특집논문 (Special Paper)
방송공학회논문지 제20권 제2호, 2015년 3월 (JBE Vol. 20, No. 2, March 2015)
http://dx.doi.org/10.5909/JBE.2015.20.2.204
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

깊이맵 업샘플링을 이용한 객관적 메트릭과 3D 평가의 비교

사이드 마흐모드포어[a)], 최 창 열[a)], 김 만 배[a)‡]

Comparison of Objective Metrics and 3D Evaluation Using Upsampled Depth Map

Saeed Mahmoudpour[a)], Changyeol Choi[a)], and Manbae Kim[a)‡]

요 약

업샘플링 깊이맵은 깊이 카메라로부터 획득된 깊이맵의 공간 해상도를 증가시키는 방법이다. 깊이맵의 성능은 입체영상, 멀티뷰의 3D 입체감과 밀접한 관계가 있다. PSNR 등의 객관적 메트릭으로 깊이맵의 업샘플링 성능을 평가하고, 생성된 입체영상은 주관적 평가를 통해서 입체감 및 시각적 피로도를 조사한다. 후자의 주관적 평가는 인적 물적 자원을 필요로 하는 반면에, 전자의 객관적 메트릭은 수학적 표현으로 정량적 수치값을 알려준다. 따라서 주관적 평가와 높은 상관관계를 가지는 객관적 메트릭이 주관적 평가를 대체할 수 있다면 많이 시간을 필요로 하는 주관적 평가가 불필요하다. 이를 위해 본 논문에서는 다양한 객관적 메트릭과 3D 주관적 평가 사이의 관계를 조사한 후에, 이용한 메트릭에 기반한 주관평가와 상관관계가 높은 객관적 메트릭을 제안한다. 업샘플링된 깊이맵의 성능을 측정하기 위해 다양한 참조영상 및 무참조영상 평가 메트릭들을 이용하였다. 주관적 평가는 DSCQS 입체영상 테스트로 수행되었다. 세 종류의 상관관계의 활용 및 분석을 통해서, SSIM과 Edge-PSNR이 주관적 평가를 대체할 수 있는 적합한 객관적 메트릭임을 실험을 통해서 검증하였다.

Abstract

Depth map upsampling is an approach to increase the spatial resolution of depth maps obtained from a depth camera. Depth map quality is closely related to 3D perception of stereoscopic image, multi-view image and holography. In general, the performance of upsampled depth map is evaluated by PSNR (Peak Signal to Noise Ratio). On the other hand, time-consuming 3D subjective tests requiring human subjects are carried out for examining the 3D perception as well as visual fatigue for 3D contents. Therefore, if an objective metric is closely correlated with a subjective test, the latter can be replaced by the objective metric. For this, this paper proposes a best metric by investigating the relationship between diverse objective metrics and 3D subjective tests. Diverse reference and no-reference metrics are adopted to evaluate the performance of upsampled depth maps. The subjective test is performed based on DSCQS test. From the utilization and analysis of three kinds of correlations, we validated that SSIM and Edge-PSNR can replace the subjective test.

Keyword : Depth map, Upsampling, Objective metric, 3D subjective test, Correlation

Ⅰ. 서 론

현재 카메라 시스템과 컴퓨터 비전 기술의 발전에 따라 3D 능동 카메라는 장면의 정확한 거리 정보를 제공할 수 있게 되었다. 고속의 Time of Flight(ToF) 카메라는 신뢰성 있는 깊이맵(depth map)을 추출한다. 그러나, 깊이맵의 공간적 해상도는 RGB 영상에 비해서 상대적으로 낮으므로, 고해상도 깊이맵 획득을 위한 다양한 깊이맵 업샘플링 기법들이 제안되었다[1-5]. 또한 HD(High Definition)에서 FHD (Full HD)로, 다시 FHD에서 UHD(Ultra High-Definition)로 발전하는 디스플레이 해상도 증가로 기 제작된 깊이맵 또한 이에 맞추어 업샘플링(upsampling) 기술을 필요로 한다. 또한 RGB+Depth 전송에서는 코딩 효율을 향상하기 위해, 서버 엔코더(encoder)에서는 깊이맵의 해상도를 줄여 압축하고, 클라이언트 디코더(decoder)에서 다시 원 깊이맵의 해상도로 높힌다[1]. 깊이맵은 3D 입체영상, 멀티뷰 영상 등의 콘텐츠 제작 및 성능에 중요한 역할을 한다.

양선형 보간법(bilinear upsampling, BLU)는 업샘플링된 깊이맵을 이루는 픽셀 보간을 위해 4개의 인접 픽셀들의 가중치 평균을 이용한다. 고등차수 보간법(bicubic upsampling, BCU)이라고 불리는 유사한 방법은 16개의 인접픽셀을 기반으로 한다. 양측 보간법(bilateral upsampling, BU)는 업샘플링 과정에서 에지 영역을 보존하기위한 레인지 필터와 공간 필터를 결합한 기법이다[2]. 또 하나의 BU를 기반으로 하는 업샘플링 방법은 색상 정보와 저해상도 깊이맵을 이용한 결합 양측 업샘플링(joint bilateral upsampling, JBU)이다[3]. 분산 양측 업샘플링(variance based upsampling, VBU)는 고정 분산을 사용하지 않고, 각 픽셀블록 단위로 계산된 분산 값을 사용한다[4]. JBU는 균일 영역에 민감하며, 가중치 함수는 비에지(non-edge) 영역에서 잘못된 분산값을 얻을 수 있다. 이 문제점을 해결하기위해, 큰 가중치는 각 에지 픽셀들의 색상에 할당하고, 반대로 작으면 비에지 픽셀들의 깊이에 할당되도록 하는 적응적 양측 업샘플링(adaptive bilateral upsampling, ABU) 방법이 제안되었었다[5]. 또한 낮은 그라디언트(gradient) 에지에서 흐림(blur)의 감소를 줄이기 위해서 거리변환 양측 업샘플링(distance transform-based bilateral upsampling, DTBU)이 있다[6].

