

H.264에서 간소화된 기법에 의한 왜곡치 예측

Simplified Approach for Distortion Estimation in H.264

박기홍*, 김윤호**

Ki-Hong Park*, Yoon-Ho Kim**

요 약

본 논문은 H.264에서 모드 결정을 위한 간소화된 왜곡치 예측 방법을 소개하였다. 왜곡치 계산은 양자화된 변환 계수와 역양자화된 변환 계수의 차이로 계산되는데, 일반적으로 이 과정은 DCT 변환, 양자화, 역양자화 및 역 DCT 변환이 수행되어야 한다. 제안하는 방식에서는 왜곡치를 계산하기 위하여 일련의 간소화된 수식을 사용함으로써 역양자화 및 역 DCT 과정을 생략하였다. 실험결과, PSNR은 거의 일치하면서도, RDO 모드 결정 시간은 기존의 방식보다 8~15 %의 감소를 보였다.

Abstract

This paper addressed an another scheme of distortion estimation method based on simplified inverse quantization in H.264/AVC. The distortion is calculated by the difference of coefficient between quantized transform coefficients and that of inverse. In general, these process included such transforms as discrete cosine transform(DCT), quantization, inverse quantization(IQ), and Inverse DCT(IDCT). In proposed approach, IQ as well as IDCT process are skipped because of replacing a couple of approximated formulas. Some simulation have been conducted and it showed that the PSNR was almost the same, and reduced the rate-distortion optimization(RDO) mode decision time of 8~15% in comparison with conventional method.

Key words : H.264, Bit-Rate, Distortion Estimation, RDO

I. 서 론

H.264에서 압축률 향상을 위해 유효-왜곡 최적화된 모드 결정 방식을 사용한다. RDO 모드 결정(RDO_MD, RDO Mode Decision) 방식은 움직임 예측 과정에서 계산된 움직임 비용으로 모드를 결정하지 않고 비트율과 왜곡치를 계산하여 유효-왜곡 비용을 최소로 가지는 매크로블록 모드를 선택하는 것이다 [1]. H.264에서 이 방식은 비트율 감소의 성능을 가지지

만 전체 부호화 시간을 증가시킨다 [2][3]. 따라서 압축률을 높이기 위한 방식들은 부호화 계산량이 증가되고, 이에 매크로블록 모드의 움직임 예측과 모드 결정에 대한 계산량을 줄이기 위한 연구가 활발히 이루어져 왔다.

Lim[4]은 블록의 경계 값을 이용하여 블록의 균일성을 이용한 움직임 예측과 모드 결정 방식을 제안하였고, Kim[5]은 양자화된 DCT 계수로부터 후보 블록의 생략 여부를 결정하였다. Tu[6]는 DCT 영역에서

* 목원대학교 일반대학원 IT공학과 박사과정

** 목원대학교 컴퓨터공학부 교수

· 제1저자 (First Author) : 박기홍

· 교신저자 (Corresponding Author) : 김윤호

· 투고일자 : 2010년 6월 8일

· 심사(수정)일자 : 2010년 6월 9일 (수정일자 : 2010년 6월 25일)

· 게재일자 : 2010년 6월 30일

양자화 계수들로부터 비트율을 예측하고, DCT 계수와 역양자화된 계수들로부터 왜곡치를 예측함으로써 엔트로피 부호화, IDCT, 역양자화에 과정에서 요구되는 계산량을 감소시켰다. 또한, Moon[7]은 양자화와 역양자화를 하지 않고 양자화 오차의 특성을 분석하여 통계적으로 왜곡치를 예측하였고, Po[8]는 반복적인 룩업 테이블을 작성하여 양자화와 역양자화한 후 왜곡치를 측정하였다. 본 논문에서는 왜곡치 계산을 위한 과정에서 필요한 역양자화 과정을 생략하는 방법을 제안한다.

