

움직임 벡터의 상관도와 참조 화면의 거리를 이용한 H.264/AVC 움직임 탐색 지점 제거

정회원 문지희*, 최정아**, 종신회원 호요성***

Removal of Search Point using Motion Vector Correlation and Distance between Reference Frames in H.264/AVC

Ji-Hee Moon*, Jung-Ah Choi** *Regular Members*, and Yo-Sung Ho*** *Lifelong Member*

요약

본 논문에서는 움직임 벡터의 상관도와 참조 화면의 거리를 이용한 H.264/AVC 움직임 탐색 지점 제거 방법을 제안한다. 결정된 탐색 영역을 이용하여 전 영역 탐색 방법에서 정화소 단위의 탐색 지점을 제거하고 향상된 예측 지역 탐색 방법에서는 불필요한 예측 움직임 벡터를 제거한다. 또한, 현재 화면에서 멀리 있는 참조 화면은 최적의 참조 화면으로 선택되는 빈도가 낮다는 특성을 이용하여 현재 화면과 참조 화면 사이의 거리를 이용한 가중평 균을 적용하여 거리가 멀어질수록 전 영역 탐색 방법의 탐색 영역의 크기를 줄여 움직임 예측에 소요되는 시간을 줄인다. 향상된 예측 지역 탐색 방법의 경우, 멀리 있는 참조 화면에서 개선 과정을 적응적으로 생략하여 움직임 예측 시간을 줄인다. 제안한 방법을 통해 전 영역 탐색 방법은 평균 57.13%의 움직임 예측 시간이 감소되었고, 향상된 예측 지역 탐색 방법은 평균 14.71%의 움직임 예측 시간이 감소되었다.

Key Words : H.264/AVC, motion estimation, enhanced predictive zonal search, full search, motion vector

ABSTRACT

In this paper, we propose the removal of search point using motion vector correlation and distance between reference frames in H.264/AVC. We remove the search points in full search method and predictive motion vectors in enhanced predictive zonal search method. Since the probability that the reference frame far from the current frame is selected as the best reference frame is decreased, we apply the weighted average based on distance between the current and reference frame to determine the final search range. In general, the size of search range is smaller than initial search range. We reduce motion estimation time using the final search range in full search method. Also, the refinement process is adaptively applied to each reference frame. The proposed methods reduce the computational throughput of full search method by 57.13% and of enhanced predictive zonal search by 14.71% without visible performance degradation.

I. 서론

H.264/AVC는 모바일 TV 환경에서 VCD급 품질을 목적으로 개발된 비디오 압축 표준으로 다양한 부호화 기술을 사용한다. 그 중 16×16부터 4×4까

지 다양한 블록 크기를 사용한 움직임 보상 기술은 이전 부호화 표준보다 압축 효율을 향상시킨 주요 기술 중 하나이다. 또한, 1/4 화소 단위의 움직임을 예측하여 이미 부호화된 이전 화면으로부터 움직임을 세밀하게 추정하였다. 좀더 정교한 움직임 예측

* LG전자(kumvi86@gmail.com) ** 광주과학기술원 실감방송연구센터({jachoi, hoyo}@gist.ac.kr), (°: 교신저자)
논문번호 : KICS2011-01-045, 접수일자 : 2011년 1월 22일, 최종논문접수일자 : 2012년 2월 13일

기술을 통해 잔여 영상의 데이터의 양을 줄일 수 있기 때문에 비디오 신호를 효율적으로 압축할 수 있다^[1].

최근 HD 및 UHD와 같은 고해상도, 고품질 영상에 대한 수요가 다양한 응용분야에서 급속히 증가하고 있다. 이러한 고해상도, 고품질 비디오는 데이터의 양이 방대하기 때문에 H.264/AVC의 압축 표준으로 전송 또는 저장에 어려움이 있다. 이러한 문제를 해결하기 위해 동화상 전문가 그룹(Moving Picture Experts Group, MPEG)과 동영상 전문가 그룹(Video Coding Experts Group, VCEG)은 2010년 1월에 협력팀(Joint Collaborative Team on Video Coding, JCT-VC)를 결성하고 공동으로 차세대 비디오 표준화(High Efficiency Video Coding, HEVC)를 시작하기로 합의하였다. HEVC는 H.264/AVC 압축 표준보다 2배 이상의 압축률 향상을 목표로 하고 있다^[2].

