

퍼지 추론을 이용한 H.263 양자화 및 비율제어

허진원*, 신경철**, 최귀열*** 이광형*

송실대학교 정보통신전자공학부*, 경남정보대학 전자정보과**, 재능대학 정보통신과***

On a Quantization and Rate-Control in H.263 Video Coder using Fuzzy Reasoning

Jin-Won Hur*, Kyung-Chul Shin**, Gwi-Yeol Choi, Kwang-Hyung Lee*

Dept. of Telecommunication & Electronics Engr., Soongsil Univ.*,

Dept. Electronics & Information, KyungNam Information College**

Dept. Information & Communication, Jaeneung College**

jwhur@dsp.soongsil.ac.kr

요약문

H.263의 시험모델인 TMN5를 최대한 적용하여 실험하였으며 분산, 엔트로피, 움직임 크기 등의 퍼지변수를 데이터 영역에서 추출하여 퍼지화하였다. 소속함수를 계산하기 위해 최소값으로 가장 분명한 퍼지값을 추출하였으며 퍼지집합을 위해서는 각 소속함수로부터의 요소를 더하는 의미에서 최대값을 선택하였다. 무게중심기법을 이용하여 최종 퍼지값도를 구하여 TMN5에 부가하였다.

1. 서론

멀티미디어기술의 발전은 과거에는 생각하지도 못하였던 새로운 분야들을 많이 열어 놓았다. 문자, 도형, 음성, 정지영상뿐만 아니라 동영상까지 다양한 표현형식을 갖는 여러 정보들은 디지털 통신, 반도체, 컴퓨터 등의 기술 발전에 힘입어 주로 영상과 음성을 중심으로 결합된 멀티미디어의 형태로 저장 또는 전송되어진다. 그러나 오늘날 우리 인간이 접하는 모든 자연, 사회 현상은 갈수록 복잡해지고 있고, 컴퓨터 기술의 발달에 힘입어 이 복잡한 환경에서 발생하는 많은 문제를 해결하고 있지만, 이는 아직도 매우 제한되어 있다. 따라서 복잡한 현상의 문제를 인간이 해결한다는 것은, 대부분이 복잡한 문제를 단순화시켜, 우리가 이해할 수 있을 정도의 간단한 문제로 만든 다음 그 문제를 해결하는 것이다. 이러한 인간의 이해 할 수 없는 많은 현상으로 인해 이러한 문제를 해결 할 수 있는 이론의 바탕으로 나오게 된 것이 퍼지이론이다. 즉, 결정론적이지 않고

불확실한 특성을 갖는 여러 응용분야에서 최적의 상태 결정을 수학적으로 다룰 수 있도록 하는 수단을 제공하는데, 여러 변수로부터 하나의 결론을 추론하는 점에서 퍼지 시스템은 주로 제어 시스템 분야에서 응용되어져 왔다. 영상의 경우에 어떠한 threshold를 선택하여 특정부분에 대한 화질 개선등을 할 경우에 퍼지 추론이 이용될 수 있다. 이 중에서 H.261보다 상대적으로 적은 비트율을 발생시키고, 64kbps이하의 전송속도를 가지는 저전송률 부호화를 위한 H.263에서 기존의 DCT 기반의 TMN5 부호기 및 복호기에서는 화면내 부호화된 DC(INTRADC) 계수값을 위한 1개의 양자화기와 화면내 부호화된 DC값을 제외한(Non-INTRADC) 계수값을 위한 31개의 균일 양자화기를 가지고 있다. 그러나 이러한 양자화기로 양자화를 하게되면 비트율의 손상과 함께 만족스럽지 못한 화질의 제공으로 인해 영상의 여러 성능 개선을 위해 퍼지 논리가 적용될 수 있다. 특정한 매개변수를 변화시켜 어느 정도의 성능개선을 이룰 수는 있지만 그에 따른 다른 부분에서의 성능변화가 일어나게 되는데 이에 따라 퍼지를 이용한 퍼지추론이 적용될 수 있는 것이다. 이에 퍼지이론을 이용한 양자화와 비율제어를 통한 성능 개선을 제안한다.