II. H.264에서의 왜곡치 예측

2-1 H.264의 RDO 모드 결정

H.264에서 RDO 모드 결정 방식은 식 1과 같은 윌-왜곡 비용이 최소가 되는 모드를 선택한다[5].

$$J_{Mode} = D_{Mode} + \lambda_{Motion} R_{Mode} \quad (1)$$

식 1에서 움직임 예측 후에 각 모드에 대해 D_{Mode} 는 해당 모드에 대한 원 블록과 복원 블록의 차의 SSD(Sum of Squared Differences)이고, λ_{Motion} 는 라그랑지 승수이다. R_{Mode} 는 헤더 정보에 해당하는 모드, 움직임 벡터, 양자화 변수 정보, 오차 블록의 비트 수를 포함한다. 그림 1은 H.264/AVC의 참조 소프트웨어 [9]에서 비트율과 왜곡치를 계산하는 과정을 나타낸다.

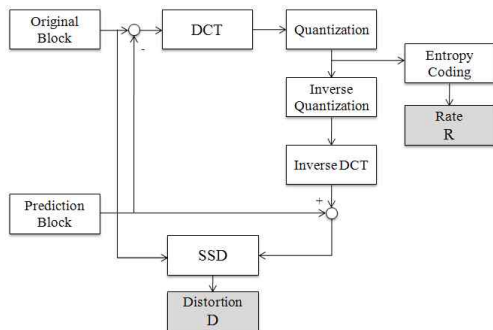


그림 1. H.264/AVC 참조 프로그램에서의 RDO 모드 결정을 위한 윌-왜곡 비용 계산

Fig. 1. Cost calculation for RDO mode decision in H.264/AVC reference software

복원 블록을 얻기 위해서는 부호화 과정인 원 블록과 움직임 예측 블록의 차인 예측 오차 블록의 DCT, 양자화, 역양자화 및 IDCT 과정이 수행된다. 따라서 복원 오차로부터 계산되는 D_{Mode} 는 DCT 영역에서 양자화 오차로 표현된다.

식 2와 식 3은 H.264에서 4×4 크기의 오차 블록 U 를 DCT 변환 계수 블록 V 를 구하는 과정과 그 역변환 과정을 나타낸다.

$$V = (C_f U C_f^T) \otimes E_f \quad (2)$$

$$\text{here, } C_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}, E_f = \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & b^2 & \frac{ab}{2} & b^2 \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & b^2 & \frac{ab}{2} & b^2 \end{bmatrix}$$

$$U = C_i^T (V \otimes E_i) C_i \quad (3)$$

$$\text{here, } C_i = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{bmatrix}, E_i = \begin{bmatrix} a^2 ab a^2 ab \\ ab b^2 ab b^2 \\ a^2 ab a^2 ab \\ ab b^2 ab b^2 \end{bmatrix}$$

여기서 \otimes 는 행렬의 각 요소별 곱하기 연산이다. 식 2에서 post-스케일링 하지 않은 정수 DCT 계수 행렬을 $Z (= C_f U C_f^T)$ 이라고 할 때, (k, l) 번째 위치의 양자화된 DCT 계수 V_q 는 다음 식과 같다.

$$V_q(k, l) = \begin{cases} \left\lceil \frac{A(k, l) + d}{\Delta} \right\rceil, & \text{if } Z(k, l) \geq 0 \\ \left\lfloor \frac{A(k, l) - d}{\Delta} \right\rfloor, & \text{else} \end{cases} \quad (4)$$

here, $A(k, l) = Z(k, l) E(k, l)$

여기서 $\lceil \cdot \rceil, \lfloor \cdot \rfloor, \Delta$ 는 각각 소수점 버림 연산, 양자화 크기를 나타내고, d 는 QP(Quantization Parameter) 값에 따라 인트라 블록일 때는 $\Delta/3$, 인터 블록일 때는 $\Delta/6$ 값을 가진다. H.264에서 DCT 계수 $V(k, l)$ 은 양자화 간격 크기(Δ (QP))에 의해 양자화 되는데, 52개의 양자화 파라미터에 대해서 일반적으로 다음과 같이 정의된다.

$$\Delta(QP) = \delta(r)2^q$$

$$\text{here, } \delta(i) = \begin{cases} \frac{10}{16}, \frac{11}{16}, \frac{13}{16}, \frac{16}{16}, \frac{18}{16} \end{cases} \quad (5)$$

$$\text{when } (0 \leq i \leq 5)$$

여기서 r 과 q 는 QP를 6으로 나눈 나머지와 몫을 나타낸다. $V_q(k,l)$ 를 역양자화한 계수를 $\widehat{V}_q(k,l)$ 라고 할 때, IDCT 후의 오차 블록은 식 6과 같다.

$$\widehat{U} = C_i^T(\widehat{V} \otimes E_i)C_i \frac{1}{64} \quad (6)$$

여기서 C_i 와 E_i 는 IDCT 행렬과 pre-스케일링 변환 행렬을 나타내고, 역 변환의 출력값은 스케일링 계수를 제거하기 위해 64로 나누어진다.