HEVC에 제안된 기술 중 기존의 16×16 화소 단위의 매크로블록보다 확장된 크기의 매크로블록을 사용한 부호화 방법은 고해상도 영상 부호화에서 뛰어난 성능을 보인다^[3]. 그러나 확장된 크기의 블록을 이용한 움직임 예측 방법은 부호화 효율을 향상시키지만, 최적의 움직임 벡터를 찾기 위해 계산해야 할 화소의 수가 증가하여 움직임 예측의 복잡도가 증가한다는 문제점이 있다. 그러므로 고해상도 영상 부호화에서는 계산량 감소를 위해 향상된 예측 지역 탐색 방법(Enhanced Predictive Zonal Search, EPZS)을 움직임 예측 모드로 사용하기를 권장하고 있다.

따라서, 본 논문에서는 움직임 예측에 소요되는 계산을 줄이기 위해 움직임 벡터의 시간적 상관도와 공간적 상관도를 이용하여 불필요한 탐색 지점을 제거하여 확장된 크기의 매크로블록을 위한 고속 움직임 예측 방법을 제안한다. 본 논문에서 제안하는 방법은 EPZS에도 적용 가능하므로 고속 EPZS에 대해서도 제안하고자 한다.

본 논문에서는 움직임 예측 기술에 대해 전체적으로 살펴보고자 한다. 2장에서는 전 영역 탐색 방법과 EPZS 방법에 대한 일반적인 내용을 소개한다. 3장에서는 제안하는 움직임 벡터 상관도와 참조 화면 거리를 이용한 H.264/AVC 움직임 탐색 지점 제거 방법에 대해 설명하고, 4장에서는 다양한 실험을 통해서 제안하는 방법의 성능을 비교하고, 마지막으로 5장에서 결론을 맺는다.

II. H.264/AVC 움직임 탐색 방법

그림 1은 H.264/AVC에서 제공하는 움직임 탐색 방법을 나타낸 것이다. 현재 부호화하고자 하는 매크로블록의 예측 움직임 벡터는 인접하는 블록에서 정해진 움직임 벡터들의 중간값을 수평 방향과 수직 방향으로 독립적으로 구한다. 결정된 예측 움직임 벡터는 움직임 벡터 부호화 및 움직임 예측의 초기 시작점으로 이용된다. 초기 시작점을 기준으로 상하좌우로 사용자가 지정한 탐색 영역의 크기만큼 탐색 영역이 정해진다. 탐색 영역을 로 정했다면 정화소 단위의 탐색 영역의 크기는 이 된다. 정화소 단위에서 움직임 예측 과정이 진행되어 결정된 최적의 정화소를 중심으로 부화소 단위로 움직임 예측 과정을 수행한다^[4,5].

