

포아송 객체 편집을 이용한 스테레오 영상 합성

백 으 뜸*, 호 요 성^o

Stereo Image Composition Using Poisson Object Editing

Eu-Tteum Back*, Yo-Sung Ho^o

요 약

본 논문에서는 포아송 영상 편집을 이용한 스테레오 3차원 영상 합성 방법을 제안한다. 3차원 영상을 만들 때 객체의 깊이 값을 고려하지 않고 색상 영상을 합성하면 매우 부자연스러운 결과를 얻는다. 본 논문에서는 그래프 컷 방법을 사용하여 추출한 전경 객체를 새로운 배경 영상에 포함시켜 자연스러운 2차원 영상을 만든다. 제안된 방법에서는 기하학적인 위치뿐 아니라 색상 톤도 고려하여 자연스러운 영상을 합성한 후, 합성된 영상의 경계선 주변에 블러링을 수행하여 전경과 배경이 잘 어우러지게 만든다. 이와 더불어, 임의로 주어진 객체의 깊이 값을 이용하여 추출된 객체의 변위를 구하고, 오른쪽 색상 영상과 깊이 영상을 자동으로 합성한다. 실험을 통해, 제안하는 방법이 자연스러운 스테레오 3차원 영상을 생성할 수 있음을 확인했다.

Key Words : Poisson image editing, Grabcut, depth map composition, blurring, stereo image composition

ABSTRACT

In this paper, we propose a stereo image composition method based on Poisson image editing. If we synthesize images without considering their depth values, it may lead to unwanted consequences. When we segment an image into its background and foreground regions using Grabcut, we take into account their geometric positions to mix color tones; thus, the image is composited more naturally. After synthesizing images, we apply a blurring operation around object boundaries; then, the foreground object and background are composited more seamlessly. In addition, we can adjust the distance of the object by setting arbitrary depth values and generating right color and depth images automatically. Experimental results show that the proposed stereo image composition method provides naturally synthesized stereo images. Improved portions were subjectively confirmed as well.

1. 서 론

영상을 합성할 때 사용자가 직접 객체의 경계를 선택해 주는 일은 어려우며, 이를 이용하는 것도 많은 시간이 요구된다. 이전에 제안된 영상 합성 방법들은 대부분 특정 상황을 고려했지만, 객체의 기하학적 위치까지는 고려하지 않았다. 최근 이들의 문제점을 보

완하는 방법이 연구되고 있다.

Patrick Pérez가 제안한 포아송 영상 편집 방법(Poisson Image Editing)은 배경영상(Background image)의 관심영역(Region of Interest: ROI)으로 전경영상(Foreground image)의 일부분을 자연스럽게 붙여 넣는 영상보간 기법으로, 이차 편미분 방정식인 포아송 방정식을 사용한다^[1,2]. 비록 포아송 영상 편집

* First Author : Department of Information and Communications Gwangju Institute of Science and Technology (GIST), eutteum@gist.ac.kr, 학생회원

^o Corresponding Author : Department of Information and Communications Gwangju Institute of Science and Technology (GIST), hoyo@gist.ac.kr, 종신회원

논문번호 : KICS2014-03-096, Received March 19, 2014; Revised May 23, 2014; Accepted August 4, 2014

방법의 성능이 우수하지만, 몇 가지 약점이 있다. 만약 합성하려는 배경영상의 색상이 복잡하거나 배경과 전경의 색상 차이가 크면 좋은 합성 결과를 얻을 수 없다. Rother가 제시한 그래프 컷(Graph cut) 방법을 이용한 이진영상 분할 방법이다^[3]. 이 방법은 사용자가 객체를 포함하는 영역을 설정하면, 선택된 영역의 외부를 배경으로 설정하고, 내부 영역과 외부 영역에서 GMM (Gaussian Mixture Model)을 추정한 후, 그래프 컷 기법을 반복적으로 적용한다.