그동안 제안되었던 많은 깊이맵 업샘플링들의 성능 평가(quality assessment, QA)로 일반적으로 사용되는 방법은 PSNR(Peak Signal-to-Noise Ratio)이다. 일반적으로 깊이맵의 성능 평가는 수학적 공식인 PSNR을 이용하여 제안 방법의 상대적 우수성을 검증하는 반면에, 입체영상의 3D 평가는 RGB 영상과 깊이맵으로부터 생성되는 입체영상에 DSCQS, SSCQS 등의 주관 평가(subjective evaluation)를 실행되는데, 이 과정은 많은 인적 자원 및 장기간의 실험을 필요로 한다. 상대적으로 수학적으로 구해지는 객관적 메트릭인 QA의 장점은 공식이 수학적으로 정의되어 있기 때문에, 빠른 시간내에 컴퓨터에서 결과를 얻을 수 있다는 것이다.

이렇게 깊이맵 평가 방법은 단순히 PSNR을 기반으로 하지만, 실제 이 깊이맵이 3D 입체영상과 어떠한 관계를 가지는 지에 대한 연구는 많지 않았다. 관련 연구로 [7]은 PSNR, 블러 메트릭(Blur metric) 및 선명도(sharpness degree)을 이용하여 객관적 방법과 3D 주관적 평가와의 관련성을 조사하였다. 만일 주관적 평가 결과를 예측할 수 있는 객관적 평가 메트릭이 존재한다면, 시간적 소요가 많은 주관적 평가없이 객관적 평가만으로 간단히 3D 입체감을 예측할 수 있다.

본 연구에서는 깊이맵 업샘플링의 응용이 필요한 분야에서 생성된 입체영상의 3D 성능을 객관적으로 예측할 수 있는 객관적 평가 방법을 제안한다. 이를 위해서 7 가지의 업샘플링 방법으로부터 얻어진 3D 평가 결과와 객관적 메트릭과의 상관관계(correlation)을 조사하여, 가장 상관관계가 높은 메트릭을 찾고자 한다. 즉, 이 메트릭은 주관적 평가와 유사한 결과를 주기 때문에, 인적 물적 자원을 필요로 하는

a) 강원대학교 컴퓨터정보통신공학과(Dept. of Computer and Communications Engineering, Kangwon National University)
‡ Corresponding Author : 김만배(Manbae Kim)
E-mail: manbae@kangwon.ac.kr
Tel: +82-33-250-6395
ORCID:http://orcid.org/0000-0002-4702-8276
※ 본 연구는 산업통상자원부/산업기술평가관리원의 산업융합원천기술개발사업(정보통신) [KI002058, 대화형 디지털 홀로그램 통합서비스 시스템의 구현을 위한 신호처리요소기술 및 하드웨어 IP 개발], 지식경제부 및 산업기술평가관리원의 시스템반도체 상용화기술사업 [10041082] 및 강원대학교의 지원을 받음.
※ 위 논문의 연구결과 중 일부는 “2014년도 한국방송공학회 추계학술대회”에서 발표한 바 있음.
· Manuscript received December 2, 2014; revised January 27, 2015; accepted January 27, 2015.

주관적 평가를 대체할 수 있다. 객관적 메트릭으로는 참조 메트릭 (full-reference metric) 및 무참조 메트릭 (no-reference metric)을 이용한다. 참조 메트릭으로는 PSNR, SSIM, 선명도(sharpness degree)가 있고, 무참조 메트릭 으로는 sharpness degree, blur metric, BIQI, NIQE 등이 있다. [7]에서 사용한 메트릭보다 더 우수한 성능을 주는 메트릭들을 추가하였고, 상관관계도 피어슨, 스피어만 및 켄달 상관계수를 포함하여 좀더 심도있는 연구를 진행한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안 평가 방법을 설명한다. 3장에서는 참조 및 무참조 객관적 평가 방법들을 소개하고, 4장에서는 실험에서 사용한 주관적 평가 방법을 설명한다. 5장에서는 3장과 4장에서 얻은 측정값들의 유사성을 알아내는 상관관계 방법을 설명하고, 6장의 실험에서는 가장 유사성이 높은 객관적 메트릭을 찾는다. 마지막으로 결론은 7장에서 정리한다.

Ⅱ. 제안 평가 방법

서론에서 언급했듯이, 본 연구의 목적은 업샘플링된 깊이맵으로 3D 입체영상을 생성할 때에, 3D 입체감 평가를 대체할 수 있는 객관적 메트릭를 찾는 것이다. 즉, 이 메트릭은 3D 입체감과의 상관관계가 높다.

그림 1은 본 연구의 수행 방법을 보여준다. RGB 영상과 저해상도 깊이맵(low-resolution depth map, LRD)이 주어지면, 업샘플링을 이용하여 고해상도 깊이맵(high-resolution depth map, HRD)을 얻는다. 모든 업샘플링 방법을 이용하는 것은 어려우므로, 실험에서는 서론에서 소개한 7가지 방법을 선택하여 사용한다. HRD에 객관적 메트릭를 적용하여 각 수치 값을 얻는다. DIBR(depth image based rendering)을 이용하여 RGB 영상과 HRD로부터 입체영상을 생성하고, 주관적 평가로[1][5] 스케일의 평가 점수를 구한다. 그림 1(a)처럼 QA 메트릭의 수치 값과 주관적 평가 점수와의 가장 우수한 상관관계(correlation)을 가지는 객관적 메트릭을 구하면, 이것은 그림 1(b)처럼 이 객관적 메트릭이 주관적 평가를 대체할 수 있다. 피어슨(Pearson), 스피어만(Spearman), 켄달(Kendall)의 상관관계는 각 객관적 결과와 주관적 QA 메트릭 사이의 유사성 측정을 위한 적절한 접근법들이다.