부호기는 영상전화기에 응용될 수 있도록 설계된 H.263이며 64Kbps 이하에서 실험되었다. 양자화 변수를 제어하며 매크로블록단위의 비율제어 기법을 유지하면서 퍼지함수를 이에 적용하였다. 주어진 비트를 최대한 사용하고 프레임내 국부 특성에 따라 최적으로 비트를 할당하여 주로 움직이는 부분에서 오차가 크게 발생하고 시각적 감도도 높기 때문에 이 점을 고려하였다.

공간영역과 시간영역에서 퍼지 감도를 적용하여 프레임 단위의 오차보다 프레임내 및 프레임간 감각적 화질을 유지하는 데에 연구의 주안점이 있다. 이를 위해 시각 특성에 적합한 소속함수를 설계하고 퍼지화하여 H.263의 QP를 제어하고 버퍼상태를 제어하였다.

2. H.263에서 퍼지 시각적 소속함수의 결정

2.1 H.263/H.263+에서의 비율제어

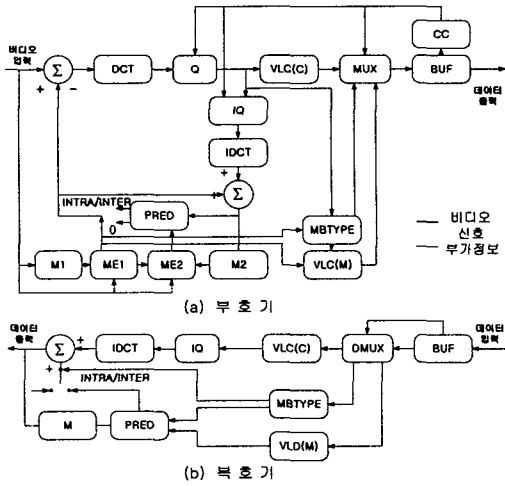


그림 1. H.263 부호기(a)와 복호기(b) 블록도

H.263의 테스트 모델인 TMN5의 비율제어는 두가지의 주요한 단계로 구성된다. 첫째로, 프레임을 부호화한 후 스킵의 수를 결정하는 것이고, 두 번째는 각각의 매크로블록에 대하여 양자화 매개변수를 결정하는 일이다. 각각의 프레임이 부호화되는동안, 프레임단위로 각각의 매크로블록의 양자화 크기를 결정하는 것이다. TMN5에서는 먼저 부호화된 프레임의 평균 양자화 스텝 크기 QP_{prev} 를 기초로 양자화 스텝 크기를 결정한다.

첫 번째로, G_{adj} 는 아래와 같다.

$$G_{adj} = \frac{B_{n-1} - f_{adj} - b_{target}}{2 \cdot B_{target}} \quad (1)$$

두 번째로, L_{adj} 는 다음과 같다.

$$L_{adj} = \frac{12 \cdot (B_{n,k} - \frac{k}{MB} B_{target})}{R} \quad (2)$$

여기서 k 는 매크로 블록의 수이며, 매크로블록 라인에서 각 매크로블록은 같은 양자화 크기를 사용하여 계산된다.

$$QP_{n,k} = QP_{prev} (1 + G_{adj} + L_{adj}) \quad (3)$$

화면내에서의 양자화 스텝 크기의 각각의 조정은 2로 제한된다.

$$QP_{n,k} = QP_{n,k-1} = \pm \max(2, |QP_{n,k} - QP_{n,k-1}|) \quad (4)$$

여기서 $QP_{n,k-1}$ 은 조정후 사용된 마지막 양자화 스텝 크기이다.

2.2 시각적 소속함수의 결정

적용적인 양자화 부호화에 관하여 crisp한 분류 방법을 사용하고, 퍼지 시스템은 각 매크로블록의 감도 인자에 대하여 연속적인 값을 만드는데 사용한다.

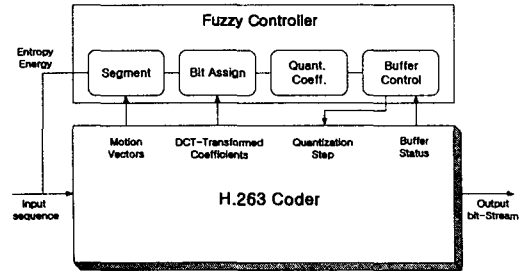


그림 2. 퍼지 제어기 구조

시각적 감도는 명암 및 주파수 특성에 주로 의존하므로 이에 근거하여 변수를 설정하고 소속함수를 결정한다.