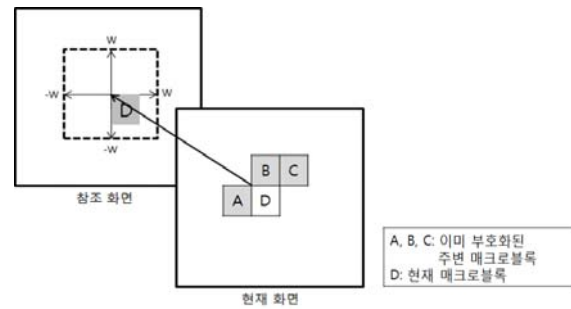


그림 1. H.264/AVC 움직임 탐색 방법
Fig. 1. H.264/AVC Motion estimation

2.1 향상된 예측 지역 탐색 방법

예측 움직임 벡터 선택은 향상된 예측 지역 탐색 방법에서 중요한 역할을 한다. 사용자가 지정한 탐색 영역 내에서 여러 개의 예측 움직임 벡터 후보들을 생성한다. 예측 움직임 벡터 후보들은 움직임 벡터의 공간적 상관도, 시간적 상관도, 매크로블록 타입 등을 기반으로 생성된다. 생성된 모든 예측 움직임 벡터 후보들에 대해서 비용함수 값을 계산한다. 후보들 중에서 최소의 비용함수 값을 가진 예측 움직임 벡터와 두 번째로 작은 비용함수 값을 가지는 예측 움직임 벡터를 결정한다. 최소의 비용함수 값이 초기 종료값보다 크면 2개의 예측 움직임 벡터에 대해서 개선 과정이 수행된다. 그렇지 않으면, 최소의 비용함수 값을 가지는 예측 움직임 벡터가 최종 움직임 벡터로 결정된다.

개선 과정은 최소의 비용함수 값을 가진 예측 움직임 벡터를 기준으로 하여 향상된 예측 지역 탐색 방법의 패턴을 이용하여 수행된다. 탐색 패턴은 사

움직임에 의해 설정되는데, 매크로블록의 타입과 참조 화면의 거리를 고려하여 조절된다. 각각의 탐색 패턴은 그림 2와 같다. 그림 2의 패턴을 적용해 2개의 예측 움직임 벡터에 대해서 개선 과정을 수행한다. 개선 과정을 거친 뒤, 최소 비용함수값을 가진 움직임 벡터가 최종 움직임 벡터로 결정된다⁶⁾.

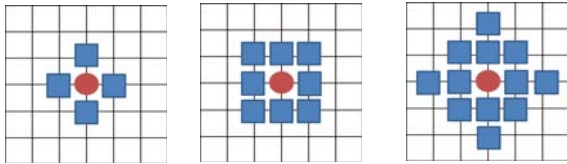


그림 2. 향상된 예측 지역 탐색 방법의 탐색 패턴
Fig. 2. Refinement process patterns.

III. 제안하는 고속 움직임 예측 방법

3.1 공간적 상관도를 고려한 탐색 지점 제거

한 화면에 위치하는 인접한 화소 사이에는 높은 상관도가 존재하는데, 이는 움직임 벡터에서도 동일하다. 그러므로 인접하는 매크로블록의 움직임 크기에도 공간적 상관도가 높다. 불필요한 비용함수 값을 계산하는 과정을 줄이기 위해 인접하는 블록 사이의 움직임 벡터의 통계적 특성을 이용하여 수평, 수직 방향으로 각각의 움직임 탐색 영역을 결정한다. 결정된 탐색 영역을 벗어난 탐색 지점에 대해서는 비용함수 계산 과정을 수행하지 않는다.

표준편차는 평균과 어느 정도 차이가 있는지를 나타내는 값이다. 움직임이 적은 영상에서는 인접한 블록들의 값이 거의 비슷하므로 작은 표준편차를 가질 것이라 예상할 수 있다. 움직임이 많은 영상에서는 인접한 블록의 움직임의 크기와 방향이 서로 유사하지 않아 큰 표준편차를 가질 것이라 예상할 수 있다. 다양한 영상을 통해 실험한 결과, 예상한 대로 움직임이 일정한 영상은 표준편차 값이 작고, 움직임이 일정하지 않은 영상은 표준편차 값이 크다는 것을 확인할 수 있었다. 이러한 사실을 통해, 제안한 알고리즘에서는 계산한 표준편차를 통해 현재 블록의 수평방향, 수직 방향의 움직임의 크기를 예측한다.