Ⅲ. 객관적 평가 방법

객관적 평가방법으로는 두 개의 큰 부류로 참조영상 평가(full-reference quality assessment, FRQA) 및 무참조 영

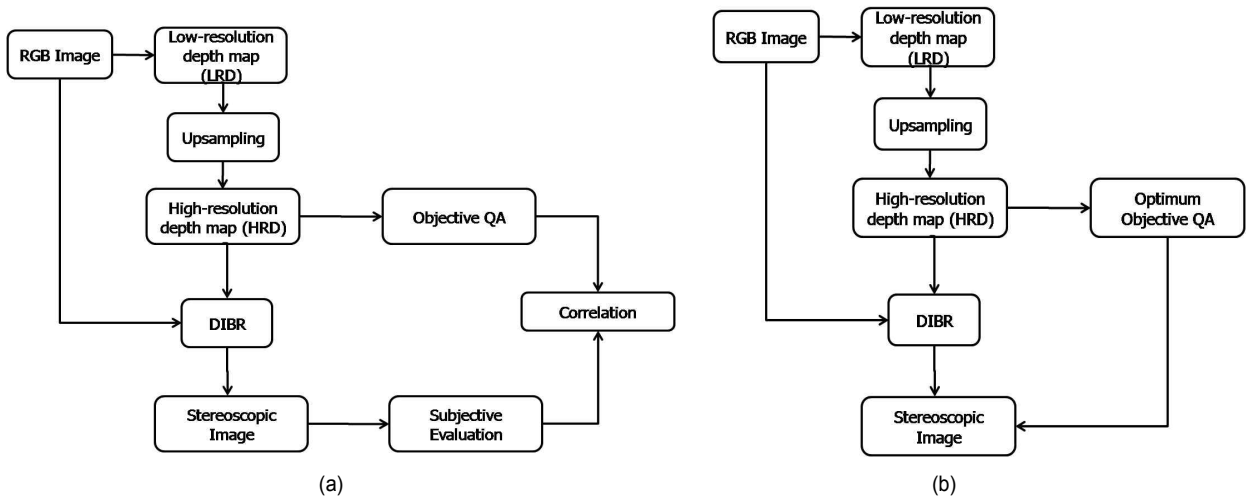

그림 1. 실험 방법의 블록도. (a) 두 측정 결과의 상관관계 계산 및 (b) 객관적 메트릭의 이용한 입체영상의 3D 입체감 측정
Fig. 1. Block diagram of the experiment methodology. (a) Correlation between two measurements and (b) measurement of 3D perception using objective metric

상 평가(no-reference quality assessment, NRQA)가 있다. 참조영상 평가에서는 업샘플링된 고해상도 깊이맵과 원 고해상도 깊이맵을 비교한다. 반대로 무참조 영상평가는 원 깊이맵에 대한 정보없이, 업샘플링된 고해상도 깊이맵만을 활용한다. 본 논문에서는 다음의 FRQA와 NRQA를 사용하여 업샘플링 방법들의 성능을 측정한다.

1. FRQA 방법

1.1 PSNR

PSNR(Peak Signal to Noise Ratio)는 성능 평가를 위한 가장 일반적인 도구 중 하나로 식 (1)로 정의된다.

$$PSNR = 10 \cdot \log\left[\frac{\sum (D^h - D^u)^2}{255^2}\right] \text{ (in dB)} \quad (1)$$

여기서 D^h와 D^u는 각각 원 고해상도 깊이맵과 업샘플링된 고해상도 깊이맵이다. 에지를 포함한 경계영역은 입체감에 큰 영향을 주기 때문에, PSNR은 전체 영상 PSNR, Edge PNSR, 및 Non-edge PNSR로 구분하여 실험에 사용한다.

1.2 SSIM

영상 품질을 평가하는 방법으로 두 영상간의 유사성을 측정하는 SSIM(structural similarity index measure)이 있다[8]. SSIM은 밝기 왜곡의 결합, 대비 왜곡, 상관관계의 손실처럼 영상 왜곡 모델링을 바탕으로 한다. 두 영상 f, g의 SSIM은 다음과 같이 표현된다.

$$\mathrm{SSIM} = \mathrm{l(f,g)c(f,g)s(f,g)}, \quad (2)$$

여기서,

$$l(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}, c(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2}, s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f + \sigma_g + C_3}$$

l(f,g), c(f,g) 및 s(f,g)는 밝기, 대조와 구조 비교 함수들이다. σ_f와 σ_g는 표준편차를 나타내며,μ_f와 μ_g는 평균값이고, σ_{fg}는 공분산(covariance)이다. C_1, C_2 및 C_3는 분모가 0이 되지 않도록 더해주는 상수이다. SSIM은 [0,1]의 값을 가지고, 1에 가까울수록 유사성이 높다.

1.3 VIF

VIF(Visual Information Fidelity)는 영상 충실도(fidelity) 측정을 위해 정보이론 기준을 사용한 참조영상 평가 지표이다[9]. 참조영상으로부터 이상적으로 추출될 수 있는 정보와 왜곡으로 인한 정보의 손실은 자연적 장면 통계(natural scene statistics), HVS(human visual system) 및 영상왜곡 채널로 모델링한다.

2. NRQA 평가 방법

2.1 선명도

선명도(Sharpness degree)[10]는 영상의 선명도를 나타내기 위해 사용되며 식 (3)으로 정의된다.

픽셀 $D(x,y)$에 대한 좌측 픽셀 $D(x-1,y)$과의 차이와 상단 픽셀 $D(x,y-1)$과의 차이의 제곱 및 평균을 냄으로써 선명도를 계산한다.

2.2 블러 메트릭

블러(blur)의 양을 측정하는 객관적 화질평가인 블러 메트릭(blur metric)은 에지의 시작과 끝 사이의 폭을 측정함으로써 블러의 정도를 추정하는 방법이다[10]. 먼저 깊이맵으로부터 에지를 검출한 후에, 각 에지 픽셀에서 국부 최대(local maximum)를 갖는 픽셀의 위치 p_s, 국부 최소(local minimum)를 갖는 픽셀의 위치 p_f를 이용하여 에지의 폭 $w_k = \|p_s - p_f\|$를 계산한다. 에지는 1 픽셀 이상의 폭을 가진다. 모든 스캔라인에서 w_k를 계산한 후에, 블러 메트릭 BM은 다음 식 (4)와 같이 w_k의 합으로 구한다.

$$\mathrm{SD} = \frac{1}{MN}\sum_{y=0}^{M-1}\sum_{x=0}^{N-1}[(D(x,y) - D(x-1,y))^2 + (D(x,y) - D(x,y-1))^2] \quad (3)$$

$$\mathrm{BM}=\sum_{i=1}^{M}\sum_{k=1}^{K}w_{k,i}\,/\,\sum_{i=1}^{M}K_i \tag{4}$$

2.3 BIQI

BIQI(blind image quality index)은 영상에서 왜곡이 식별되면, NSS를 기반으로 이 왜곡을 수치화한다[11]. 왜곡 영상이 주어지면, 먼저 알고리즘은 JPEG, JPEG2000, white noise(WN), Gaussain Blur(Blur) 및 Fast fading(FF)으로 이루어진 영상에서 왜곡의 유무를 추정한다. 영상에서 각 왜곡의 확률은 p_i로 나타낸다. 영상의 품질은 확률 및 가중치의 합으로 표현된다.

$$\mathrm{BIQI}=\sum_{i=1}^{5}p_i\cdot q_i, \tag{5}$$

여기서 q_i는 각각 5개 왜곡영상에 상응하는 5개의 품질 평가에서 얻은 수치이다.