첫 번째로, 에너지 E_k 는 다음과 같이 정의된다.

$$E_k = \frac{1}{N^2} \sum_{(i,j) \in MB_k} (f_{i,j} - \bar{f}_k)^2 \quad (5)$$

여기서 N 은 매크로블록의 크기로서 16의 값이 적용되며, $f_{i,j}$ 는 픽셀 (i,j) 의 휘도값이며, \bar{f}_k 는 같은 매크로블록내 화소의 평균휘도값이다. 이를 이용한 소속함수를 그림 4(a)에 나타냈다.

두 번째로, 엔트로피 H_k 는 아래와 같이 정의된다.

$$H_k = - \sum_{i=0}^{L_{MAX}} p_i \log_2 p_i \quad (6)$$

여기서 p_i 는 매크로블록내 발생 확률이며, L_{MAX} 는 최대 휘도값, 즉 255이다. 엔트로피가 주파수특성과 관계한다는 것은 다음과 같이 설명된다. 주파수가 낮다는 것은 블록내 분포가 임펄스 분포 즉 변화방향이 적음을 의미하며 최대 엔트로피는 균일분포 즉 매크로블록의 크기가 16×16 이며 명암의 변화가 $0 \sim 255$ 이므로 모두 한번씩 발생함을 의미한다. 이 경우 공간주파수는 최대가 되며 관측거리와 샘플링 비율을 고려할 때 공간주파수는 약 40 cyc/deg에 해당한다. 이러한 특성을 그림 3 b)와 같이 소속함수 결정에 이용한다. 동영상에 대하여 고려해야 할 세 번째 특성은 시간적 주파수인데, 이것은 부호기에서 움직임보상에 의해 충분히 고려되고 있다. 다만 움직임의 크기에 따라 움직임보상 오차가 달라지고 감도에 차이가 발생하므로 이것을 또 하나의 퍼지 변수로 삼는다. 현재 동작에 대한 활동도는 다음과 같이 정의된다.

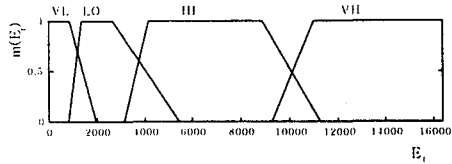
$$MV_k = u_k^2 + v_k^2 \quad (7)$$

여기서 u_k^2 와 v_k^2 은 매크로 블록 k 의 수직과 수평 이동 벡터를 가리킨다.

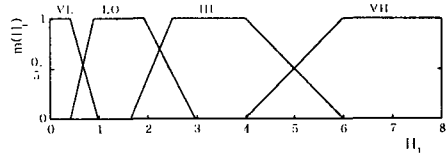
네 번째로 고려할 특성은 이전의 움직임에 대한 활동도를 고려하는 것이다.

$$PMV_k = \overline{u_k^2} + \overline{v_k^2} \quad (8)$$

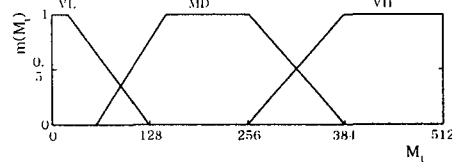
에너지와 엔트로피는 공간영역에서 인간 시각시스템의 특징과 관련되어있다. 움직임 벡터는 신호의 시간적인 특성을 추정하는데 사용되고, DCT 계수는 신호의 스펙트럼 특성을 추정하는데 사용된다. 인간 시각 시스템은 공간-주파수 영역에서 대역 통과 필터, 시간-주파수 영역에서 저대역 통과 필터에 적용되어진다[2],[3]. 퍼지 규칙은 시간영역과 공간영역에서 독립적으로 정의된다. 이것은 퍼지 규칙 집합의 결과이고, 퍼지 연관 메모리와 같이 정의 될 수 있다[4] 결정된 FAM를 그림 4에 나타냈다.



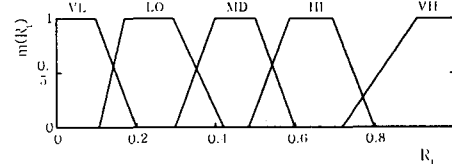
a)



b)



c)



d)

그림 3. 감각적 소속함수의 결정: a) 분산, b) 엔트로피, c) 움직임에 대한 소속함수와 d) 출력 감도 특성.