먼저, 움직임 벡터의 통계적 특성을 이용하고자 인접하는 블록의 움직임 벡터들의 평균값을 계산한다. 이를 통해 현재 부호화하고자 하는 매크로블록 주변 블록의 수직방향과 수평방향의 움직임의 크기 정도를 예측할 수 있다. 예를 들어 수평방향의 표준

편차가 수직방향의 표준편차보다 크면 현재 부호화하고자 하는 매크로블록 주변에 수평방향의 움직임이 크다는 것을 알 수 있다. 이를 이용하여 수평방향, 수직방향에 대해 각각 서로 다른 크기의 탐색 영역을 결정하게 된다. 수직방향과 수평방향의 움직임 벡터의 표준편차를 이용하여 적정한 영역 내에서 비용함수 값을 계산하므로 계산량 절감이 가능하다. 전 영역 탐색 방법의 경우 정해진 탐색 영역 내에서만 움직임 예측을 수행한다. EPZS의 경우에는 적정한 영역을 벗어난 예측 움직임 벡터를 제거하여 불필요한 탐색 지점을 제거한다. 그림 3은 EPZS에서 제안된 탐색 영역을 이용하여 불필요한 예측 움직임 벡터가 제거되는 것을 나타내고 있다.

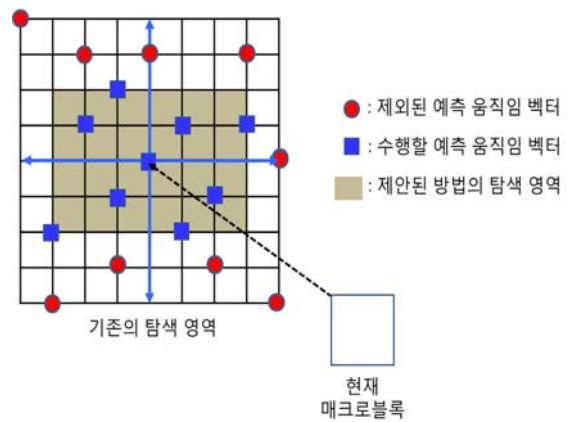


그림 3. 제안된 탐색 영역을 이용해 불필요한 탐색 지점을 제거하는 방법
Fig. 3. Reduction of motion vector predictors.

기존 방법의 성능을 유지하기 위해서 결정된 탐색 영역은 최적의 움직임 벡터로 결정될 가능성이 있는 지점을 포함해야 한다. 본 논문에서는 기존의 전 영역 탐색 방법과 EPZS의 성능을 보장하기 위하여 비선형 함수를 이용하여 탐색 영역을 결정한다. 움직임 벡터의 표준편차 값이 지정된 탐색 영역의 크기를 넘어가면 비선형 함수를 사용하여 최대 탐색 영역의 값을 가지도록 조절한다. 또한, 단순히 표준편차 값을 탐색 영역의 크기를 결정하는 방법보다 비선형적인 값으로 탐색 영역의 크기를 조정하여 움직임 벡터의 표준편차 값보다 더 큰 값을 탐색 영역의 크기로 결정하여 움직임 예측 성능을 보장할 수 있다. 수평방향과 수직 방향의 탐색 영역의 결정식은 다음과 같이 정의될 수 있다.

$$R = \alpha \sqrt{\text{standard_deviation}(MVs)} \quad (1)$$

$$\alpha = search_range/2\sqrt{search_range} \quad (2)$$

여기서 α 는 표준편차를 이용하여 결정된 탐색 영역이 사용자가 설정한 최대 탐색 영역을 넘어가는 것을 막기 위한 크기 조정 인자이다.

3.2 공간적 상관도를 고려한 탐색 지점 제거

공간적인 특성을 이용하여 결정된 탐색 영역은 모든 참조 화면에서 동일한 크기로 적용된다. 그러나 실제 최적의 참조 화면은 현재 부호화하고자 하는 매크로블록에서 가장 가까운 화면으로 선택될 확률이 높다. 그러므로 부호화하고자 하는 매크로블록에서 가장 멀리 있는 참조 화면에서 동일한 움직임 예측 과정을 수행하는 것은 불필요한 계산량 증가를 의미한다. EPZS의 경우, 모든 참조 화면에서 개선 과정을 수행한다. 최적의 참조 화면으로 선택될 확률이 낮은 멀리 있는 참조 화면에서 개선 과정을 수행함으로써 계산량이 증가한다. 그러므로, 전 영역 탐색 방법과 EPZS의 계산량을 줄이기 위해 가까운 참조 화면에 대해서는 세밀한 움직임 예측을 수행하고 참조 화면이 현재 부호화하고자 하는 매크로블록에서 멀어질수록 탐색 영역의 크기를 줄이거나 개선 과정을 생략한다. 전 영역 탐색 방법에서 최종 탐색 영역은 다음과 같이 결정된다.

