

<http://dx.doi.org/10.7236/IIBC.2015.15.4.161>

IIBC 2015-4-20

SURF 특징과 블록 모양을 이용한 동영상 압축 방법

Video Compression Method Using SURF Feature and Block Shaping

전재현*, 김민준**, 장용석***, 안철웅****, 김승호*****

Jae-Hyun Jun*, Min-Jun Kim**, Yong-Suk Jang***, Cheol-Woong Ahn****,
Sung-Ho Kim*****

요약 오늘날 컴퓨터 네트워크 인프라 구축이 일반화되고 인터넷 사용자의 수가 증가하는 가운데 망에는 대용량의 트래픽이 흐르고 있다. 네트워크 트래픽은 2014년에 월간 47 엑사바이트(Exabyte), 2015년 월간 62 엑사바이트로 엄청난 증가를 보인다. 인터넷 트래픽 중에서 동영상 트래픽의 비중은 전체 60%를 차지하여 인터넷 트래픽 상승의 견인차 역할을 수행하고 있다. 엄청난 인터넷 트래픽의 성장에 비해서 네트워크 인프라는 이를 지원하기 턱없이 부족하여 모바일 네트워크에서 음성통화의 품질저하라는 극단적인 부작용이 나타나고 있다. 미래에 예상되는 잠재적인 위험을 사전에 방지하기 위해 SURF 특징과 블록 모양을 이용한 동영상 압축 방법을 제안한다. 제안한 방법은 기존의 방법보다 좋은 성능을 보이는 것을 확인하였다.

Abstract Today, computer network infrastructure is commonplace and internet user increase. Transmission bandwidth for network traffic is increasing. The data amount of 2014 and 2015 year measured 47 and 62 exabyte, it shows a tremendous increase. Video traffic occupy 60 percent among internet traffic. Enormous growth of Internet traffic than the network infrastructure is not enough to support them, so mobile network occur side-effects such as fall of voice communication. In this paper, we proposed the video compression method using SURF feature and block shaping in order to avoid potential damage. Our proposed method can provide enough performance than previous method.

Key Words : SURF Features, Video Compression, H.264/AVC

1. 서 론

최근 초고속 인터넷과 디지털 기술이 급속하게 발전에 힘입어 네트워크는 거대한 발달을 이루어왔다. 특히

국내에서는 BcN (Broadband Convergence Network) 구축 사업에 의한, 통합 광대역 네트워크 인프라는 더욱 확대 되었다. 네트워크 기술의 성장과 더불어 네트워크를 이용한 다양한 서비스의 등장은 인터넷 이용자 수와 이

*준회원, 경북대학교 컴퓨터학부

**정회원, (주)에이투텍 R&D부서

***정회원, (주)다올디엔에스 R&D부서

****정회원, 계명문화대학 디지털콘텐츠학부

*****정회원, 경북대학교 컴퓨터학부 (교신저자)

접수일자 2015년 7월 28일, 수정완료 2015년 8월 7일

게재확정일자 2015년 8월 7일

Received: 28 July, 2015 / Revised: 7 August, 2015

Accepted: 7 August, 2015

*****Corresponding Author: shkim@knu.ac.kr

School of Computer Science and Engineering, Kyungpook National University, Korea

용량의 증가라는 결과를 가져왔다. 네트워크 서비스를 지원하는 하드웨어 기술의 발전으로 인해 인터넷 이용자 수와 증가를 가속화 시키고 있다.

네트워크 이용은 네트워크 기술의 발전과 무선 인터넷 기술의 발전 등으로 인해 지속적으로 증가하고 있다. 활성화된 네트워크를 바탕으로 다양한 서비스를 제공받고자 하는 이용자의 요구도 증가하고 있다.

비디오는 컴퓨터 산업에서 중요한 가치가 있는 데이터 중 하나이다. 오늘날, 컴퓨터를 이용하는 사람들은 비디오 장면을 통해 자신의 요구를 만족시키고 있다. 이런 경향은 앞으로 더욱 가속화 될 전망이다^[1].