2.4 NIQE

NIQE(natural image quality evaluator)은 자연영상에서 관찰되는 통계적 정형화로부터 측정가능 편차만 사용하는 무참조 평가방법이다[12]. BIQI와 NIQE의 수치는 [0,100]의 값으로 표현되는데, 0에 근접할수록 우수한 품질을 의미한다.

Ⅳ. 주관적 3D 입체감 평가

DIBR로 만들어진 입체영상의 3D 입체감 평가는 주관적 평가실험을 이용한다. 42″3DTV로 입체영상을 시청하였고 DSCQS(Double Stimulus Continuous Quality Scale)를 적용하였다[14],[15]. 평가실험에는 12명의 피실험자가 참가하였다. 이중 7명은 입체영상 시청 경험이 있고, 나머지 5명은 경험이 없는 피실험자이다. 주관적 3D 성능평가는 크게 3D 입체감 평가 및 시각적 피로도 평가로 나누어진다. 그러나, 원 고해상도 깊이맵과 업샘플링으로 획득한 고해상도 깊이맵은 유사한 깊이값을 가지고 있기 때문에 작은 차이는 입체감에는 큰 영향을 미치지 못한다. 그러나 작은 차이라 할지라도 피로도에는 큰 영향을 미칠 수 있기 때문에 실험에서는 시각적 피로도(visual fatigue)를 이용하여 측정한다. 피실험자들은 원 깊이맵에 의해 생성된 입체영상과 실험 업샘플링 방법으로부터 얻어진 깊이맵에 의해 생성된 입체영상을 약 10초동안 번갈아 보면서 평가 점수를 기록한다. 여기서 전자의 입체영상을 5.0을 기준으로 하고, 상대적인 평가를 하도록 하였다. 표 1은 주관적 평가에 사용되는 점수표를 보여준다.

표 1. 시각적 피로도 평가
Table 1. Visual fatigue evaluation

평가단계	점수
피로함 없음	5
약간 (slight) 피로함	4
조금 (moderate) 피로함	3
피로함	2
매우 (severe) 피로함	1

Ⅴ. 상관 관계 측정

3장의 객관적 평가 방법의 결과와 4장에 얻은 3D 주관적 평가의 연관성을 조사하기 위해 다음 3개의 상관관계를 이용한다.

1. 피어슨 상관 관계 (Pearson Correlation)

두 변수의 연관성을 측정하기 위한 보편적인 지표로 피어슨 상관 계수(Pearson's correlation coefficient)가 있다. 피어슨의 방법은 [-1,1]의 범위를 가지는 계수를 사용하여 상관 관계의 크기를 계산한다. 양의 계수는 두 개의 변수의 비례 관계를 의미하고, 음의 값은 반비례 관계를 의미한다. 두 변수간의 비례관계가 없으면 0의 값을 가진다. 피어슨 계수 ρ_p는 다음과 같이 정의된다.

$$\rho_p=\frac{Cov(x,y)}{\sigma_x\sigma_y} \tag{6}$$

$$\rho_s = \frac{\sum_{i=1}^{n}((rank(x_i) - \overline{rank(x)})((rank(y_i) - \overline{rank(y)})}{\sqrt{\sum_{i=1}^{n}((rank(x_i) - \overline{rank(x)})^2 \sum_{i=1}^{n}((rank(y_i) - \overline{rank(y)})^2}} \quad (7)$$

여기서 Cov(x,y)는 두 그룹간의 공분산이며, σ_x와 σ_y는 표준 편차이다.

2. 스피어만 상관관계 (Spearman Correlation)

이 상관계수는 피어슨 상관계수의 순위기반 형태이다. n개의 변수를 가지는 각 그룹의 표본들은 가장 작은 값은 1로, 가장 큰 표본은 n으로, 1에서 n까지 순위가 매겨진다. 스피어만 상관계수 ρ_s는 다음과 같이 계산된다.

$$\tau = \frac{\sum_{i=1}^{n}\sum_{i=1}^{n} sgn(x_i - x_j) sgn(y_i - y_j)}{n(n-1)} \quad (8)$$

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

여기서 $rank(x_i)$와 $rank(y_i)$은 표본에서의 순위이다.

3. 켄달 타우 상관관계 (Kendall's Tau Correlation)

켄덜 타우(τ)의 범위는 [-1, +1]이고, 이전의 상관계수와 비슷하게 두 변수 관계의 강도를 계산한다. 이 상관계수는 일치하는 수와 불일치하는 쌍의 차이를 수량화한다.

Ⅵ. 실험 결과

업샘플링 방법의 성능은 MSR, Middlebury[13], HHI, 및 GIST의 16 개의 실험 깊이맵을 사용하여 평가하였다. 실험 RGB 영상과 깊이맵들은 그림 2에서 보여진다. 저해상도 깊이맵을 획득하기 위하여, 원본 데이터를 다운샘플링(downsampling)하고, 다시 업샘플링하여 고해상도 깊이맵을 만들었다. 그림 3 및 4는 Bowling, Ballet에 대해서 7가지 방법의 업샘플링된 깊이맵을 보여준다.