2-2 기존의 왜곡치 예측

Tu는 post-스케일링 변환을 하지 않은 정수 DCT 계수 행렬을 Z 와 역양자화된 \widehat{Z} 를 이용하여 왜곡치를 계산하였고, $\widehat{Z}(k,l)$ 을 다음과 같이 정의하였다.

$$\widehat{Z}(k,l) = Z(k,l)M(k,l) \frac{1}{64}$$

$$\text{here, } M(k,l) = \begin{cases} 16, & \text{when } (k,l) \in G0 \\ 25, & \text{when } (k,l) \in G1 \\ 20, & \text{when } (k,l) \in G2 \end{cases} \quad (7)$$

여기서 $G0$, $G1$ 및 $G2$ 다음과 같이 4×4 블록 내의 위치들에 대한 그룹을 나타낸다.

$$\begin{cases} G0, & \text{when } (k,l) \in \{(0,0), (0,2), (2,0), (2,2)\} \\ G1, & \text{when } (k,l) \in \{(1,1), (1,1), (3,1), (3,3)\} \\ G2, & \text{when } (k,l) \in \{\text{나머지 위치}\} \end{cases} \quad (8)$$

따라서 Tu의 4×4 블록에 대한 DCT 영역의 왜곡치는 식 9와 같이 계산된다.

$$D_{Tu} = \sum_{k=0}^3 \sum_{l=0}^3 \{(Z(k,l) - \widehat{Z}(k,l))\}^2 E_f(k,l)^2$$

$$\text{here, } E_f(k,l) = \begin{cases} a^2, & \text{when } (k,l) \in G0 \\ b^2/4, & \text{when } (k,l) \in G1 \\ ab/2, & \text{when } (k,l) \in G2 \end{cases} \quad (9)$$

여기서 post-스케일링 변환 계수 $E_f(k,l)$ 의 각 요소 a 와 b 는 $1/2$ 와 $\sqrt{2/5}$ 이다.

Po는 양자화와 역양자화 과정의 곱셈 연산량을 줄이기 위해서 반복적인 테이블 룩업 방식을 제안하였고, Po의 양자화 간격 크기는 다음과 같이 정의하였다.

$$\Delta_{Po}(QP, k, l) = \Delta QP(k, l) / E_f(k, l) \quad (10)$$

2-3 제안하는 왜곡치 예측

본 논문에서 제안하는 율-왜곡 예측 방식을 설명하기 앞서 H.264에서의 양자화를 살펴보면 식 4에서의 나눗셈 연산을 없앤 식 11과 같이 정의된다.

$$V_q(k,l) = \{|Z(k,l)MF(r,k,l) + f| \gg \text{qbits}\}$$

$$\text{here, } MF(r,k,l) \approx \frac{E_f(k,l)}{\Delta(QP)} 2^{q+15} \quad (11)$$

$$\text{qbits} = 15 + \text{floor}(QP/6)$$

여기서 \gg 는 오른쪽 이진 쉬프트 연산이고, f 는 인트라 블록에서 $2^{\text{qbits}}/3$ 이고, 인터 블록에서는 $2^{\text{qbits}}/6$ 이다. 역양자화는 다음과 같다.

$$\widehat{Z}(k,l) = \{V_q(k,l)RF(r,k,l)\} \ll \text{floor}(QP/6)$$

$$\text{here, } RF(r,k,l) \approx \Delta(QP) \frac{1}{2^{\text{floor}(QP/6)}} E_i(k,l) 2^6$$

$$E_i(k,l) = \begin{cases} a^2, & \text{when } (k,l) \in G0 \\ b^2, & \text{when } (k,l) \in G1 \\ ab, & \text{when } (k,l) \in G2 \end{cases} \quad (12)$$

일반적인 왜곡치 계산은 양자화된 DCT 계수(V_q)와 역양자화된 DCT 계수(\widehat{V}_q)의 차이로 계산되는데, 이 과정은 DCT 변환, 양자화, 역양자화 및 IDCT 변환이 수행되어야 한다.