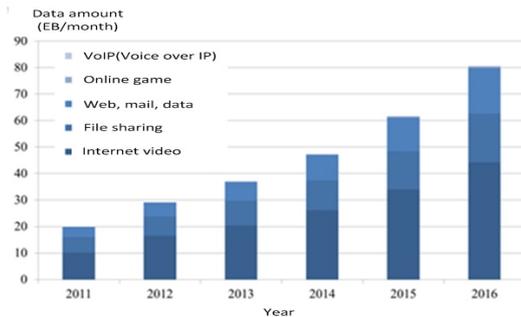


그림 1. 세계 네트워크 데이터 발생량
Fig. 1. World network data generation amount

네트워크 자원의 사용량은 그림 1과 같이 2012년 기준으로 월간 약 29 엑사바이트 (EB, Exabyte, 260 bytes) 정도였으나, 2013년 37EB, 2014년 47EB, 2015년 62EB, 2016년 80EB로 엄청난 증가를 보이면서 팽창할 것으로 예상된다. 인터넷 트래픽 중에서 비디오 트래픽의 비중은 전체 60%를 차지하여 인터넷 트래픽 상승의 견인차 역할을 수행하고 있다. 엄청난 인터넷 트래픽의 성장에 비해서 네트워크 인프라는 이를 지원하기 턱없이 부족하여 모바일 네트워크에서 음성통화의 품질저하라는 극단적인 부작용이 나타나고 있다. 미래에 예상되는 잠재적인 위험을 사전에 방지하기 위해 네트워크 인프라 확장 및 비디오 트래픽을 줄일 수 있는 고효율 압축 방법의 개발이 시급한 실정이다.

1장에서는 연구를 수행하게 된 배경에 대해 설명하였다. 2장에서는 기존 동영상 압축 방법들에 대해 소개하고, 3장에서는 본 논문에서 제안한 SURF(Speeded Up Robust Features) 특징과 블록 모양을 이용한 동영상 압축 방법에 대해 설명한다. 4장에서는 제안한 방법의 실험

결과에 대해 살펴보고 5장에서 결론을 내린다.

II. 관련 연구

오늘날 대용량의 영상을 압축하기위해 다양한 연구가 진행되고 있다^[8,9]. 이 중에서 H.264/AVC는 MPEG-2와 같은 범용적인 부호화를 목표로, 기존의 압축기술에 비해 높은 압축률을 달성하는 것을 목표로 제정된 압축기술이다. H.264/AVC는 다음의 3가지의 예측 기술을 기본 요소로 사용하고 있다^[7].

먼저 화면간 예측과 화면내 예측에 대해 알아본다. TV화면은 1초에 30장 또는 60장의 프레임으로 구성되므로 한 프레임과 이웃한 프레임 사이에는 그 차이가 작기 때문에 인간의 눈으로는 구분하기가 어렵다. 따라서 화면 A와 그 다음에 오는 화면 B가 비슷하다면 화면 A를 구성하고 있는 화소값으로부터 다음에 오는 화면 B의 화소 값을 예측할 수 있다

다음 움직임보상 화면간 예측에 대해 알아본다. 동영상은 시간적으로 연속된 프레임으로 구성되어 있기 때문에 화면간 예측은 매우 유용한 압축기술이지만, 두 화면 사이에 피사체의 큰 움직임이 있다면, 단순한 화면간 예측은 맞지 않게 된다. 그러나 형태가 변하지 않는 물체가 움직일 때, 움직임 방향과 움직임 크기를 알 수 있다면, 그 물체의 모습에 관한 정보를 이미 알고 있기 때문에 정확한 예측화면을 만들 수 있게 되어 높은 효율의 화면간 예측을 할 수 있다.

마지막으로 양방향 움직임보상 화면간 예측이다. 현재 프레임에서 물체의 위치를 과거와 미래의 양방향에서 예측하는 기술이다. 이 경우는 지금 압축 부호화할 화면을 예측하는데 있어서 시간적으로 전에 위치한 화면과, 시간적으로 후에 위치한 화면의 양방향으로부터 움직임보상 예측을 수행한다. 미래로부터의 예측은 압축부호화를 해야 하는 동영상의 화면 순서를 바꿔, 우선 선행하는 화면과 뒤에 나오는 화면을 압축부호화하고, 그 다음에 이것을 사용해 현재 화면을 예측한다.

