

논문 2010-47SP-6-4

# 녹화된 아날로그 영상의 화질 개선을 위한 잡음 연관성을 고려한 학습기반 잡음개선 기법

( Training-Based Noise Reduction Method Considering Noise Correlation  
for Visual Quality Improvement of Recorded Analog Video )

김 성 득\*, 임 경 원\*\*

( Sung Deuk Kim and Kyoung Won Lim )

## 요 약

녹화된 아날로그 영상에 내재하는 잡음을 효과적으로 제거하기 위해서는 잡음의 실제 특성과 정도를 정확히 파악하는 것이 매우 중요하다. 본 논문에서는 실제 방송되는 아날로그 영상을 녹화하여 잡음의 특성을 분석한 후, 녹화된 아날로그 영상을 위한 효과적인 학습기반 잡음개선 방법을 제안한다. 먼저 녹화된 아날로그 영상의 잡음을 분석하여 무시할 수 없는 잡음의 연관성이 존재하는 것을 보임으로써, 전통적인 부가 백색 가우시안 잡음(AWGN) 모델에 기반을 둔 잡음의 추정과 잡음 제거 방법이 가지는 한계를 설명한다. 또한 잡음의 연관성을 고려한 자기회귀 모델을 이용해서 녹화된 아날로그 영상에 내재하는 잡음을 추정하고 합성할 수 있음을 보이며, 추정된 자기회귀 모델을 이용해 학습기반 잡음제거 기법에 적용함으로써 비디오 잡음을 제거한다. 실험결과는 제안된 방법이 무시할 수 없을 정도로 잡음 연관성을 가진 실제 녹화된 아날로그 영상의 잡음 제거에 효과적으로 활용될 수 있음을 보여준다.

## Abstract

In order to remove the noise contained in recorded analog video, it is important to recognize the real characteristics and strength of the noise. This paper presents an efficient training-based noise reduction method for recorded analog video after analyzing the noise characteristics of analog video captured in a real broadcasting system. First we show that there is non-negligible noise correlation in recorded analog video and describe the limitations of the traditional noise estimation and reduction methods based on additive white Gaussian noise (AWGN) model. In addition, we show that auto-regressive (AR) model considering noise correlation can be successfully utilized to estimate and synthesize the noise contained in the recorded analog video, and the estimated AR parameters are utilized in the training-based noise reduction scheme to reduce the video noise. Experiment results show that the proposed method can be efficiently applied for noise reduction of recorded analog video with non-negligible noise correlation.

**Keywords :** analog noise reduction, trained filter, auto-regressive model

## I. 서 론

최근 들어 고해상도를 지원하는 고화질 디스플레이 장치 가 각 가정에까지 널리 보급되었다. 이로 인해 시

청자는 매우 선명한 화질을 가정에서도 느낄 수 있게 되었고, 영상 화질에 대해 매우 높은 기대수준을 가지게 되었다. 가정용 텔레비전은 캠코더, 케이블 방송, DVD, 디지털 방송 등과 같은 다양한 종류의 입력을 처리해야 하므로, 시청자의 높아진 화질 기대치를 만족하기 위해서는 각종 아날로그 잡음과 디지털 잡음을 적절히 제거하는 방안이 필수적이다. 특히 지상파 HDTV 방송의 선명한 화질에 익숙해진 시청자는 과거 녹화된 아날로그 영상이 재생될 때 나타나는 잡음으로 인한 화

\* 정회원, 안동대학교 정보전자공학교육과  
(Dept. of IT & Electronics Education,  
Andong National University)

\*\* 정회원, (주) LG전자 DTV연구소  
(Digital TV Lab., LG Electronics, Inc.)

접수일자: 2010년6월26일, 수정완료일: 2010년10월4일

질 저하를 더 쉽게 느끼게 된다. 아날로그 잡음 제거 방안들은 오래 동안 연구되어 왔지만, 잡음 제거 기술에 있어서 간과되기 쉬운 부분 중 하나는 입력되는 잡음의 정도를 추정하는 것이다. 잡음 추정은 영상에 내재하는 잡음의 정도를 판단하여 잡음 제거를 위한 필터링의 강도를 제어하는 역할을 수행한다. 따라서 녹화된 아날로그 영상에 내재하는 잡음을 효과적으로 제거하기 위해서는 잡음의 실제 특성과 정도를 파악하는 것이 매우 중요하다.