본 논문에서는 왜곡치를 계산하기 위해 수행되어야 했던 일련의 과정을 간소화 하였다. 즉 IDCT 변환된 계수(\widehat{V}_q)를 $V_q \Delta(QP)$ 로 재 정리 하였고, 양자화 간격 크기($\Delta_{PM}(QP, k, l)$)를 다음과 같이 정의 하였다.

$$\begin{aligned}
& \Delta_{PM}(QP, k, l) \\
& = E_f(k, l) Q_s \\
& = \begin{cases} a^2 S & , \text{when } (k, l) \in G0 \\ (b^2/4) S & , \text{when } (k, l) \in G1 \\ (ab/2) S & , \text{when } (k, l) \in G2 \end{cases} \quad (13) \\
& \text{here, } Q_s = \frac{2^{qbits}}{MF(r, k, l)}
\end{aligned}$$

식 11과 식 13을 이용하여 왜곡치를 계산하는 과정은 식 14와 같다.

$$\begin{aligned}
D_{PM} & = \sum_{k=0}^3 \sum_{l=0}^3 \{(Z(k, l) E_f(k, l) - V_q(k, l) \Delta_{PM}(QP, k, l))\}^2 \\
& = \sum_{k=0}^3 \sum_{l=0}^3 \left\{ (Z(k, l) MF(r, k, l) \Delta_{PM}(QP, k, l) \frac{1}{2^{qbits}} - V_q(k, l) \Delta_{PM}(QP, k, l)) \right\}^2 \\
& = \sum_{k=0}^3 \sum_{l=0}^3 \left\{ (Z(k, l) MF(r, k, l) \frac{1}{2^{qbits}} - V_q(k, l) \Delta_{PM}(QP, k, l)) \right\}^2 \quad (14)
\end{aligned}$$

III. 실험 결과 및 고찰

본 본문에서 제안하는 방법은 윈도우 기반 팬티엄-IV, CPU 3.0GHz, 2GB RAM에서 수행되었고, 성능을 검증하기 위해 참조 소프트웨어를 이용하였다.

표 1. PSNR 비교

Table 1. Comparison of PSNR

JM ver.	12.2
Test sequence	QCIF, CIF
Profile	Baseline
Frame Num.	100
GOP	IPPP...
Entropy coding	CAVLC
RDO	On
Motion estimation	On
QP	28, 32, 36, 40
Search range	16

실험 환경은 표 1과 같고 P 프레임에서는 인터 예측만을 사용하였고, 왜곡치 예측은 휘도 블록에만 적용하였다. 실험에 사용된 비디오 시퀀스는 그림 2과

같이 QCIF(176×144) 시퀀스는 'Foreman', 'Silent'를 사용하였고, CIF(352×288)는 'Akiyo', 'Coastguard'를 사용하였다.

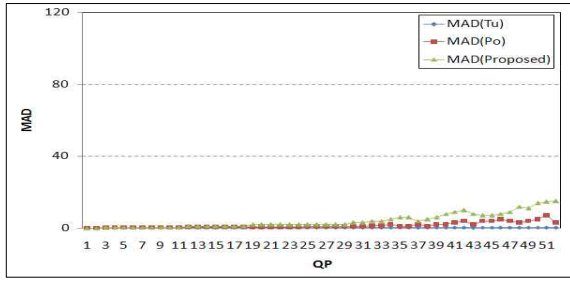


그림 2. 비디오 시퀀스
Fig. 2. Video sequence

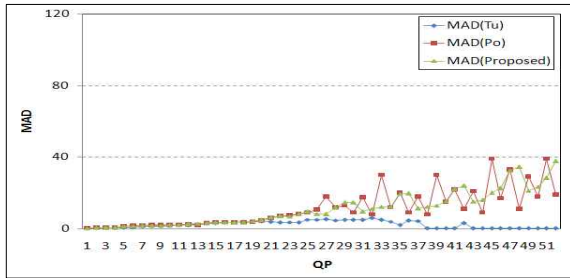
본 연구에서 계산된 왜곡치의 정확성을 검증하기 위해 참조 소프트웨어에서의 재구성된 왜곡치(D_{JM})와 4×4 블록들의 예측된 왜곡치간의 MAD(Mean Absolute Differences)를 비교한다. 기존 및 제안하는 방법의 MAD는 식 15와 같다.