기존의 비디오 압축 방법^[2,3]는 비디오 콘텐츠 고려없이 syntax에 의해 처리된다. 즉, 기존에 기반한 처리가 수행된다. 이러한 방법들은 조명의 변화, 카메라의 회전, 확대/축소에 쉽게 영향을 받는다. 그 결과, 비디오 압축을 상쇄시키고 전체적인 압축 성능 역시 감소하였다.

본 논문에서는 압축 성능 저하를 해결하기 위해 SURF 특징과 블록 모양을 이용한 동영상 압축 방법을 제안한다. 제안된 방법은 장면 검출, 에너지 점, 블록 형성, 오류 검출 처리 단계로 구성되어 있다. 각각의 프로세스를 통해, 제안된 방법은 효과적으로 비디오 콘텐츠에 따라 비디오 압축을 할 수 있다.

III. SURF 특징과 블록 모양을 이용한 동영상 압축 방법

이 장에서는 제안한 방법에 대한 내용을 소개한다. 아래 그림 2는 제안한 방법을 나타내고 있다. 동영상 압축을 위해 장면 검출, 특징점 검출, 블록 형성, 오류 처리 단계로 구성되어 있다.

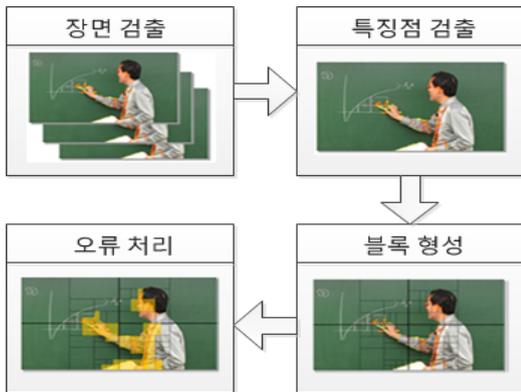


그림 2. 제안한 방법
 Fig. 2. Proposed method

먼저 장면 검출에 대해 알아본다. 장면 검출은 장면이 구분되는 첫 번째 프레임인 키(key) 프레임을 획득하는 것이다. 검출한 키 프레임과 다음에 검출되는 키 프레임 사이의 연속하는 프레임들이 하나의 장면을 구성한다. 장면 검출은 [4]를 이용해 비디오 시퀀스들을 나눈다.

장면이 분할되어 키 프레임이 검출되면 키 프레임 내의 특징점을 검출한다. 특징점 검출을 위해 SURF 특징 기술자를 이용한다[5]. SURF 특징 기술자는 비교할 영역의 지역적인 특징을 이용한 방법으로 적분 영상과 근사화 된 헤이시안 검출기에 기반한 고속 헤이시안 탐지기를 사용하여 빠르게 특징점을 탐지한다. 검출한 특징점을 그림 3과 같다.

외부 요소와 관계없는 국소 특징(local feature)를 추출하고 검색하는데 우수한 성능을 보이는 방법으로, 영상의 크기, 조명 값 등의 영향에 독립적인 특징을 추출할 수 있는 장점을 가진다. SURF 특징 기술자를 이용해 검출한 결과는 특징점으로 칭하며, 이는 비디오 압축의 요소처럼 이용할 수 있다.

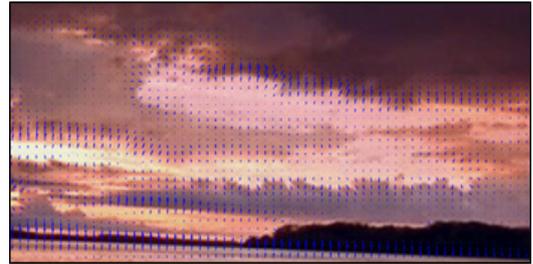


그림 3. 특징 점 검출
 Fig. 3. Feature point detection

다음으로 블록화 단계이다. 특징점은 영상의 주요한 특징을 포함하는 점들이므로 압축 효율을 위해 영상에서 특징점을 포함하고 있는 영역과 그렇지 않은 영역으로 구분한다. 특징점이 포함된 영역은 영상에서 상대적으로 주요한 부분으로 가정하며 그렇지 않은 영역에 비해 압축을 위한 추가 과정이 필요하다.