$$r = |Best_MV - Predictive_MV| \quad (3)$$

$$Final_Search_Range = \frac{R + (Distance \times r)}{Distance + 1} \quad (4)$$

여기서 Distance은 참조 화면과 현재 화면 사이의 거리를 나타내고, R은 공간적 상관도를 이용하여 결정된 탐색 영역을 나타내고, r은 시간적 상관도를 고려하여 결정된 탐색 영역을 의미한다. r은 이전에 부호화된 매크로블록에서의 움직임 벡터와 초기에 구한 예측 움직임 벡터의 차이값을 의미한다.

작은 차이값은 현재 부호화하고자 하는 매크로블록의 움직임이 작다는 것과 초기 시작점으로부터 멀리 않은 곳에 최적의 움직임 벡터가 존재한다는 것을 의미한다. 반대로 큰 차이값은 현재 매크로블록의 움직임이 화면 사이에서 크다는 것과 초기 시작점으로부터 멀리 떨어진 곳에서 최적의 움직임 벡터가 존재하는 것을 나타낸다. 실험을 통해 확인해 본 결과, R의 크기는 r의 크기보다 크다. 식 (4)을 통해 최종 결정된 탐색 영역은 참조 화면과 현재 화면 사이의 거리가 멀어질수록 시간적 상관도

를 이용한 탐색 영역에 가중치를 줌으로써 거리를 고려하여 탐색 영역의 크기를 줄였다. 그러므로 멀리 있는 참조 화면에서 불필요한 탐색 지점의 계산량을 줄일 수 있다.

EPZS의 경우, 첫 번째 참조 화면을 제외한 참조 화면에서 움직임 예측에 소모되는 시간을 줄이기 위해 불필요한 개선 과정을 생략하였다. 첫 번째 참조 화면에서 결정된 최소의 비용함수 값은 나머지 참조 화면에서 개선 과정 적용 여부를 결정하는 임계값으로 사용된다. 첫 번째 참조 화면을 제외한 나머지 화면에서 결정된 최소 비용함수 값이 임계값보다 작으면 개선 과정이 수행되어 정확한 움직임 벡터가 결정된다. 첫 번째 참조 화면이 최적의 참조 화면으로 선택될 확률이 높기 때문에, 반대의 경우에 대해서는 개선 과정을 수행하지 않는다. 그림 4는 제안하는 적응적인 개선 과정 적용 방법을 나타내고 있다. 여기서 $min_cost(i)$ 는 i 번째 참조 화면에서의 최소의 비용함수 값을 나타낸다.