이 논문에서는 단일 색상영상에서 관심이 있는 전경을 추출하여 스테레오 영상에 자연스럽게 합성하는 방법을 제안했다. 배경영상은 스테레오 색상 정보와 그에 상응하는 깊이 정보를 포함한다. 이 논문에서 제안하는 방법은 그래프 컷을 활용하여 객체를 추출하고, 임의로 깊이 값을 정하여 추출된 객체를 우영상으로 옮긴다. 좌우영상에서 각각 깊이 정보를 이용한 포아송 영상 편집을 적용하여 전경과 배경을 합성한다. 보다 자연스러운 영상 합성을 위해서 후처리 과정으로 경계 영역을 흐리게 하는 블러링 방법을 추가했다.

II. 스테레오 매칭 방법

제안하는 영상 합성 방법을 사용하기 위해서는 색상영상 뿐만 아니라 깊이영상이 필요하다. 좌우영상이 주어지면 스테레오 정합을 이용하여 깊이 정보를 구할 수 있다. 좌우영상으로부터 깊이 정보를 얻는 과정은 영상 획득, 특징 추출, 스테레오 정합, 변위 추정, 변위로부터의 거리 계산 등으로 이루어지는데, 특징으로 사용될 정합요소의 선택과 그 요소에 적절한 정합 전략을 정하는 것이 중요하다.

스테레오 정합 방법은 화소의 변위를 계산하기 위해 정합 에너지 함수를 정의하고, 정합 오차를 계산하면서 정합 오차가 가장 작은 값을 가질 때의 최적의 변위를 구한다. 보통 스테레오 정합 에너지 함수는 두 가지 요소를 포함하는데, 첫 번째는 색상의 유사성이고, 두 번째는 주변 화소들과의 변위 변화도이다. 이 두 가지 요소는 정합 에너지 함수에서 데이터 항과 평활화 항으로 표현된다.

$$D(x, y) = \operatorname{argmin}_d \{E(x, y, d)\} \quad (1)$$

$$E(x, y, d) = E_{data}(x, y, d) + E_{smooth}(x, y, d) \quad (2)$$

여기서 $D(x, y)$ 는 화소 (x, y) 의 변위를 말하고, 정

합함수 $E(x, y, d)$ 는 데이터 항 $E_{data}(x, y, d)$ 과 평활화 항 $E_{smooth}(x, y, d)$ 로 구성된다. 데이터 항과 평활화 항은 다음과 같이 정의된다.

$$E_{data}(x, y, d) = U\{I_L(x, y), I_R(x - d, y)\} \quad (3)$$

$$E_{smooth}(x, y, d) = \sum_{(p, q) \in N} W(p, d) \quad (4)$$

여기서 $U(I_L(x, y), I_R(x - d, y))$ 는 좌영상과 우영상의 색상값 차이를 나타내는데, 거리차 제곱 합계(SSD: Sum of Squared Differences), 절대 거리차 합계(SAD: Sum of Absolute Differences), 평균 상관 계수(NCC: Normalized Cross Correlation) 등이 사용된다. 또한, N은 주변 화소의 집합을 말하며, $W(p, d)$ 는 주변 화소들과의 변위 차이를 나타낸다.

III. 스테레오 영상 합성 방법

그림 1은 이 논문에서 제안하는 알고리즘의 전체 과정을 보여준다. 사용자는 그래프 컷 알고리즘을 사용하여 단일 색상영상의 전경과 배경을 분리한 후, 추출된 전경의 영역과 깊이 값을 이용하여 배경의 깊이영상과 같이 합성한다. 단일 색상영상을 합성할 때 배경의 깊이 값을 이용하여 포아송 영상 편집을 적용하며, 합성된 영상에 전경의 경계선 영역을 얻어 후처리 작업을 수행한다. 좌영상이 합성된 뒤에 자동적으로 우영상의 깊이영상과 색상영상이 합성된다.

