도 높기 때문에 이 점을 고려하였다. 공간영역과 시간영역으로 퍼지감도를 적용하여 프레임 단위의 오차보다 프레임 내 및 프레임간 감각적 화질을 유지하는 것에 주안점을 두었다. 이를 위해 시각 특성에 적합한 소속함수를 설계하고 퍼지화하여 양자화 변수를 제어하도록 하였으며, 목표 비트율에 적합하고 적정 화질을 유지하도록 하였다.

본 논문의 구성은 II장에서 H.263 부호기의 양자화와 울제어 방법에 대하여 고찰하고, III장에서는 퍼지 양자화와 울제어 방법에 대하여 서술하였으며, IV에서는 컴퓨터 시뮬레이션 및 검토를 통하여 제안한 퍼지 감도를 이용한 방법과 기존의 방법을 비교·분석하였으며, V장에서 결론을 맺었다.

II. H.263 부호기의 양자화와 울제어

영상을 압축하는데 가장 중요한 출발점은 영상 내에 존재하는 공간적, 시간적 상관성을 이용하는 것이다. 이 상관성은 영상의 특성에 따라 다르기 때문에 일률적인 기준을 적용하면 영상마다 다른 압축율을 얻게된다. 그러나 전송로에서 주어진 전

송속도에 맞추기 위해서는 어떤 방식으로든지 전송율을 일정히 해야한다. 특히 고정 비트율 전송시에 이 문제는 절대적이며 가변 비트율 전송에서도 어느 정도의 비트율 기준화는 필요하다[3,4,5].

H.263에서는 여타의 H.261, MPEG 등과 같은 표준안과 마찬가지로 부호기 출력단에 버퍼를 두고 상태를 측정하여 양자화 행렬을 조절하는 기법을 사용한다. 이 경우 버퍼량에 따라 일률적으로 양자화 레벨을 크거나 작게 비트율을 조절할 수 있으며, 이러한 경우, 영상의 국부적인 특성을 고려해야 한다. 즉 H.263에서 기본적으로 주어진 것은 HRD(hypothetical reference decoder)이며, HRD와 부호기는 같은 QCIF 만큼의 동일한 시각 주파수를 가진다[3,4].

H.263에서 새로운 양자화 변수는 다음 (1)식과 같이 이전 프레임의 평균 양자화 변수값에 각각 프레임간 및 프레임내의 비트 변화량을 고려하여 결정한다[3,4].

$$QP_{new} = \overline{QP}_{i-1} \left(1 + \frac{\Delta_1 B}{2B} + \frac{12\Delta_2 B}{R} \right) \quad (1)$$

여기서, $\Delta_1 B = B_{i-1} - \overline{B}$, $\Delta_2 B = B_{i, mb} - \frac{mb}{MB} \overline{B}$ 이다. 이때 \overline{QP}_{i-1} 는 이전 영상에 사용되었던 평균 양자화 변수이며, $B_{i, mb}$ 는 영상에서 현재 매크로블록까지 사용한 비트수, mb 는 현재 매크로블록 번호이다. 또한 MB 는 영상 내 매크로블록 수를 의미한다.

III. 퍼지 양자화 및 울제어

1. 퍼지 입력변수 추정

동영상 부호기에 사용되는 양자화 방법으로서 퍼지추론을 이용한 퍼지 양자화 방법을 제안한다. 여기에는 분산과 엔트로피, 움직임 벡터 크기 등을 입력변수로 갖는 FAM(fuzzy associative memory) 뱅크가 이용되며, 이는 세 가지의 입력변수

를 갖고 하나의 출력변수를 갖기 때문에 4차원 벡크가 된다. 이러한 퍼지 양자화 방법에서 효과적인 결과를 얻기 위해서는 FAM 벡크내 퍼지규칙의 구성과 또한 그 규칙내의 퍼지집합의 소속함수 형태와 배치 등이 매우 중요한 역할을 하게 된다. 본 논문은 퍼지 추론법과 역퍼지화 방법으로 각각 Mamdani와 무계중심법을 이용하였다.