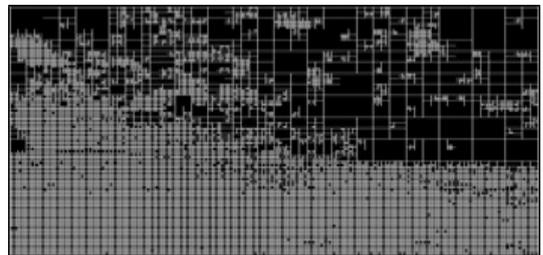


그림 4. 기억 블록과 비 기억 블록
 Fig. 4. Memorial and non-memorial blocks

블록화 단계에서는 그림 4와 같이 에너지 점 유무에 따라 기억 블록과 비기억 블록으로 나눈다. 우리는 4진트리(quad-tree)[6]를 이용하여 기억 블록인지 비기억 블록인지 발견한다. 비기억 블록은 기억 블록보다 더욱 중요한 에너지 점을 가지고 있기 때문에, 추가 코딩 처리가 처음 필요하다.

블록 형성이 종료되면 프레임은 크게 기억 블록과 비기억 블록으로 분류되며, 오류 처리 과정을 통해 이 두

블록 영역에 대한 압축이 수행된다. 오류 처리 과정은 기존 압축 표준의 P 프레임 또는 B 프레임의 화면 간 예측 과정과 유사하다. 참조 영상으로부터 현재 블록의 움직임을 예측해내는 화면 간 예측과 같이 오류 처리 과정도 참조 영상으로부터 해당 블록의 영상을 표현할 수 있는 방법을 찾는 단계이다.

기억 블록은 특징점이 존재하지 않는 영역이므로 프레임 내에서 비기억 블록에 비해 상대적으로 중요도가 낮은 부분이다. 하지만 영상의 질을 최대한으로 보장되어야 하므로, 현존하는 압축 표준인 H.264/AVC [7] 압축 표준의 화면간 예측을 사용한다.

비 기억블록은 특징점을 포함하는 영상의 중요한 부분이므로 압축 성능이 보장되어야 한다. 특징 점 검출과 매칭 과정을 통과하면 특징점의 이동 벡터가 먼저 추출된다. 이 벡터는 참조 프레임에 존재하는 에너지 점이 현재 프레임에서 얼마만큼 이동한 것인지에 대한 정보를 나타낸다. 이동 벡터를 이용하여 현재 비기억 블록을 압축하기 위해 참조해야 할 위치를 파악할 수 있다.

IV. 실험 및 결과

이 장에서는 본 논문에서 제안한 SURF 특징과 블록 모양을 이용한 동영상 압축 방법의 성능에 대해 알아본다. 제안한 방법의 성능을 측정하기 위해, 우리는 PSNR (Peak-Signal-to-Noise Ratio)과 압축률로 측정한다. 각 실험 결과는 H.264/AVC [7]와 비교하였다.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (1)$$

$$여기서, MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2$$

$$압축률 = \frac{압축된 크기}{원본 크기} \times 100 \quad (2)$$

식 1의 MSE(Mean Square Error)는 두 영상 간의 화소 간 차이의 제곱의 평균을 의미한다. MSE는 두 영상 간의 상대적인 오류를 구하는 것이라면, PSNR은 신호 중 최고 수치 MAX와의 비교를 통해 절대적인 오차를 계산한다. 신호 처리나 영상 처리에서 PSNR은 객관적인 오류에 대한 평가 기준으로 이용되며 본 논문에서도 복원 영상의 객관적인 화질 비교를 위해 사용한다.

압축 성능에 대한 평가 척도로써 압축률(compression ratio)을 사용하였다. 압축률은 원본 영상의 크기와 비교하여 압축된 영상의 크기의 비율을 나타낸 것으로 식 2와 같이 계산된다. 두 영상의 비를 백분율로 나타낸 것이 압축률이다. 압축률의 수치가 작을수록 더 많은 압축이 이루어진 것으로 판단할 수 있다.