$$\begin{aligned}
MAD_{Proposed} & = \frac{1}{N} \sum_{i=0}^{N-1} |D_{JM_i} - D_{Proposed_i}| \\
MAD_{Tu} & = \frac{1}{N} \sum_{i=0}^{N-1} |D_{JM_i} - D_{Tu_i}| \quad (15) \\
MAD_{Pb} & = \frac{1}{N} \sum_{i=0}^{N-1} |D_{JM_i} - D_{Pb_i}|
\end{aligned}$$

그림 3에서 'Akiyo' 시퀀스는 움직임이 거의 없는 영상으로 MAD가 고르게 증가하는 반면에 'Coastguard' 시퀀스는 움직이는 많은 영상으로 MAD 그래프가 변화가 심하는 것을 확인 할 수 있다. 즉, DCT 변환, 양자화, 역양자화 및 IDCT 변환 과정을 모두 수행하는 Tu의 방법의 MAD 그래프에 비해 제안하는 방법은 역양자화 과정을 생략함으로써 약간의 오차 범위를 허용하여 균일하지 않는 MAD 그래프를 보였다.



(a) Akiyo (CIF, 10Hz)



(b) Coastguard (CIF, 10Hz)

그림 3. CIF 시퀀스의 MAD 그래프
Fig. 3. MAD graph of CIF sequence

표 2. PSNR 비교

Table 2. Comparison of PSNR

Sequence	Hz	PSNR		
		Tu's	Po's	Proposed
Foreman (QCIF)	10	-0.023	-0.022	-0.068
	30	0.007	0.009	0.004
Silent (QCIF)	15	0.005	0.004	-0.023
	30	0.018	0.042	0.015
Akiyo (CIF)	10	0.004	0.017	0.009
	30	0.007	0.006	0.005
Coastguard (CIF)	10	-0.013	0.005	-0.012
	30	-0.004	-0.003	-0.016

기존의 방식과 제안하는 방식의 성능이 표 1에서와 같이 PSNR이 거의 일치함을 알 수 있고, 음수는 PSNR의 감소를 의미한다. 기존 및 제안 방법을 참조 소프트웨어의 RDO_MD에 적용했을 때의 부호화 시간 감소량은 식 16과 같다.

$$\begin{aligned}
 & Save\ Time_{RDO_MD}(\%) \\
 &= \frac{Time_{RDO_MD} - Time_{with\ Estimated\ Distortion}}{Time_{RDO_MD} - Time_{with\ jout\ RDO_MD}} \times 100
 \end{aligned}
 \tag{16}$$

여기서 $Time_{RDO_MD}$ 와 $Time_{with\ jout\ RDO_MD}$ 는 참

조 소프트웨어에서 RDO_MD를 사용하는 경우와 사용하지 않는 경우의 전체 부호화 시간을 나타낸다. 또한 $Time_{with\ Estimated\ Distortion}$ 는 기존의 방법과 제안 방법으로 예측한 왜곡치를 RDO_MD에 적용한 경우의 전체 부호화 시간을 나타낸다 표 2는 QP가 28, 32, 36인 경우의 기존의 방법과 제안 방법의 RDO 모드 결정의 시간 감소량을 나타낸다.