동영상 부호기의 전체 및 국부 특성을 나타내는 곳은 변환영역이다. 그러나 변환계수는 원영상 자체의 변환에 의한 것이 아니라 차동부호화된 차분값의 변환계수이기 때문에 원영상 자체를 표현하지 않는다. 더욱이 예측에 이용된 참조영상은 움직임 보상된 것이어서 이로 인한 보상오차가 첨가되어 있으므로 변환계수를 퍼지변수화하는 대신에 원영상을 이용한다. 이와 같이 시각적 감도는 명암 및 주파수 특성에 주로 의존하므로 이에 근거하여 변수를 설정하고 소속함수를 결정한다.

첫 번째로 고려할 명암 특성은 웨버의 법칙이라 일컬어지는 콘트라스트 비율로부터 유도된다. 주위 밝기에 대한 대상 물체(매크로블록)의 명암 세기에 따라 감도가 변화함을 의미하는데, 이를 단순화하여 매크로블록 내 명암의 변화량 즉 다음 (2)식과 같은 분산량으로 표현된다.

$$E_k = \frac{1}{N \times M} \sum_{(i,j) \in MB_k} (x_{i,j} - \bar{x}_k)^2 \quad (2)$$

두 번째로 고려할 엔트로피의 개념은 주어진 영역 내 명암의 크기에 관계없이 변화량을 나타낸다. 이를 정보량으로 표현하며, 변화가 많으면 정보량이 많음을 의미하고, 적으면 정보량도 적게 된다. 매크로블록 내 엔트로피는 다음 (3)식과 같이 정의된다.

$$H_k = \sum_{i=0}^{L-1} p_i \log_2 \frac{1}{p_i} \quad (3)$$

동영상은 공간영역 뿐만 아니라 시간영역으로 진행되기 때문에, 세 번째로 고려할 특성은 시간적 주파수이다. 이것은 부호기에서 움직임보상에 의해 충분히 고려되고 있다. 다만 움직임의 크기

에 따라 움직임보상 오차가 달라지고 감도에 차이가 발생하므로 이것을 또 하나의 퍼지변수로 고려할 수 있다. 움직임 에너지는 다음 (4)식과 같이 정의된다.

$$M_k = MV_{k,h}^2 + MV_{k,v}^2 \quad (4)$$

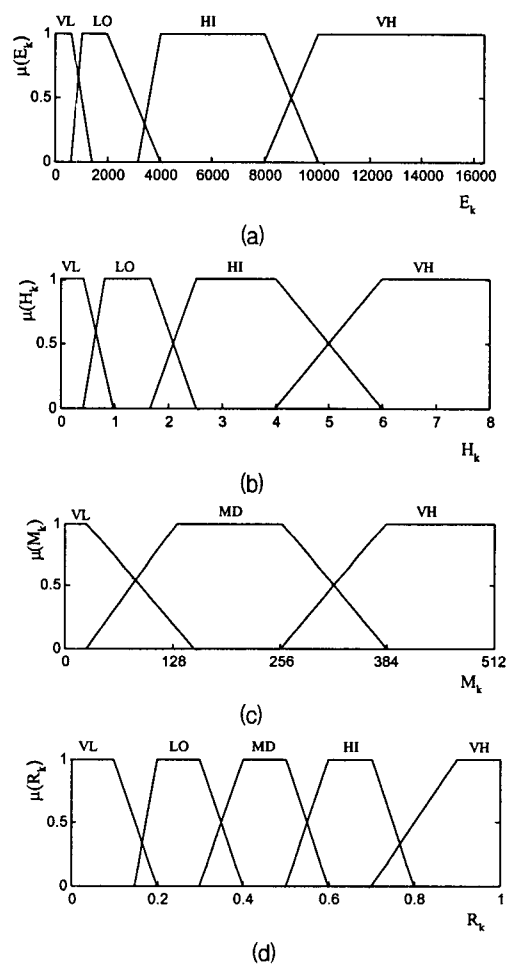


그림 1. 감각적 소속함수 결정 (a) 분산, (b) 엔트로피, (c) 움직임 에너지, (d) 출력감도 특성
Fig. 1 The membership functions for (a) variance, (b) entropy, (c) the motion energy and (d) the output relevance using fuzzy classifier