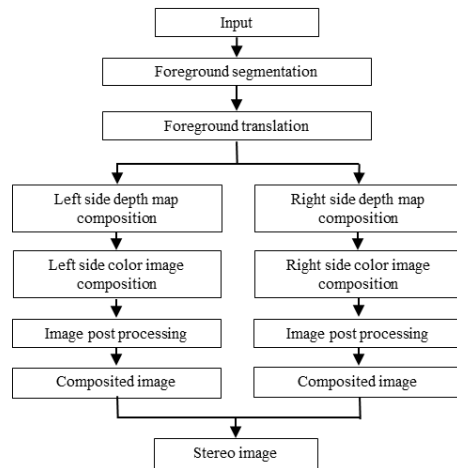


그림 1. 알고리즘 과정 흐름도
Fig. 1. Flowchart of the whole process

3.1 그랩컷과 전처리 방법

그랩컷 알고리즘은 사용자가 객체를 포함하는 영역을 지정하여 트라이 맵(trimap) T 를 얻고, 영상의 객체와 배경영역의 분포를 나타내기 위해 GMM (Gaussian Mixture Model)을 구성한다^[3].

그림 2는 그랩컷과 트라이 맵을 보여 주는데, 물체와 배경의 GMM은 트라이 맵 영역의 배경영역인 T_B , 객체와 배경이 불분명한 영역인 T_U , 객체 영역인 T_F 로부터 각각의 가우시안 요소(Gaussian components)로 구성된다. 그 다음, GMM을 갱신하고 윤곽선 정보 패널티를 기반으로 그래프를 형성한다. 그래프 컷을 활용하여 에너지 함수를 최적화하는 가중치를 구하는데, 가중치에 따라 분류되지 않은 화소들에 대해서 양방향 상태전이를 수행하여 객체나 배경으로 분류한다.

그랩컷을 적용할 때 전경 주위에 잡음이 생긴다. 전경 주위의 잡음을 제거하기 위해 마스크를 만들어 문턱치 기법과 모폴로지 처리인 침식(Erosion)과 팽창(Dilation)을 사용했다. 침식을 실행한 후 팽창의 순서로 모폴로지 기법을 영상에 적용한다. 침식과 팽창은 실행하는 순서가 달라지면 영상의 전처리 결과에 큰 영향을 주기 때문에 순서를 잘 지켜야 한다.

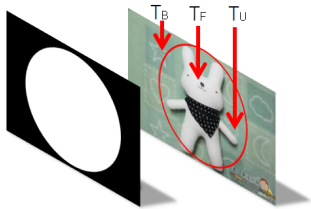


그림 2. 그랩컷과 트라이 맵
Fig. 2. Grabcut and Trimap

3.2 깊이 값과 변위의 관계를 이용한 전경 이동

이 논문에서는 스테레오 색상영상에 스테레오 정합 방법을 적용하여 깊이영상을 구할 때 좌영상을 기준 영상으로 삼는다. 추출된 전경영상에 임의의 깊이 값을 선택하여 좌영상에서 우영상으로 전경을 이동시킨다. 이때 전경영상에서 임의로 선택된 깊이 값이 우영상의 변위(disparity) 값으로 변환되며, 이를 이용하여 전경영상을 좌측으로 이동시킨 후, 단일 색상영상과 단일 깊이영상에 합성한다. 변위는 좌측과 우측 눈에 보이는 객체 위치 사이의 차이를 나타낸다.

임의로 정해진 깊이 값을 가지고 식 (5)에서 식 (8)로 바꿀 수 있다^[4]. 식 (5)와 식 (6)을 합쳐서 식 (7)과 식 (8)을 만든다. 그림 3에서 b 는 두 카메라의 중심 사이의 거리이고, f 는 초점거리, d 는 변위, 그리고 z 는

실제 깊이 값이다. 식 (7)에 근거하여 영상의 깊이 값은 변위와 같은 의미로 사용할 수 있다.

$$z = \frac{bf}{(x'_l - x'_r)} \tag{5}$$

$$d = x'_l - x'_r \tag{6}$$

$$z = \frac{bf}{d} \tag{7}$$

$$d = \frac{bf}{z} \tag{8}$$

식 (9)에서 dx 는 변위이고, T_d 는 변위 이동변환 행렬이다. 식 (9)를 활용하여 객체를 변위만큼 움직인다.