표 3. RDO_MD의 시간 감소량

Table 3. Time reduction of RDO_MD

Sequence	Hz	Method	$Save\ Time_{RDO_MD}(\%)$		
			QP 28	QP 32	QP 36
Foreman (QCIF)	10	Tu's	10.67	17.74	16.57
		Po's	15.00	24.59	26.90
		Proposed	25.67	32.03	32.30
	30	Tu's	11.79	19.24	11.77
		Po's	22.76	25.09	22.45
		Proposed	26.13	34.84	31.96
Silent (QCIF)	15	Tu's	14.87	14.88	12.43
		Po's	21.75	25.06	24.08
		Proposed	33.49	27.33	27.91
	30	Tu's	14.29	22.42	19.70
		Po's	26.14	29.63	25.86
		Proposed	35.04	27.03	30.79
Akiyo (CIF)	10	Tu's	22.56	19.10	22.10
		Po's	27.39	23.06	25.33
		Proposed	37.78	33.46	33.57
	30	Tu's	21.37	19.92	19.91
		Po's	25.08	26.04	24.05
		Proposed	34.37	32.56	33.77
Coastguard (CIF)	10	Tu's	7.64	11.65	12.24
		Po's	14.04	19.14	19.27
		Proposed	23.69	27.47	27.85
	30	Tu's	11.10	12.38	14.48
		Po's	18.30	19.31	19.70
		Proposed	24.34	25.67	27.95

IV. 결 론

본 논문에서는 H.264에서 모드 결정을 위한 왜곡치 예측 방법을 제안하였다. 왜곡치 계산은 양자화된 변환 계수와 역양자화된 변환 계수의 차이로 계산되는데, 이 과정은 DCT 변환, 양자화, 역양자화 및 IDCT 변환이 수행되어야 한다. 제안하는 왜곡치 예측 방식에서는 왜곡치를 계산하기 위해 수행되어야 했던 역양자화 과정을 생략하였다. 실험결과,

영상의 PSNR이 거의 일치하면서, 기존의 방식보다 8~15 %의 RDO 모드 결정의 시간 감소를 보였다.

참 고 문 헌

[1] "Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods", *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Document JVT-K049, Munich, Germany*, Mar. 2004.

[2] J. Ostermann, J. Bormans, P. List, D. Marpe. M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC:Tools, Performance, and Complexity", *IEEE Curcuits and Systems Magazine, First Quarter*, pp. 7-28, 2004.

[3] Lain E.G. Richardson, H.264 and MPEG-4 Video Compression, John Wiley & Sons Inc. 2003.

[4] K. P. Lim, S. Wu, D. J. Wu, S. Rahardja, X. Lin, F. Pan, S. G. Li, "Fast Inter Mode Selection", *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Document JVT-I020, San Diego, US*, 2-5 Sept. 2003.

[4] Y. H. Kim, J. W. Yoo, S. W. Lee, J. Shin, J. Paik, and H. K. Jung, "Adaptive mode decision for H.264 encoder", *IEEE Electronics Letters, Vol. 40, No. 19*, Sep. 2004.

[6] Y. K. Tu, J. F. Yang, and M. T. Sun, "Efficient Rate-Distortion Estimation for H.264/AVC Coder", *IEEE Trans. Circuits Syst. Video Technol., vol 16, no. 5*, pp. 600-611, May 2006.

[7] J. M. Moon and J. H. Kim, "A Rate-Distortion Estimation Method for Fast Mode Decsion in H.264", *Journal of KICS, vol. 33, no. 1*, pp. 34-44, Jan. 2008.

[8] I. Pao and M. Sun, "Modeling DCT Coefficients for Fast Video Encoding", *IEEE Trans. Circuits Syst. Video Technol., vol 9, no. 4*, pp. 608-616, Jun. 1999.

[9] Joint Video Team (JVT) Reference Software. Available: <http://bs.hhi.de/~suehring/tml/download/>

박 기 홍(朴起弘)



2004년 2월 : 목원대학교 컴퓨터공학과 (공학사)
2006년 2월 : 목원대학교 공학석사
2006년 8월~현재 : 목원대학교 IT 공학부 박사과정
관심분야 : 영상처리, 컴퓨터비전

김 윤 호 (金允鎬)



2003 - 현재 : 목원대학교 컴퓨터 공학부 교수
2005-2006: University of Auckland. NZ, CITR Lab Research Fellow.
2008-현재: ISO/TC223 Societal Security Korea Delegate.

IEEE, 대한전자공학회, 한국통신학회 정회원/ 한국정보 기술학회, 한국해양정보통신학회, 한국 모바일학회, 한국항행학회, 사회안전학회 종신회원.
관심분야 : 영상처리, 컴퓨터비전, 사회안전 표준화, IT정책, 방재정보통신 등.