움직임 벡터의 크기가 적으면 인접한 프레임에서 높은 상관도가 존재함을 의미하므로 움직임보상에 의해 효율적으로 부호화될 수 있다. 움직임

벡터가 커지면 움직임보상 오차가 커지기 시작한다. 따라서, 이것을 줄이는 방향으로 부호화를 해야 한다. 즉 매우 적은 움직임 블록에서는 시각적 감도가 낮고 반대로 매우 큰 영역에서도 감도는 상대적으로 낮아진다. 분산값, 엔트로피, 움직임 벡터 크기 등에 따른 각각의 소속함수 및 퍼지감도 결정을 그림 1에 나타내었다.

2. FAM بانک 설계

움직임과 관련하여 대상으로 하는 매크로블록의 움직임과 특성 (PM_k)이 감도에 직접적인 영향을 주도록 한 것은 시간적 마스킹을 고려한 것이다. 현재 매크로블록을 부호화 한 후에 다음 매크로블록을 위해 M_k 를 PM_k 로 하여 얻는다 이전에 큰 움직임 있다가 갑자기 적어지면 움직임이 멈춘 것을 의미하며 이것은 장면전환의 경우에 해당한다. 이때 높은 감도를 보이는 것으로 보이지만 이것은 시간 마스킹을 나타내는 것이며, 무조건 높은 감도를 보이는 것은 아니며 대역통과 특성을 고려하여 현재 매크로블록의 움직임과 함께 FAM بانک 결정에 이용한다. 즉 FAM بانک 설계에 이용된 퍼지규칙은 다음과 같다.

분산과 엔트로피가 중간영역에서 감도는 높고 너무 크거나 낮은 경우는 감도가 낮도록 한다. 또한 움직임이 적은 경우는 감도가 낮고, 적절한 움직임이 존재하거나 장면전환이 일어나는 경우는 감도를 높게 설정한다. 그러나 움직임이 매우 큰 경우에는 시각적 특성을 고려하여 감도를 낮게 설정한다.

3. 퍼지 양자화 및 율제어

율제어는 화질 왜곡과 목표 비트율 사이의 관계로부터 최상의 화질을 얻을 수 있는 비트 할당 방법을 객관적인 기준에 의거해서 찾았다. 그러나 영상 시퀀스는 영상 그 자체만으로 좋은 화질을 전송할 수 없으며, 관측자의 시각 및 주변환경 등에 따라서 변화할 수 있다. 따라서 본 논문에서는 HVS와 퍼지 논리를 이용하여 퍼지감도를 구하고 이를 양자화 및 율제어에 이용하였다.

정해진 감도특성은 H.263의 양자화를 위한 양자화 변수 QP의 변화에 이용한다. QP는 1에서 31까지 변화하며 버퍼상태를 고려하여 전체적 및 국부적으로 제어된다. 전체적 제어는 프레임 단위의 변화를 의미하며 평균 비트율과 이전 프레임에서 사용된 비트를 고려하여 변화된다. 국부적 제어는 프레임 내에서 주어진 가용 비트 자원을 각 매크로블록에 효율적으로 배분하는데 목적이 있다. 이러한 목적에 맞도록 새로운 양자화 변수값을 결정하도록 다음 식(5)과 같이 하였다.

$$QP_{new} = \overline{QP}_{i-1} (1 + G_{adj} + L_{adj}) \quad (5)$$

여기서, G_{adj} 와 L_{adj} 는 각각 다음과 같다.