$$T_d = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{9}$$

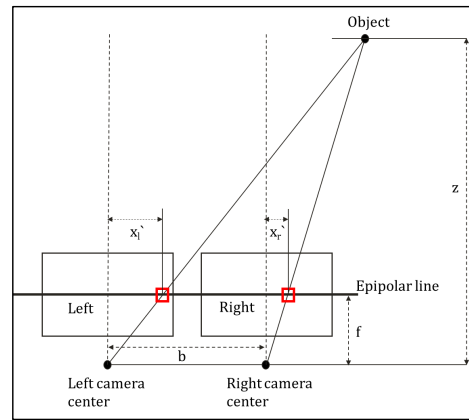


그림 3. 깊이 값과 변위에 관련된 그림
Fig. 3. Diagram about relationship between depth value and disparity

3.3 포아송 영상 편집

기존의 영상 편집 방법들은 합성 영역의 경계선 부분에서 최대한 부드러운 이음매가 형성되도록 한다. 만약 영상 합성시 전경영상과 배경영상을 직접 붙이면 배경과 전경영상의 환경이 서로 다르고 색상톤의 차이로 인해 이질감을 줄 수 있다.

반면, 포아송 영상 편집에서는 영상을 그대로 붙이는 대신에, 영상을 보간하여 배경과 전경을 자연스럽게 합성한다. 배경영상의 한 화소에 4를 곱한 값에서 이웃한 네 화소 값의 합을 빼는데, 이것은 각 화소에

서 선택된 화소와의 변화도(Gradient) 값을 빼는 것이다. 전경영상 역시 각 화소에서 선택된 화소와의 변화도 값을 빼다. 배경영상의 각 화소에서 선택된 화소와의 변화도 값을 빼 값과 전경영상의 변화도가 같을 때 해의 근사값을 찾아서 배경과 전경이 자연스럽게 합성되도록 한다. 입력은 2차원 영상이고 이산적인 성질을 가지고 있기 때문에 이산 포아송 방정식을 이용한 다. 식 (10)은 이산 포아송 방정식이다.

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p \cap \Omega} f_q^* + \sum_{q \in N_p} V_{pq} \quad (10)$$

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p} V_{pq} \quad (11)$$

식 (10)과 식 (11)은 Ω 안에 있는 모든 화소 p에 관한 것으로, 식 (10)은 전경과 배경의 화소가 경계선에 위치하고 있을 때 사용하며, 식 (11)은 전경과 배경의 경계선에 걸쳐 있지 않을 때 적용한다. 식 (10)과 식 (11)에서 Ω 는 추출된 전경영상의 범위, p 는 Ω 안에 위치한 화소 중의 하나, 그리고 $|N_p|$ 는 경계선 안의 이웃한 화소의 개수이다. f_p 는 Ω 안에 위치한 화소 중의 하나이며, q는 p에 이웃한 화소이다. f_q 는 Ω 안에 위치한 화소 p의 이웃한 화소이고, f_q^* 는 경계선에 위치한 화소 p의 이웃한 화소 중의 하나이다.

$$V_{pq} = g_p - g_q \quad (12)$$

여기서 V_{pq} 는 전경의 변화도 필드(Gradient field)이다. g_p 는 Ω 안에 위치한 화소이고, g_q 는 Ω 안에 위치한 화소 p의 이웃한 화소 중의 하나이다.

그림 4는 식 (12)에 대한 기호를 보여준다. 식 (10)과 식(11), 그리고 식 (12)를 해결하기 위해 본 논문에서는 자코비 반복법(Jacobi method)을 활용해서 해의 근사값을 구한다. 자코비 반복법은 $Ax=b$ 형태의 선형 연립 방정식을 구하는 방법 중의 하나이다. $Ax=b$ 의 형태를 $x_n = Cx_{n-1} + d$ 의 형태로 바꾼 뒤, 임의의 초



그림 4. 전경과 배경의 합성 기호
Fig. 4. Notations for Poisson image equations

기 값 x_0 를 선정하여 다음의 계산을 수렴할 때까지 또는 지정한 횟수만큼 반복한다.

일반적인 포아송 영상 편집은 전경의 변화도가 작은 부분, 즉 사용자가 합성을 원하지 않은 부분에 대해서 배경에 손실을 준다. 따라서 포아송 영상 편집을 수행하기 전에 그래프를 적용하여 전경을 효과적으로 잘라서 영상을 합성하는 방법도 제안되었다^{5,6)}.