실험 영상은 저화질 영상과 고화질 영상으로 나뉜다. 저화질 영상은 41개의 저화질 CCTV 영상으로 가로 720 세로 480 크기이다. 11개의 일반화질 영상은 가로 1080 세로 720이고 15개의 고화질 영상은 가로 1980 세로 1080의 영상에 대해 실험하였다.



그림 5. 저화질 영상 샘플
Fig. 5. low image quality video samples

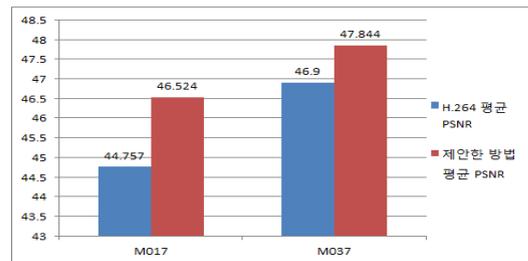


그림 6. 각 방법의 평균 PSNR
Fig. 6. Average PSNR of each method



그림 7. 각 방법의 크기
Fig. 7. Size of each method

그림 5~7은 저화질 영상에 대한 샘플과 평균 PSNR 그리고 압축 후 크기 값이다. 그림 6과 7은 41개의 저화질 영상 중 압축률이 가장 작은 영상과 가장 큰 영상을 비교한 것이다. 그림 6에서 왼쪽은 M017 실험 영상에 대한 결과로 H.264 방법의 평균 PSNR은 44.757dB 이고 제안한 방법은 46.524dB이다. M037 실험 영상에서 H.264는 46.9dB이고 제안한 방법은 47.844dB이다. 저화질 영상에서 PSNR은 제안한 방법이 H.264방법보다 약 1~2dB 높은 것을 확인하였다. 그림 7은 압축 후 크기에 대해 비교한 것이다. M017영상의 경우 H.264의 크기가 6290kb 인 것에 반해 제안한 방법의 크기는 4585kb 이다. 이 두 값의 이득은 4.642%이다. M037 영상의 경우 H.264의 크기는 5998kb 인 것에 반해 제안한 방법의 크기는 3555kb 이다. 저화질 영상에서 제안한 방법의 압축률이 더 높은 것을 볼 수 있다.



그림 8. 일반화질 영상 샘플
 Fig. 8. Normal quality video samples

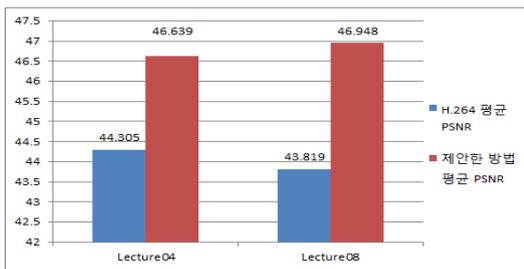


그림 9. 일반 화질에서 각 방법의 평균 PSNR
 Fig. 9. Average PSNR of each method on normal image quality