$$G_{adj} = \frac{(B_{i-1} - \overline{B})}{2B} \times \left(\frac{F_{sum}}{MB} + \alpha \right)$$

$$L_{adj} = \frac{12 \times (B_{i,mb} - \frac{F_{i,mb}}{F_{sum}} \times \overline{B})}{R} \times (F_i + \beta)$$

(5)식에서 보인 제안한 양자화 변수 결정에서 프레임간은 G_{adj} 에 나타낸 바와 같이 프레임 단위의 퍼지감도를 반영하였으며, 프레임 내에서는 L_{adj} 와 같이 매크로블록 단위로 퍼지감도의 변화량을 반영하도록 하였다.

IV. 컴퓨터 시뮬레이션 및 고찰

본 논문에서는 영상의 국부적 특성을 고려하도록 각 매크로블록 단위로 QP 변화를 처리하였다. 또한, 프레임 건너뛴을 하지 않도록 하였다. 사용된 영상 시퀀스는 176×144 화소의 QCIF 형식의 Foreman이다.

분산, 엔트로피, 움직임 벡터 크기 등의 퍼지변수를 데이터 영역에서 추출하여 퍼지화하였으며, 퍼지화 과정에서 min-max 기법 즉 소속함수를 계산하기 위해 최소값으로 가장 분명한 퍼지값을 추출하였으며, 퍼지집합을 위해서는 각 소속함수

로부터 요소를 더하는 의미에서 최대값을 선택하였다. 비퍼지화를 위해 무게중심 기법을 이용하여 최종 퍼지감도를 구하였다.

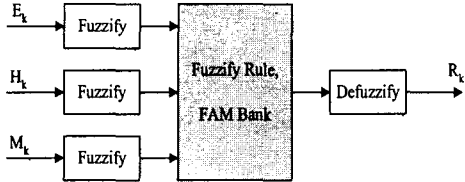


그림 2. 퍼지 시스템
Fig. 2 Fuzzy system

그림 2에 퍼지화 및 역퍼지화 과정을 나타내었으며, 퍼지시스템에서 얻어진 퍼지감도를 양자화 및 울제어에 적용하였다.

그림 3과 4는 각각 프레임별로 버퍼상태의 변화과정과 평균 양자화 변수값의 변화상태를 나타내었다. 퍼지 방법에 의한 버퍼제어 및 양자화 변수 제어 방법이 H.263의 제어 방식을 유지하면서도 그림 5에 나타낸 바와 같이 객관적 성능 비교가 가능한 SNR이 향상됨을 알 수 있으며, 주관적 화질 평가에서도 향상된 결과를 얻었다. 이러한 결과는 시각시스템에 적합한 퍼지함수를 유도하여 각 매크로블록 단위로 퍼지감도에 의한 버퍼 및 양자화 과정을 제어함으로써 보다 더 효과적인 부호화가 이루어졌다. 주어진 가용 자원을 최대한 사용한 결과를 그림 6에 보였으며, 움직임에 따른 퍼지감도를 적용하여 가장 좋은 성능을 보인 54번째 프레임에 대한 움직임 에너지와 퍼지감도의 관계를 그림 7에 나타내었다.

TMN5 실험에 이용한 영상 시퀀스들은 모두 150 프레임까지 할당된 비트를 모두 사용하지 못하고 누적되어 남는 비트수가 영상 시퀀스에 따라 3,000에서 10,000까지의 비트수가 남게 됨을 알 수 있다. 제시된 결과에서 원영상의 프레임율이 30[Hz]일 때 전송 프레임율 및 비트율을 각각 10[Hz]와 30[Kbps]이다면, 프레임당 목표 비트율은 3,000[bits]가 된다. TMN5는 이것에 제한을 두어 울제어를 수행하므로 목표 비트율을 초과하지 않지만 가용자원이 효율적으로 사용되지 못하고 남는 결과를 보여주고 있다. 이러한 결과는 영상