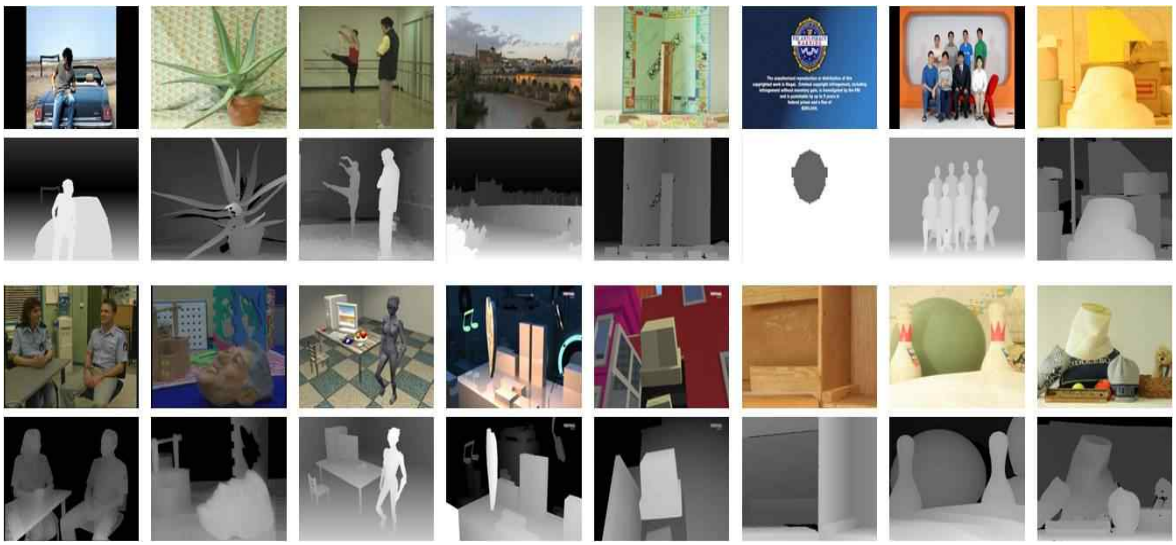

그림 2. 실험 RGB 영상과 깊이맵. 실험데이터는 MSR, Middlebury, GIST 및 HHI에서 제공 받았음.
Fig. 2. RGB images and related depth maps. (Provided by MSR, Middlebury, GIST and HHI)

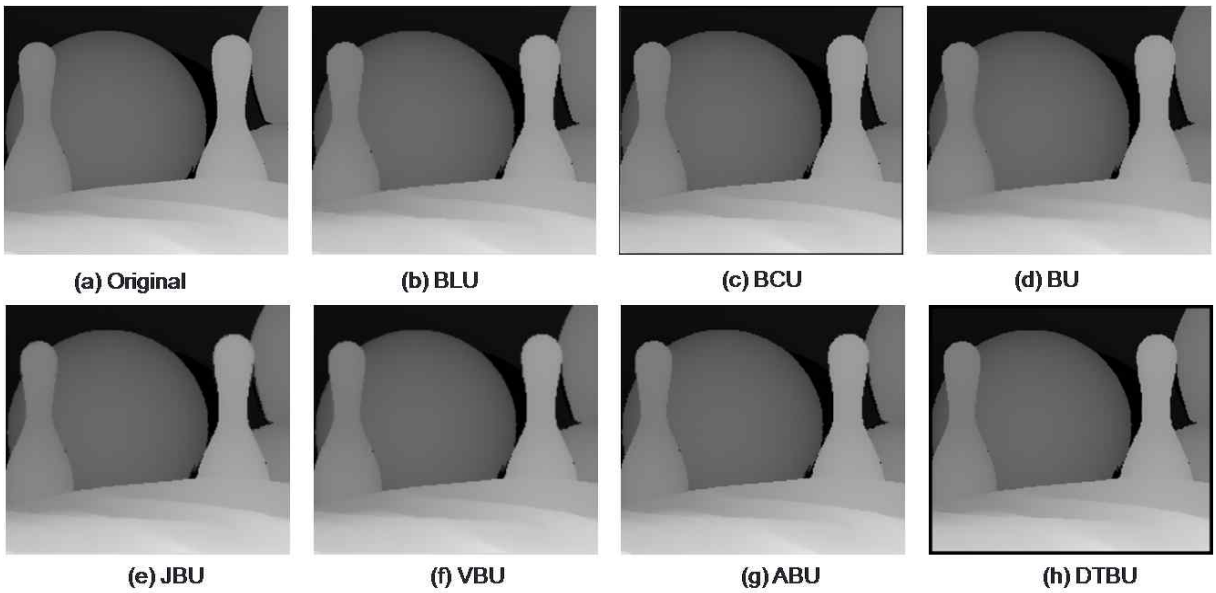

그림 3. Middlebury Bowling 깊이맵의 업샘플링된 깊이맵. (a)는 원깊이맵이고 (b)~(h)는 7개의 업샘플링방법으로 얻어진 깊이맵임. (BLU는 양선형 보간법, BCU는 고등차수 보간법, BU는 양측 업샘플링, JBU는 결합 양측 업샘플링, VBU는 분산 기반 양측 업샘플링, ABU는 적응적 양측 업샘플링, 그리고 DTBU는 거리변환 기반 양측 업샘플링임)

Fig. 3. Upsampled depth maps of Middlebury Bowling using seven upsampling methods. (a) is an original depth map and (b)~(h) are upsampled depth maps obtained by seven methods. (BLU= bilinear upsampling, BCU=bicubic upsampling, BU= bilateral upsampling, JBU=joint bilateral upsampling, VBU=variance-based bilateral upsampling, ABU=adaptive bilateral upsampling, and DTBU=distance transform-based bilateral upsampling)