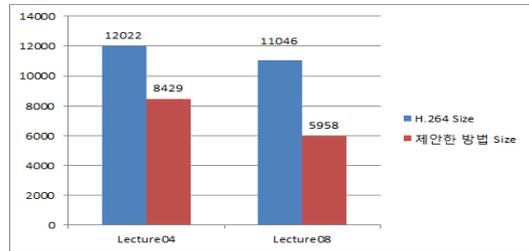


그림 10. 일반 화질에서 각 방법의 크기
 Fig. 10. Size of each method on normal image quality

그림 8~10은 일반 영상에 대한 샘플과 평균 PSNR 그리고 압축 후 크기 값이다. 그림 8~10은 11개의 일반화질 영상 중 압축률이 가장 작은 영상과 가장 큰 영상을 비교한 것이다. 그림 8에서 왼쪽은 Lecture04 실험 영상에 대한 결과로 H.264 방법의 평균 PSNR은 44.305dB 이고 제안한 방법은 46.639dB이다. Lecture08 실험 영상에서 H.264는 43.819dB이고 제안한 방법은 46.948dB이다. 일반화질 영상에서 PSNR은 제안한 방법이 H.264방법보다 약 2~3dB 높은 것을 확인하였다. 그림 9은 압축 후 크기에 대해 비교한 것이다. Lecture04영상의 경우 H.264의 크기가 12022kb 인 것에 반해 제안한 방법의 크기는 8429kb 이다. 이 두 값의 이득은 29.887%이다. Lecture08 영상의 경우 H.264의 크기는 11046kb 인 것에 반해 제안한 방법의 크기는 5958kb 이다. 배경이 고정된 영상의 경우 제안한 방법의 압축률이 월등히 높은 것을 볼 수 있다.



그림 11. 고화질 영상 샘플
 Fig. 11. High quality video samples

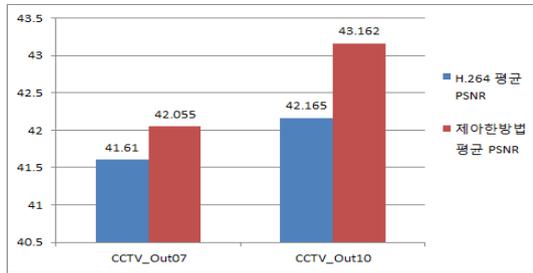


그림 12. 고화질에서 각 방법의 평균 PSNR
Fig. 12. Average PSNR of each method on high image quality

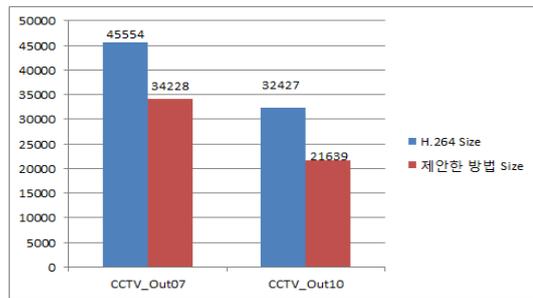


그림 13. 고화질에서 각 방법의 크기
Fig. 13. Size of each method on high image quality

그림 11~13은 고화질 영상에 대한 샘플과 평균 PSNR 그리고 압축 후 크기 값이다. 그림 11~13은 15개의 고화질 영상 중 압축률이 가장 작은 영상과 가장 큰 영상을 비교한 것이다. 그림 11에서 왼쪽은 CCTV_Out07 실험 영상에 대한 결과로 H.264 방법의 평균 PSNR은 41.61dB 이고 제안한 방법은 42.055dB이다. CCTV_Out10 실험 영상에서 H.264는 42.165dB이고 제안한 방법은 43.162dB이다. 고화질 영상에서 PSNR은 제안한 방법이 H.264방법보다 약 0.5~1dB 높은 것을 확인하였다. 그림 13은 압축 후 크기에 대해 비교한 것이다. CCTV_Out07영상의 경우 H.264의 크기가 45554kb 인 것에 반해 제안한 방법의 크기는 34228kb 이다. 이 두 값의 이득은 24.86%이다. CCTV_Out10 영상의 경우 H.264의 크기는 32427kb 인 것에 반해 제안한 방법의 크기는 21639kb 이다. 제안한 방법의 압축률이 월등히 높은 것을 볼 수 있다.

V. 결론

본 논문에서는 SURF 특징과 블록 모양을 이용한 동영상 압축 방법을 제안하였다. 실험 결과 제안한 방법이 H.264/AVC보다 PSNR이 1dB이상 높은 것을 보였으며 압축률 또한 더 높은 것을 확인하였다. 특히 배경이 고정되어 있는 강의영상 경우는 H.264/AVC보다 약 20%이상 압축률을 보이는 것을 확인하였다.

향후 추가로 연구되어야 할 부분으로 H.265/HEVC를 메모리얼 블록에 적용하여 더 높은 압축률을 연구해야 할 계획이다. 이를 적용하면 지금보다 더 높은 성능을 보일 것으로 예상되고, 추가로 비디오 시간 중복성 제거에 여러 기술들을 다각도로 검증함으로써 압축률 향상에 초점을 맞춰 향후 연구를 진행할 계획이다.