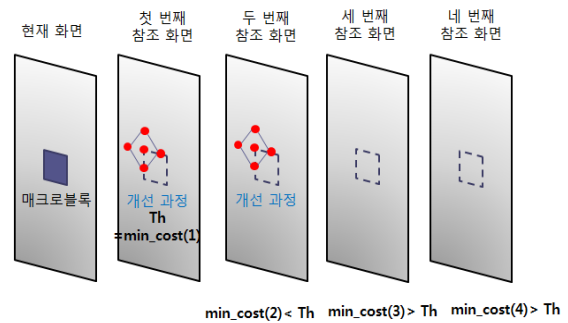


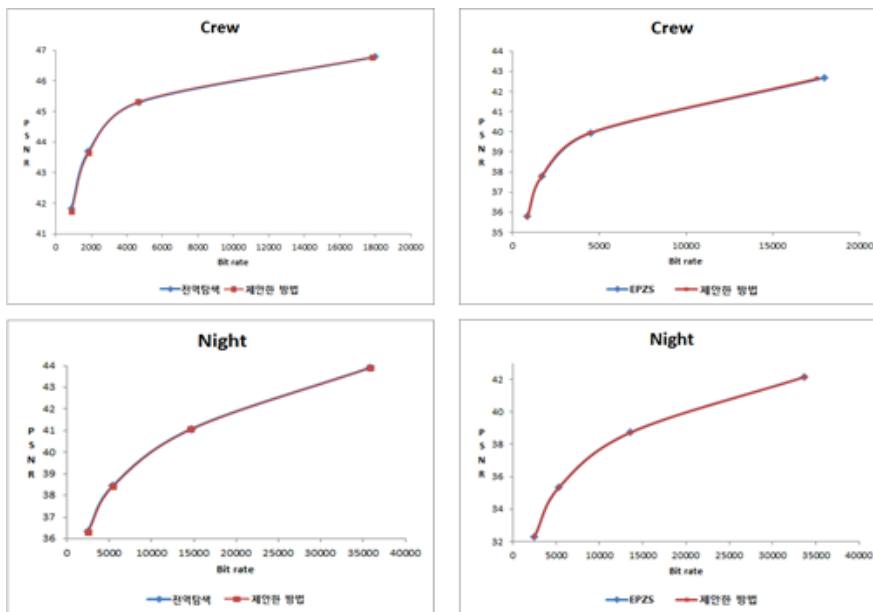
그림 4. 적응적인 개선 과정 적용 방법
Fig. 4. Proposed refinement process.

IV. 실험 결과 및 분석

제안하는 방법의 성능을 평가하기 위해 KTA 2.7 소프트웨어를 이용하여 실험했다. KTA 2.7 소프트웨어는 JM 11.0을 기반으로 한다. 전 영역 탐색 방법에서는 탐색 영역의 크기를 128로 하였고, EPZS에서는 탐색 영역의 크기를 96으로 하여 실험을 수행했다. 자세한 부호화 변수는 표 1과 같다.

표 1. 부호화 변수
Table 1. Encoding parameters

실험 변수	값
Profile	100 (high profile)
QP	22,27,32,37
SymbolMode	1 (CABAC)
FrameStructure	IPP...P
UseExtMB	2 (64×64)
NumberReferenceframes	4



(a) 전 영역 탐색 방법

(b) EPZS

그림 5. 비트율-왜곡 곡선
Fig. 5. Rate distortion curves.

표 2. 움직임 예측 시간 비교

Table 2. Processing time saving comparison

실험 영상	QP	전 영역 탐색 방법			EPZS		
		전 영역 탐색 방법	제안한 방법	△시간(%)	EPZS	제안한	△시간(%)
Crew	22	29600.69	13037.08	-55.96	45969.38	39287.40	-14.54
	27	26040.99	99736.00	-61.70	36089.86	30024.73	-16.81
	32	21645.74	81258.28	-62.46	32471.74	26532.78	-18.29
	37	16576.37	63817.24	-61.50	29387.82	23877.96	-18.75
Night	22	22628.35	12321.41	-45.55	47791.38	41094.92	-14.01
	27	21322.94	10527.71	-50.63	41057.20	34547.12	-15.86
	32	19614.12	88564.16	-54.85	35688.50	29536.75	-17.24
	37	17359.79	73733.32	-57.53	31103.24	25667.72	-17.48
Raven	22	237334.45	100689.86	-57.57	36224.93	28972.75	-20.02
	27	200256.95	77036.11	-61.53	29918.33	23900.96	-20.11
	32	161336.53	59629.66	-63.04	27251.87	22010.67	-19.23
	37	128489.13	47923.26	-62.70	25243.53	20986.38	-16.86
Jets	22	267731.64	123622.39	-53.83	30501.10	29048.10	-4.76
	27	186280.81	78575.79	-57.82	24022.12	22760.17	-5.25
	32	122048.62	51440.21	-53.85	21893.72	20360.19	-7.00
	37	83916.60	38963.54	-53.57	20598.72	18702.99	-9.20
평균				-57.13			-14.71