		분산				
		VL	LO	III	VH	
엔트로피	엔트로피	VL	MD	III	HI	MD
	로퍼	LO	HI	MD	MD	HI
	로퍼	HI	HI	MD	MD	HI
	로퍼	VH	MD	HI	HI	MD

		분산				
		VL	LO	III	VH	
엔트로피	엔트로피	VL	LO	MD	MD	LO
	로퍼	LO	MD	HI	HI	MD
	로퍼	HI	MD	HI	HI	MD
	로퍼	VH	LO	MD	MD	LO

그림 4. 움직임벡터 크기에 따른 분산, 엔트로피의 감도 FAM 맵

- 움직임이 적은 경우 ($M_i = VL, PM_i = VL$)
- 적당한 움직임인 경우 ($M_i = MD, PM_i = MD$)
- 움직임이 매우 큰 경우 ($M_i = PM_i = VH$),
- 움직임이 정지했거나 시작하는 경우 ($M_i = VL, PM_i = VH$ 또는 $M_i = VH, PM_i = VL$).

3. 퍼지감도를 이용한 양자화 및 비율제어

정해진 감도특성은 H.263의 양자화를 위한 양자화 변수 QP의 변화에 이용한다. QP는 1에서 31까지 변화하며 비퍼상태를 고려하여 전체적으로 및 국부적으로 제어된다.

제한한 시스템은 이전까지 사용되고 남은 전체 비트를 대상으로 가용 비트 B_{res} 를 최대한 사용토록 한다. 전체적 비트율 제어는 감도의 합 F_{sum} 을 매크로블록 수로 나눈 평균값으로 균준하여 다음과 같이 설정한다.

$$G_{adj} = -\frac{B_{res}}{2B} * (\frac{F_{sum}}{MB} + 0.5) \quad (9)$$

따라서 지금까지 사용하고 남은 비트를 가용 자원으로 하고 프레임 단위의 평균 감도에 의해 기준을 삼는 시각적 제어가 가능하게 된다. 프레임내 국부제어는 다음과 같이 적용된다.

$$L_{adj} = \frac{12\Delta_2 B}{R} * (F_i + 0.5) \quad (10)$$

$$\Delta_2 B = B_{i,mb} - \frac{F_{i,mb}}{F_{sum}} * B_{res} \quad (11)$$

여기서 F_i 는 i -번째 매크로블록에 대한 퍼지감도를, $F_{i,mb}$ 는 i -번째 매크로블록까지 누적된 퍼지감도를 의미한다. 즉 감도에 따라 양자화를 하며 비트를 분산시키는 효과가 있다. 최종 양자화를 위한 $QP_{n,k}$ 는 식 (3)과 같다.

양자화는 비트율 제어와 밀접한 관계가 있으며 최적 양자화는 최적 비트율 제어를 전제로 하고 있기 때문에 퍼지감도에 의한 QP 제어가 특정 매크로블록에 대해서

는 받아들여지지 않을 수도 있다. 즉, 가용 비트 자원이 없으면 감도에 의한 QP 제어는 무시된다. 그러나 허용 범위내에서 감도에 따라 가용 자원을 최대한 사용토록 한다.

IV. 실험 및 결과

이 논문에서 제안한 기법들은 H.263의 시험모델인 TMN5를 최대한 적용하였다. TMN5와 이 논문에서 적용한 기법들과 다른점은 매크로블록의 행 단위로 처리하는 QP 변화를 각 매크로블록 단위로 처리하도록 하였으며 프레임 건너뛴을 하지 않도록 하여 프레임 단위별로 두 시스템을 직접적으로 비교하였다. QP가 달라지면 건너뛰는 프레임이 달라지고 직접 비교가 불가능하기 때문이다. 일부 수정이 가해져도 TMN5의 성능은 저하되지 않고 적용 방법이 달라진다는 것을 전제로 하였다. 사용된 영상 시퀀스는 176×144 화소의 QCIF 영상으로 foreman, Claire, Carphone 등이다.