시퀀스의 움직임이 적은 경우일수록, 배경이 밝을수록 여분의 비트수가 더 크게 증가함을 알 수 있었다. 그러나 퍼지시스템은 여분의 비트가 증가하지 않고 목표 비트 근처에서 변화함을 보여주고 있으며 최종적으로는 영(zero)에 수렴함을 알 수 있다. 이는 할당된 비트가 프레임 내에서 충분히 이용되고 있음을 알 수 있다. 본 논문에서 제안한 퍼지감도를 이용한 양자화 및 울제어 방법은 초저전송을 동영상 부호기에 분산에너지, 엔트로피에 대한 공간적 소속함수와 움직임에 대한 시간적 소속함수를 유도하고 퍼지 감도에 따라 비트율을 제어하는 퍼지 논리를 적용함으로써 동영상의 시각적 감도에 따라 적절한 비트할당이 이루어짐을 증명하였으며, 그 결과 시각적으로도 매우 양호한 결과를 얻을 수 있었다.

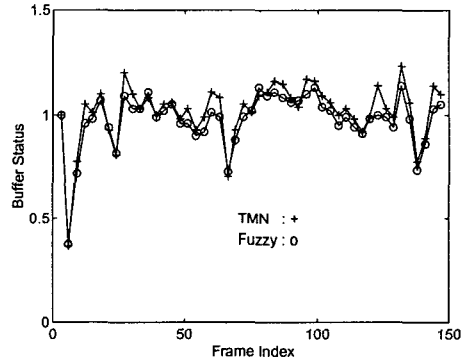


그림 3. 프레임별 버퍼상태
Fig. 3 Buffer status in frames

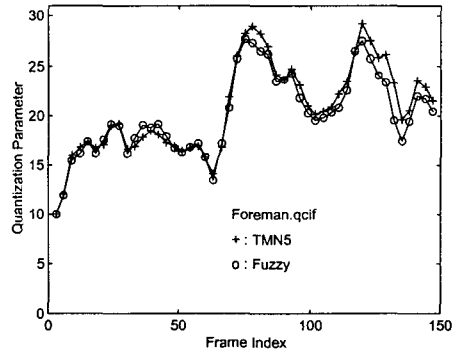


그림 4. 프레임별 평균 양자화 변수
Fig. 4 Mean quantization parameter in frames

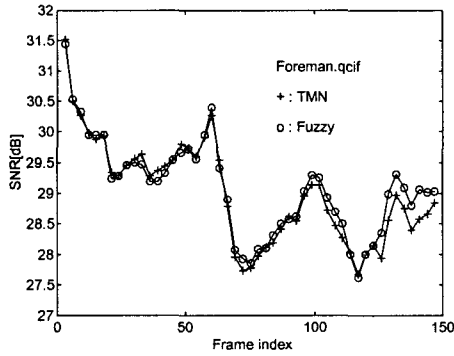


그림 5. 프레임별 성능(SNR) 비교
Fig. 5 SNR comparison in frames

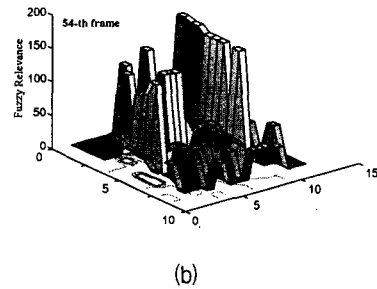


그림 7. 가장 성능이 향상된 54번째 프레임(30[Kbps])
(a) 움직임 에너지, (b) 퍼지감도

Fig. 7 Motion energy and fuzzy relevance of the 54-th frame with the best performance, 30[Kbps]