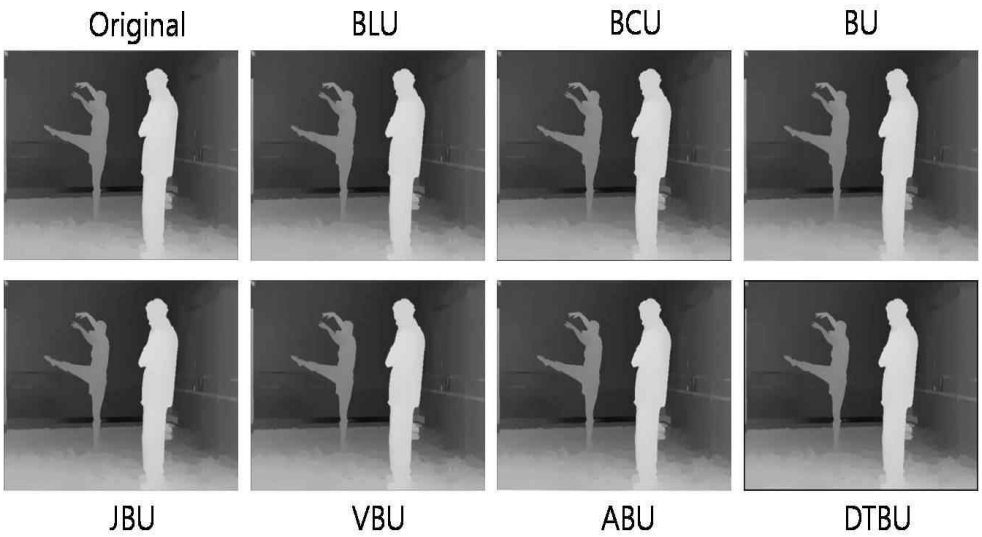

그림 4. MSR Ballet 깊이맵의 업샘플링된 깊이맵. (a)는 원깊이맵이고 (b)~(h)는 7개의 업샘플링방법으로 얻어진 깊이맵임.

Fig. 4. Upsampled depth maps of MSR Ballet using seven upsampling methods. (a) is an original depth map and (b)~(h) are upsampled depth maps obtained by seven methods

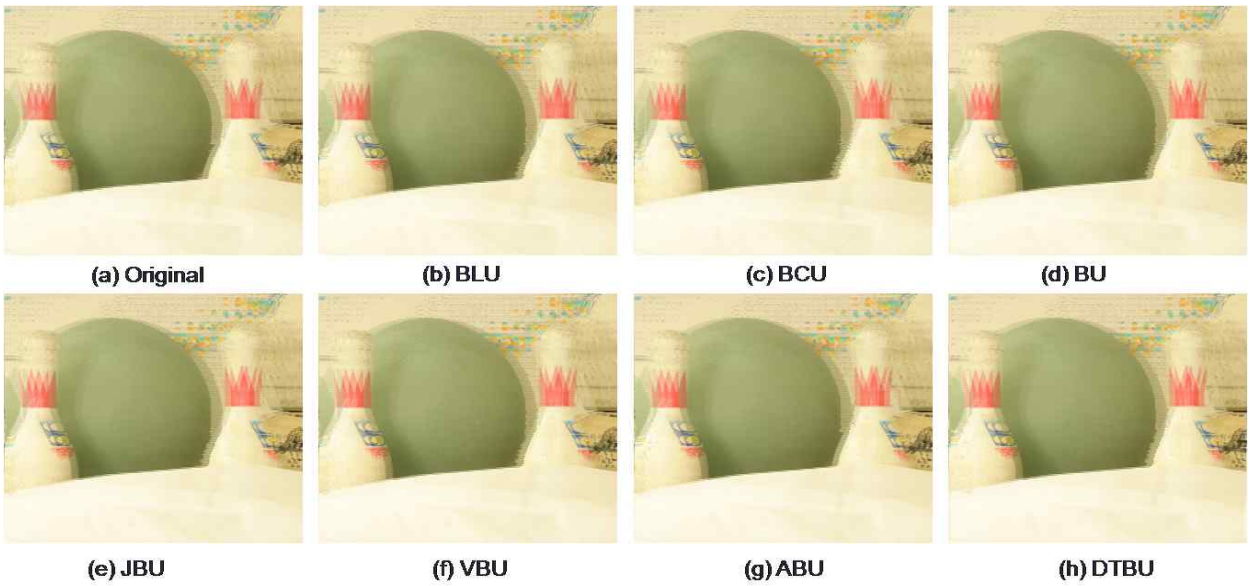

그림 5. 원 RGB 영상과 업샘플링된 깊이맵으로부터 생성된 좌우 입체영상
Fig. 5. Stereoscopic images in interlaced format generated by original RGB images and upsampled depth maps.

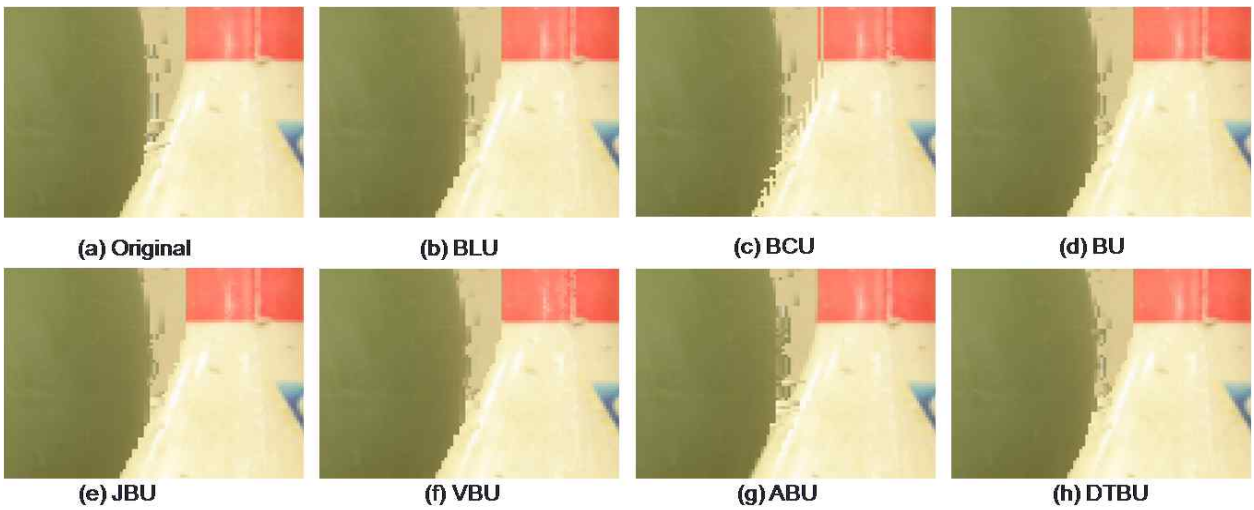

그림 6. 그림 5의 확대 영상
Fig. 6. Close-ups of images in Fig. 5

그림 5는 DIBR로 만들어진 입체영상을 보여준다. 그림 5에 있는 영상의 하단 부분을 확대한 영상이 그림 6에서 보여지는데, 원 깊이맵과 달리 왜곡이 발생하는 것을 관찰할 수 있다. 이러한 왜곡은 입체영상의 화질 저하로 시각적 피로도를 유발하게 된다. 업샘플링된 고해상도 깊이맵에 대해, Image PSNR, Edge PSNR, Non-edge PSNR, Sharpness Degree (SD), Blur Metric (BM), SSIM, VIF, BIQI, NIQE의 객관적 평가 수치를 측정하였고, 16장의 평균값은 표 2에서 보여진다.