3.4 단일 색상영상과 단일 깊이영상의 합성

그래프적으로 추출된 전경영상에 임의의 깊이 값 (0~255)을 넣고 단일 색상영상과 단일 깊이영상에 합성한다. 단, 배경영상은 깊이영상이 있다는 가정 하에 합성한다.

식 (13)과 식 (14)에서 Dep_{fore} 는 전경의 깊이 값이고, Dep_{back} 는 배경의 깊이 값을 나타낸다. Col_{fore} 는 전경영상의 색상 값, Col_{back} 는 배경영상의 색상 값이다. 식 (13)처럼 배경과 전경영상의 깊이 값을 비교하여, 배경의 깊이 값이 크면 배경영상의 깊이 값을 유지하고, 작으면 전경의 깊이 값으로 값을 바꾸어서 깊이영상을 만든다. 색상영상에는 식(14)를 적용하여 배경영상과 전경영상의 깊이 값이 달라짐에 따라서 전경영상이 배경영상에 의해 가려지기도 하고 배경영상이 나타나기도 한다.

$$\begin{cases} Dep_{fore} > Dep_{back} , Dep_{back} = Dep_{fore} \\ Dep_{fore} \leq Dep_{back} , Dep_{back} = Dep_{back} \end{cases} \quad (13)$$

$$\begin{cases} Dep_{fore} > Dep_{back} , Col_{back} = Col_{fore} \\ Dep_{fore} \leq Dep_{back} , Col_{back} = Col_{back} \end{cases} \quad (14)$$

색상영상의 합성은 깊이영상의 합성과 유사한데, 전경과 배경의 깊이 값을 참조하여 만약 전경의 깊이 값이 배경의 깊이 값보다 작다면 배경의 색상 값을 유지하고, 전경의 깊이 값이 배경의 깊이 값보다 크다면 전경의 색상 값으로 합성한다.

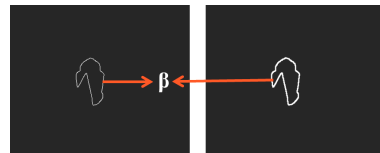


그림 5. 경계선 영역 블러링을 위한 팽창화
Fig. 5. Dilatation for boundary region blurring

3.5 경계선 영역의 블러링

포아송 영상 편집은 합성된 부분에서 경계선에 대한 철저한 처리를 해 주지만, 좀 더 자연스러운 영상을 합성하기 위해서 전경의 외곽과 배경이 만나는 경계선 영역에 블러링을 적용한다. 합성된 영상의 경계선 영역을 얻고 경계선 영역에 모폴로지(Morphology) 영상처리 기법인 팽창(Dilatation)을 적용한다. 포아송 영상 편집을 수행한 영상에서 얻어진 영역의 좌표에 평균값 필터(Mean filter)를 이용한다. 따라서 배경과 전경의 색상톤이 어우러지게 되고 계단화 현상이 제거된다. 그림 5는 경계선 부분 β 를 팽창화를 시킨 영상이다. 이후 이미 합성된 영상에서 β 에 속한 모든 화소 p 에 평균값 필터를 수행한다.

IV. 실험 결과 및 분석

4.1 경계선 영역의 블러링 결과

이 논문에서는 경계선 영역을 추출한 뒤 침식 마스크와 팽창 마스크를 사용하여 모폴로지 영상처리를 적용했다. 그림 6은 경계선 영역 블러링을 활용한 영상과 활용하지 않은 영상을 비교한 것이다. 그림 6(a) 영상에서 알 수 있듯이, 포아송 영상 편집만 실행한 영상은 전경과 배경의 경계 부근에서 계단화 현상이 심하게 발생한다. 하지만 경계선 영역 블러링을 적용한 그림 6(b)은 경계부근 생기는 계단화 현상이 제거되었다.