분산, 엔트로피, 움직임 크기 등의 퍼지변수를 데이터 영역에서 추출하여 퍼지화하였다. 퍼지화 과정에서는 min-max 기법 즉, 소속함수를 계산하기 위해 최소값으로 가장 분명한 퍼지값을 추출하였으며 퍼지집합을 위해서는 각 소속함수로부터의 요소를 더하는 의미에서 최대값을 선택하였다. 역퍼지화를 위해 무게중심 기법을 이용하여 최종 퍼지감도를 구하였다. 이것을 양자화 및 비율제어에 적용하였으며 TMN5에 추가하였다. 그림 5와 6은 TMN5와 적용한 퍼지의 SNR을 보인다.

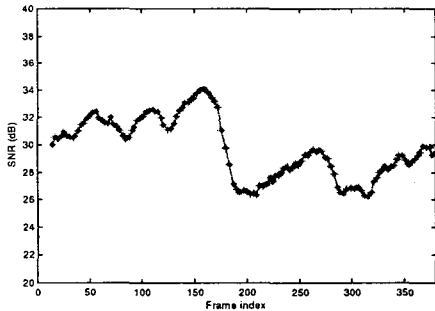


그림 5. TMN5의 SNR(carphone.qcf)

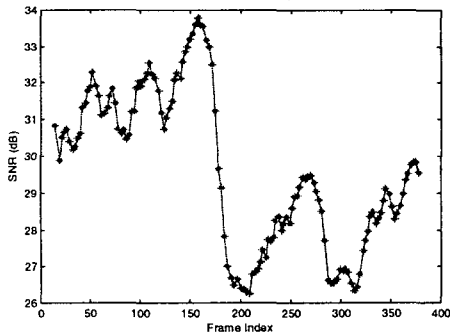


그림 6. 적용한 퍼지의 SNR(carphone.qcf)

V. 결론

64kbps 이하의 영상전화 및 회의에 이용되는 저진송률 동영상 부호기에 퍼지추론을 적용하는 기법을 제안하였다. 전송률이 낮아질수록 파형 분석에 의한 압축은 효과가 줄어들게 되며 블록현상이 커지게 된다. H.263이 수십 Kbps급의 전송률에 맞추어 설계된 것이지만 실제로 깨끗한 화질을 보이지 못하는 이유가 여기에 있다. 따라서 시각적인 감도의 분산이나 적절한 분배가 필요해진다. 이 논문에서는 시각적인 특성에 적합하게 퍼지변수를 결정하고 소속함수를 유도하였다. 소속함수의 결정은 출력 영상의 주관적 및 객관적 화질에 결정적인 영향을 미치므로 시각적 감도에 근거한 소속함수의 결정은 매우 의미있는 것이었다. 정지영상에 대한 시각 특성은 명암의 세기 및 공간주파수적으로 주어진다. 이것을 각각 분산 에너지와 엔트로피에 대응하여 적절한 소속함수를 유도하였다. 시간적 주파수에 대한 감도는 움직임벡터의 크기를 소속함수로 유도하여 이용하였다. 그러나 색성분에 대한 시각적 감도는 명암에 비해 감도가 낮은 것으로서 고려하지 않았다. 이는 대부분의 동영상 부호기에서 움직임벡터를 명암만으로 추출하고 있음과 관계된다.

소속함수의 퍼지감도는 min-max 기법에 의해 유도되었으며 역퍼지화는 무게중심기법을 사용하였다.

참고 문헌

- [1] ITU Telecom. Standardization Sector of ITU, "Video Coding for Low Bitrate Communication," ITU-T Recommendation H.263, March 1996
- [2] J.T. Kim, H. J. Lee, and J. S. Choi, "Subband coding using human visual characteristics for image signals," IEEE J. Select. Areas Commun., vol. 11, pp. 59-64, Jan. 1993
- [3] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second-generation image-coding techniques," Proc. IEEE, vol. 73, pp. 549-574, Apr. 1985.
- [4] B. Kosko, Neural Networks and Fuzzy Systems. Englewood Cliffs, NJ : Prentice-Hall, 1992.
- [5] Pohsiang Hsu and K. J. Ray Liu, "A Predictive H.263 bit-rate control scheme based on scene information", IEEE Int. Conf. ICME2000, New York, vol. III, pp. 1737-1736, July 2000