표 2. 객관적 평가의 평균 결과. PSNR은 dB임. (BLU는 양선형 보간법, BCU는 고등차수 보간법, BU는 양측 업샘플링, JBU는 결합 양측 업샘플링, VBU는 분산 기반 양측 업샘플링, ABU는 적응적 양측 업샘플링, 그리고 DTBU는 거리변환 기반 양측 업샘플링임)
Table 2. Average results of objective QA metrics. (BLU= bilinear upsampling, BCU=bicubic upsampling, BU= bilateral upsampling, JBU=joint bilateral upsapling, VBU=variance-based bilateral upsampling, ABU=adaptive bilateral upsampling, and DTBU=distance transform-based bilateral upsampling)

Depth map	BLU	BCU	BU	JBU	VBU	ABU	DTBU
Image PSNR	35.85	35.71	35.64	34.15	35.64	33.16	34.86
Edge PSNR	23.68	23.55	23.66	22.82	23.38	20.97	22.93
Non-edge PSNR	38.07	37.94	37.78	37.50	37.93	35.43	36.92
Sharpness Degree	39.5	42.2	49.51	49.09	31.92	88.31	68.14
Blur Metric	8.48	11.38	10.29	10.87	10.51	9.00	9.89
SSIM	0.946	0.955	0.975	0.956	0.971	0.962	0.972
VIF	0.518	0.539	0.424	0.422	0.478	0.398	0.438
BIQI	57.8	66.34	61.11	32.81	41.94	29.15	72
NIQE	15.95	13.11	13.94	11.82	12.47	13.41	13.82

표 3. 주관적 평가의 평균 결과. scale=[1,5]임.
Table 3. Average results of 3D subjective test at scale [1,5]

	BLU	BCU	BU	JBU	VBU	ABU	DTBU
visual Fatigue	3.76	3.64	3.89	3.84	4.03	3.46	3.99

표 3은 주관적 평가에 대한 각 업샘플링 방법의 평균 평가 점수를 보여준다. 시각적 피로도 값은 3.76(BLU), 3.64(BCU), 3.89(BU), 3.84(JBF), 4.03(VBU), 3.46(ABU) 및 3.99(DTBU)이다. 모든 값들은 0에서 1사이로 정규화되며, 피어슨, 스피어만, 켄달 상관계수 값은 표 4에서 보여준다. 그림 7은 다이어그램 형태로 평가 메트릭의 상관계수 값을 보여준다.

표 4에 따르면, 3개의 PSNR중에서 Edge-PSNR이 PSNR과 non-edge PSNR 보다 더 높은 상관관계를 가진다. 이 결과는 에지에서의 깊이맵 품질이 시각적 피로도와 밀접한 관계가 있다는 것을 알려준다. 선명도와 블러 메트릭은 각각 음수 및 양수 상관관계 값을 보여준다. 이 결과는 높은 고주파(high frequency) 성분은 시각적 불편의 원인이라는 [14]와 실험 결과와 일치한다. SSIM은 PSNR과 유사하게 피어슨 상관계수는 다른 두 상관계수보다 더 높다. SSIM은 메트릭중에서 가장 높은 스피어만 계수 값을 가지며, 3개의 상관관계의 합도 가장 높다. 그러므로 시각적 피로도와 가장 상관관계가 높다는 것을 알 수 있다. VIF 결과는 시각적 피로도와 낮은 상관관계를 보인다. BIQI는 선명도처럼 음의 값을 가지고 있다. 이는 BIQI가 낮을수록, 시각적 피로도는 개선되는 역비례를 가진다. NIQI의 상관관계는 블러 메트릭과 VIF와 유사한 값을 보여준다. 따라서 상기 3개는 시각적 피로도와 연관성은 낮다고 할 수 있다. 실험 결과로 부터 주관적 평가를 대체할 수 있는 객관적 메트릭은 SSIM 및 Edge-PSNR이라고 할 있다.

표 4. 시각적 피로도와 객관적 메트릭의 피어슨, 스피어만, 켄달 상관관계
Table 4. Correlations of Pearson, Spearman and Kendall for visual fatigue and objective metrics

	Pearson	Spearman	Kendall	Sum
PSNR	0.582	0.0357	0.0476	0.6653
Edge-PSNR	0.608	0.1429	0.1429	0.8938
Non-edge PSNR	0.554	0.0357	0.0476	0.6373
Sharpness Degree	-0.522	-0.3214	-0.1429	-0.9863
Blur Metric	0.293	0.1429	0.0476	0.4835
SSIM	0.505	0.3571	0.1429	1.005
VIF	0.019	0.1071	0.1429	0.269
BIQI	-0.339	-0.3214	-0.2381	-0.8985
NIQE	0.132	0.1071	0.1429	0.382

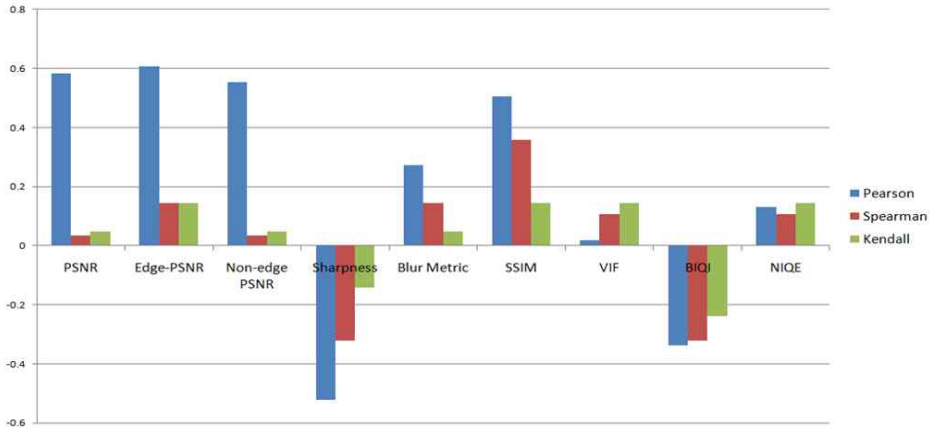

그림 7. 피어슨, 스피어만, 켄달 상관계수의 바 다이아그램
Fig. 7. Bar diagram of correlations of Pearson, Spearman and Kendall