(a) 포아송 이미지 에디팅만 적용된 영상
(b) 경계선 영역이 평활화가 적용된 영상

그림 6. 경계선 영역 블러링 결과
Fig. 6. Result of Boundary Region Blurring

4.2 깊이 값에 따른 단일 색상영상 합성 결과

그림 7에 보인 것처럼, 깊이 값을 임의로 변경하면 깊이 값에 의해 전경이 가려지기도 하고 나타나기도 한다. 또한, 전경의 깊이 값이 같더라도 위치를 바꾸어 합성하면 배경의 깊이 값이 변해서 전경의 보이는 부분과 가려지는 부분이 달라지는 것을 알 수 있다. 그림 7(a)부터 그림 7(d)은 각각 다른 깊이 값을 이용하여 합성한 영상이다. 색상영상에 깊이영상을 고려하면 전경과 배경의 기하학적 위치까지 반영할 수 있어서, 보다 자연스러운 합성 영상을 얻을 수 있다.



(a) 깊이 값: 120

(b) 깊이 값: 255



(d) 깊이 값: 190

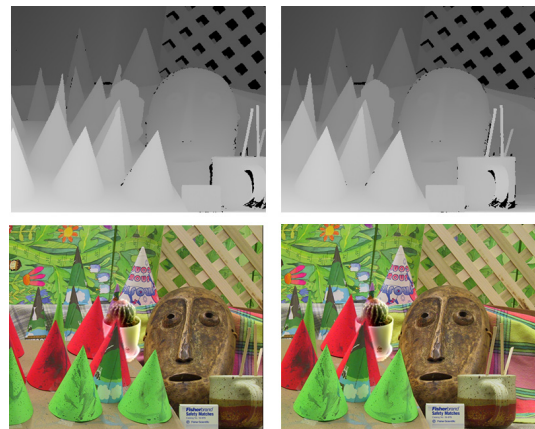


(c) 깊이 값: 130

그림 7. 깊이 값 변화에 따른 합성 결과 비교
Fig. 7. Comparison of Composited Color Images

4.3 스테레오 영상 합성 결과

그림 8은 이 논문에서 제안한 포아송 영상 합성 방법을 이용하여 얻은 스테레오 영상의 합성 결과를 보여준다.



(a)Left image

(b)Right image

그림 8. 스테레오 영상 합성 결과
Fig. 8. Result of composited Stereo Image

V. 결 론

이 논문에서는 고품질의 스테레오 영상을 자연스럽게 합성하기 위한 포아송 영상 편집 기술을 설명했다. 본 논문에서 제안한 방법은 이진 세그멘테이션을 수행하여 객체를 영상에서 추출하고, 임의로 깊이 값을 정해 추출된 객체를 좌영상에서 우영상으로 옮겼다. 스테레오 영상의 깊이 값과 사용자가 임의로 정한 깊이 값을 비교하여 영상 안에서의 전경의 위치를 예측했고, 포아송 영상 편집 방법을 적용하여 전경과 배경을 합성했다. 또한, 경계선 영역에 블러링을 수행하여 전경과 배경 주변에 발생하는 계단화 현상을 감소시켰다. 실험적으로, 제안한 방법은 기존의 스테레오 영상 합성 방법에 비해 좀더 자연스러운 결과를 얻을 수 있었다.

References

- [1] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *Proc. ACM SIGGRAPH*, vol. 22, no. 3, pp. 313-318, Jul. 2003.
- [2] X. Li, "The study on digital image composition by water-wave image," M.S. Thesis, Sungsil Univ., Korea, 2008.
- [3] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut-interactive foreground extraction using iterated graph cuts," in *Proc. ACM SIGGRAPH*, vol. 23, no. 3, pp. 309-314, Aug. 2004.
- [4] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*, NY: McGraw-Hill, 1995.
- [5] E. Baek and Y. Ho, "Image composition for object based on poisson image editing," in *Proc. KSPC Conf.*, vol. 26, pp. 69-72, 2013.
- [6] J. Jia, J. Sun, C. Tang, and H. Shum, "Drag-and-drop pasting," in *Proc. ACM Trans. Graphics (TOG)*, vol. 25, no. 3, pp. 631-637, Jul. 2006.

백 으 뜸 (Eu-Tteum Baek)



2012년 : 전북대학교 컴퓨터 공학과 졸업 (학사)
 2013년~현재 : 광주과학기술원 정보통신공학과 석사과정
 <관심분야> 디지털 영상처리, 3D 영상합성

호 요 성 (Yo-Sung Ho)



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