제안된 알고리즘의 성능을 평가하기 위해서 1280×720 해상도의 영상을 이용하여 실험을 진행했다. 표 2는 기존의 움직임 예측 방법과 제안하는

방법의 움직임 예측 시간 비교를 나타낸다. 표 2를 통해 제안한 방법을 이용하면 전 영역 탐색 방법의 움직임 예측 시간을 평균 57.13% 감소함을 확인할

수 있다. 또한 EPZS의 움직임 예측 시간을 평균 14.71% 감소함을 확인할 수 있다. 기존 방법의 성능과 유사함은 그림 5의 비트율 왜곡 곡선을 통해 확인할 수 있다.

V. 결 론

본 논문에서는 움직임 벡터의 상관도를 이용하여 불필요한 탐색 지점을 제거하는 방법을 제안했다. 인접하는 블록의 움직임 벡터의 표준편차의 비선형 함수를 이용하여 최종 탐색 영역 및 향상된 예측 지역 탐색 방법을 제안했다. 제안한 최종 탐색 영역과 향상된 예측 지역 탐색 방법은 전 영역 탐색 방법과 동일한 성능을 유지하면서 각각 평균 57.13%, 평균 14.71%의 움직임 예측 시간을 감소시켰다.

참 고 문 헌

- [1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Transactions on Circuits and System for Video Technology, Vol. 13, no. 7, pp. 560-579, July 2003.
- [2] Joint Call for Proposals on Video Compression Technology, ISO/IEC MPEG 91st Meeting, Kyoto, Japan, Document N11113, April 2010.
- [3] K. Lee, E. Alshina, J. Park, W. Han and J. Min, "Technical consideration for Ad Hoc Group on New Challenges in Video Coding Standardization," ISO/IEC JTC/SC29/WG11 doc.M1558, July 2008, Hannover, Germany
- [4] I. Richardson, H.264 and MPEG-4 Video Compression, Wiley, 2003, pp. 175-177.
- [5] J.H. Moon and Y.S. Ho, "Fast Motion Estimation with Extended Block Sizes for High-definition Video Coding," Asia-Pacific Signal and Information Processing Association (APSIPA) 2010, pp. 619-622, Dec. 2010.
- [6] A. Tourapis, "Enhanced Predictive Zonal Search for Single and Multiple Frame Motion Estimation," Visual Communications and Image Processing, pp.1069-1079, Jan. 2002.

문 지 희 (Ji-Hee Moon)

정회원



2008년 세종대학교 전자공학과 졸업(학사)
 2011년 광주과학기술원 정보통신공학과 졸업(석사)
 2011년~현재 LG전자 <관심분야> 디지털 영상처리, H.264/AVC, HEVC

최 정 아 (Jung-Ah Choi)

정회원



2007년 한국항공대학교 항공전자 및 항공전자공학과 졸업(학사)
 2008년 광주과학기술원 정보통신공학과 졸업(석사)
 2008년~현재 광주과학기술원 정보통신공학과 박사과정 <관심분야> 디지털 영상처리, H.264/AVC, HEVC

호 요 성 (Yo-Sung Ho)

중신회원



1981년 서울대학교 공과대학 전자공학과 졸업(학사)
 1983년 서울대학교 대학원 전자공학과 졸업(석사)
 1989년 Univ. of California, Santa Barbara, Dept. of Electrical and Computer Engineering.(박사)

1983년~1995년 한국전자통신연구소 선임연구원
 1990년~1993년 미국 Philips 연구소, Senior Research Member
 1995년~현재 광주과학기술원 정보통신공학과 교수
 2003년~현재 광주과학기술원 실감방송연구센터 센터장
 <관심분야> 디지털 신호처리, 영상 신호 처리 및 압축, 디지털 TV와 고선명 TV, 멀티미디어 시스템, MPEG 표준, 3차원 TV, 실감방송, H.264/AVC, HEVC