Ⅶ. 결론 및 향후 연구

3D 입체영상은 그동안 주관적 평가에 의해 3D 입체감 및 시각적 피로도를 측정해왔다. 객관적 메트릭으로 영상의 시각적 평가를 대체하는 연구는 2D 비디오 분야에서는 오랜 동안 진행되었던 것과는 달리, 3D 분야에서는 아직 괄목할만한 연구 성과가 나오지 않고 있다. 본 논문에서는 깊이맵 업샘플링을 이용하여 생성된 입체영상의 주관적 입체감 평가를 대체할 수 있는 객관적 메트릭을 조사하였다. 다양한 객관적 메트릭으로부터 얻어진 객관적 수치와 주관적 시각적 피로도 점수를 얻은 후에, 세 가지 상관관계를 통해서 SSIM이 가장 주관적 평가점수와 근접하다는 것을 얻을 수 있었다. 따라서 향후 인적 물적 자원이 요구되는 주관적 평가의 실험을 이 SSIM이 대체할 수 있다는 결론을 얻었다.

본 논문에서 조사한 객관적 기법이외에도 다음과 같은 객관적 화질 평가 기법이 있다: IW-PSNR, NQM, UQI, MS-SSIM, IWSSIM, IFC, VIF, M-SVD, PSNR-HVSM, VSNR, MAD, FSIM, GSM, IGM, GMSD, 및 VSQA. 이 메트릭들의 심층 조사 및 단일 메트릭들의 혼합 방법도 중요한 연구 주제이다.

참 고 문 헌 (References)

[1] A. B. Dehkordi, M. T. Pourazad, P. Nasiopoulos, "A study on the relationship between depth map quality and the overall 3D video quality of experience," IEEE 3DTV Conference, 2013.
[2] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Image", In Proc. IEEE Int. Conf. on Computer Vision, pp.836-846(1998).
[3] J. Kopf, M. F. Cohen, D. Lischinski and M. Uyttendaele, "Joint bilateral upsampling", ACM Trans. on Graphics, Vol. 26, No. 3, pp.1-6, 2007.
[4] C. Pham, S. Ha, and J. Jeon, "A local variance-based bilateral filtering for artifact-free detail and edge-preserving smoothing", PSIVT, Part Ⅱ, LNCS 7088, 99.60-70, 2011.
[5] D. Yeo, E. Haq, J. Kim, M. Baig, H. Shin, "Adaptive Bilateral Filtering for Noise Removal in Depth Upsampling", SoC Design Conf., pp. 36-39, 2010.
[6] S. Jang, D. Lee, S. Kim, H. Choi, M. Kim, "Depth Map Upsampling with Improved Sharpness", Journal of Broadcast Engineering, Vol. 17, No. 6, pp. 933-944, 2012.
[7] J. Gil, S. Mahmoudpour and M. Kim, "Analysis of relationship between objective performance measurement and 3D visual discomfort in depth map upsampling", Journal of Broadcast Engineering, Vol. 19, No. 1, pp. 31-43, Jan. 2014.
[8] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Trans. On Image Processing, Vol. 13, No. 4, 2004.
[9] H. R. Sheikh and A. C. Bovik, "Image information and image quality", IEEE Trans. On Image Processing, Vol. 12, No. 2, pp. 430-444, 2006.
[10] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics:applicaation to JPEG2000", Int. Workshop on Multimedia Signal Processing, pp. 403-408, 2008.
[11] A. K. Moorthy and A. Bovik, "A two-step framework for constructing

blind image quality assessment", IEEE Signal Processing Letters, Vol. 17, No. 5, pp. 513-516, 2010
[12] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind' image quality analyzer", IEEE Signal Processing Letters, Vol. 20, No. 3, pp. 209-212, 2013.
[13] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", Int. J. of Computer Vision, Vol. 47, No.1-3, pp. 7-42, 2002.
[14] E. Lee, H. Heo, and K. Park, "The comparative measurements of eye strain caused by 2D and 3D displays", IEEE Trans. on Consumer Electronics, Vol. 56, No. 3, pp. 1677-1683, 2010.
[15] D. Kim and K. Sohn, "Visual fatigue prediction for stereoscopic image", IEEE Trans. Circuits Syst. Video Technol., Vol. 21, No. 3, pp. 231-236, 2011.

저 자 소 개

사이드 마흐모드포어

- 2007년 9월 : B.S degree, Islamic Azad University, Iran
- 2011년 9월 : M.S degree, Islamic Azad University, Iran
- 2013년 3월 ~ 현재 : Ph.D student, Kangwon National University, Repubic of Korea
- 주관심분야 : 3D image processing, 3D human factor, human detection

최 창 열

- 1979년 : 경북대학교 전자공학과 학사
- 1981년 : 경북대학교 전자공학과 석사
- 1995년 : 서울대학교 컴퓨터공학과 공학박사
- 1984년 ~ 1996년 : ETRI 컴퓨터연구단 책임연구원 / 연구실장
- 2009년 ~ 2011년 : 강원대학교 IT대학 학장
- 1996년 ~ 현재 : 강원대학교 컴퓨터정보통신공학과 교수
- ORCID : http://orcid.org/0000-0002-8340-4195
- 주관심분야 : 모바일전송, 3D데이터처리, 미디어서비스

김 만 배

- 1983년 : 한양대학교 전자공학과 학사
- 1986년 : University of Washington, Seattle 전기공학과 공학석사
- 1992년 : University of Washington, Seattle 전기공학과 공학박사
- 1992년 ~ 1998년 : 삼성종합기술원 수석연구원
- 1998년 ~ 현재 : 강원대학교 컴퓨터정보통신공학과 교수
- ORCID : http://orcid.org/0000-0002-4702-8276
- 주관심분야 : 3D영상처리, 깊이맵처리, 입체변환