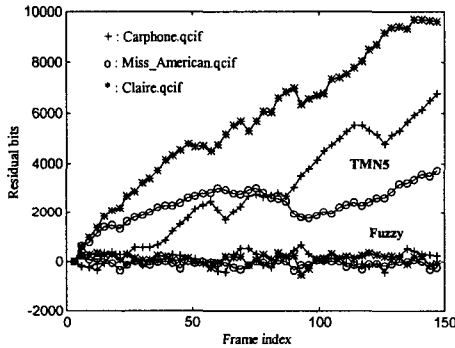
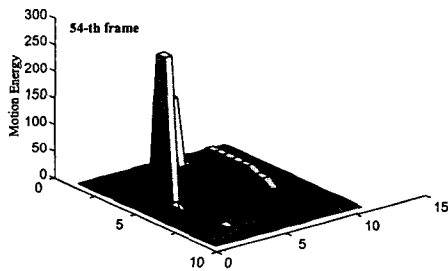


그림 6. 각 프레임에서 사용하고 남은 가용 비트 자원(30[Kbps])

Fig. 6 An available residual bit resource in frames(30[Kbps])



(a)

V. 결론

영상전화 및 회의에 이용되는 초저전송율 동영상 부호기에 퍼지 논리를 적용하는 기법을 제안하였다. 전송율이 낮아질수록 파형 분석에 의한 압축 효과가 줄어들게 되며 블록현상이 커지게 된다. H.263이 수십[Kbps]급의 전송율에 맞추어 설계된 것이지만 실제로 깨끗한 화질을 보이지 못하는 이유가 여기에 있다. 따라서 시각적인 감도의 분산이나 적절한 분배가 필요해진다.

본 논문에서는 시각적인 특성에 적합하게 퍼지 변수를 결정하고 소속함수를 유도하였다. 소속함수의 결정은 출력 영상의 주관적 및 객관적 화질에 결정적인 영향을 미친다. 따라서, 시각적 감도에 근거한 소속함수의 결정은 타당하다. 시간적 주파수에 대한 감도는 움직임 벡터의 크기를 소속함수로 유도하여 이용하였다. 그러나 색성분에 대한 시각적 감도는 명암에 비해 감도가 낮은 것으로서 고려하지 않았다. 이는 대부분의 동영상 부호기에서 움직임 벡터를 명암만으로 추출하고 있음과 관계된다.

참고 문헌

- [1] A. Leone, A. Bellini, R. Guerrieri, "Fuzzy-controlled perceptual coding of videophone

- sequences," IEEE trans. on Fuzzy System Vol.5, No.2, pp.294-303, May 1997.
- [2] K. A. Wen, C.Y. Lu, M.C. Tsai, "The transform image codec based on fuzzy control and human visual system," IEEE trans. on Fuzzy Systems, Vol.3, No.3, pp.253-259, Aug. 1995.
- [3] ITU-T, Draft Recommendation H.263 (Video coding for low bit rate communication), Dec. 1995.
- [4] ITU-T SG15 WP15/1, Video codec test model, TMN5, Jan. 1995.
- [5] K.N. Ngan, D. Chai, A. Millin, "Very low bit rate video coding using H.263 coder," IEEE trans. on CAS Video Tech., Vol.6, No.3, pp.308-312, June 1996.
- [6] T. Wiegand, M. Lightstone, S.K. Mitra, etc, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," IEEE trans. on CAS Video Tech., Vol.6, No.2 Apl. 1996.
- [7] K. R. Rao and J. J. Hwang, Techniques and standards for image, video, and audio coding, Prentice-Hall, 1996.
- [8] The MathWorks, Fuzzy logic toolbox, ver. 5.2, 1998.

저자 소개



양근호(Keun-Ho Yang)

1991년 8월 전북대학교 전자공학과, 석사

1999년 8월 전북대학교 전자공학과, 박사

1995년 3월~현재 한려대학교 멀티미디어정보통신공학과

※ 관심분야 : 영상신호처리 및 영상통신, 의료영상 처리, 디지털신호처리, DSP 응용