References

- [1] Cisco Systems, Inc., "www.cisco.com".
- [2] ITU Telecommunication Standardization Sector, "www.itu.int/en/ITU-T/pages/default.aspx".
- [3] The International Organization for Standardization, "www.iso.org/iso/home.html".
- [4] S. Lian, "Automatic Video Temporal Segmentation based on Multiple Features," *Soft Computing - A Fusion of Foundations, Methodologies and Applications - Special Issue on Digital Information Forensics*, Vol. 15, No. 3, pp. 469-482, Mar. 2011.
- [5] H. Bay, T. Tuytelaars and L. V. Vool, "SURF: Speeded Up Robust Features," *European Conference on Computer Vision*, Vol. 3951, pp. 404-417, 2006.
- [6] H. Samet, "Application of Spatial Data Structure," Addison-Wesley, New York, 1990.
- [7] ITU-T SG16 and ISO/IEC JCT1/SC29/WG11, "Information Technology - Coding of Audio-Visual Objects - Part 10: Advanced Video Coding," ISO/IEC 14496-10, 2012.
- [8] C. Kim, and J. W. Oh, "BTC-based Image Compression using Pattern", *The Journal of The*

Institute of Internet Broadcasting and Communication,
 Vol. 15, No. 3, pp. 85-90, June 2015.

[9] I. H. Jee, "Effective Compression Technique of Multi-view Image expressed by Layered Depth Image", The Journal of The Institute of Internet Broadcasting and Communication, Vol. 14, No. 4, pp. 29-37, Aug. 2014.

저자 소개

전 재 현(준회원)



- 2009년 : 대구가톨릭대학교 전자공학과 (공학사)
- 2011년 : 경북대학교 전자전기컴퓨터학부 (석사)
- 2011년 ~ 현재 : 경북대학교 컴퓨터학부 박사과정

<주관심분야 : 트래픽 분류, 디도스 공격 탐지, 멀티미디어, 기계학습>

김 민 준(정회원)



- 2005년 : 경북대학교 컴퓨터공학과 (공학사)
- 2007년 : 경북대학교 컴퓨터공학과 (공학석사)
- 2012년 : 경북대학교 전자전기컴퓨터학부 (공학박사)
- 2012년 ~ 현재 : ㈜에이투텍 책임연구원

<주관심분야 : 트래픽 분류, 침해 트래픽 특성 연구, 멀티미디어, 기계학습>

장 용 석(정회원)



- 1999년 : 경북대학교 컴퓨터공학과 (공학사)
- 2001년 : 경북대학교 컴퓨터공학과 (공학석사)
- 2007년 : 경북대학교 컴퓨터공학과 (공학박사)
- 2007년 ~ 현재 : ㈜다올디엔에스 대표이사

<주관심분야 : 트래픽 분류, 침해 트래픽 특성 연구, 멀티미디어, 기계학습>

안 철 웅(정회원)



- 1993년 : 경북대학교 컴퓨터공학과 (공학사)
- 1995년 : 경북대학교 컴퓨터공학과 (공학석사)
- 2009년 : 경북대학교 컴퓨터공학과 (공학박사)
- 2001년 ~ 현재 : 계명문화대학교 디지털콘텐츠학부 교수

<주관심분야 : 멀티미디어, 멀티미디어통신, 이미지 처리, 이미지 검색, 빅데이터 처리>

김 승 호(정회원)



- 1981년 : 경북대학교 전산학과 (공학사)
- 1983년 : 한국과학기술원 전산학과 (공학석사)
- 1994년 : 한국과학기술원 전산학과 (공학박사)
- 1986년 ~ 현재 : 경북대학교 컴퓨터학부 교수

<주관심분야 : 멀티미디어, 다시점영상, 트래픽 분류, 동시식 네트워크>

※ 본 논문(저서)은 교육부 및 한국연구재단의 BK21 플러스 사업 (경북대학교 컴퓨터학부 Smart Life 실현을 위한 SW인력양성사업단)으로 지원된 연구임 (21A20131600005)