전통적인 아날로그 잡음 제거 방안에서는 잡음을 AWGN로 가정하고 잡음의 표준편차를 추정한 후, 잡음의 정도를 고려하여 인접한 화소들을 활용한 공간영역 필터링 혹은 시간영역 필터링을 수행한다<sup>[1~9]</sup>. 잡음 추정 방법으로는 영상 내에서 매우 평탄한 영역을 선별하고, 해당 국부 영역의 분산으로 잡음 분산을 추정하는 방안이 널리 사용된다. 국부 분산을 활용하는 방법으로는 국부 분산의 히스토그램(histogram)을 구한 후 적절한 임계치를 활용해서 잡음의 분산 (혹은 표준편차)을 추정하는 방안<sup>[1~3]</sup>과 영상을 블록으로 나눈 후 선별된 평탄한 블록의 분산의 평균값 혹은 최소값에 기반을 둔 예측값을 사용하기도 한다<sup>[4]</sup>. 블록 기반의 잡음 분산 예측에서는 적절히 평탄한 블록을 선택하는 방법과 선택된 블록의 분산으로부터 최종적인 잡음 분산을 어떻게 예측하는가가 중요하므로, 적응적인 Gaussian 필터링을 이용해 잡음 성분을 구분하기도 하고 3차원 구조를 활용하여 시공간 정보를 함께 사용하기도 한다<sup>[5~6]</sup>.

잡음의 추정이 부정확하여 실제 잡음 분산보다 낮게 추정되면, 원하는 수준만큼 잡음 제거 성능을 보일 수 없다. 반대로 지나치게 높게 잡음 분산이 예측되면 지나친 필터링으로 인해 원치 않는 영상 번짐 현상을 초래할 수 있기 때문에 잡음 추정의 정확도를 높이는 것은 실제 응용에서 매우 중요한 요소가 될 수 있다.

국부 분산의 히스토그램을 이용하는 잡음 추정에 있어서 주목할 만한 어려움 중 하나는 히스토그램의 어떤 부분이 잡음으로 인해 발생했다고 판단하기 어렵다는 점이다. 평탄한 영역에서의 잡음 성분이 국부 분산의 히스토그램에 지배적인 영향을 주고, 국부 분산의 분포는 Gaussian 분포를 따를 것이라는 가정이 잘 성립해야만 정확한 추정이 가능하다. 하지만, 일반적으로 잡음의 영향만 존재하는 일부 영역을 자동으로 구분하는 것은 어려운 일이며, 영상에 내재하는 텍스처 혹은 경계 등의 영향이 필연적으로 개입된다. 따라서 영상에 내재하

는 텍스처 혹은 경계 등의 영향을 줄이기 위해서 영상 자체의 국부 분산을 직접 이용하는 대신 시공간적 예측 오차 정보에 해당하는 국부 분산을 활용하여 보다 정밀한 예측을 수행하기도 한다<sup>[7~8]</sup>. 또한, 국부 분산의 히스토그램의 분포가 정확히 Gaussian이 되지 않는 경우를 대비하여 최적화 도구를 활용하여 히스토그램을 정합한 후 최적의 파라미터를 추정하는 방안도 시도되었다<sup>[8~9]</sup>.

다양한 시도들이 AWGN 모델에 바탕을 두고 제안되었으나, 실제 방송되는 아날로그 방송을 녹화하여 과연 실제 잡음이 AWGN 모델을 만족하는지와 실제 환경에서 충분히 신뢰할 수 있는 예측이 이루어지는지에 대해 검증하는 것은 실제 응용에서 매우 중요하다.

본 논문에서는 실제 방송되는 아날로그 방송 영상을 녹화하여 잡음의 특성을 분석하는 과정을 수행함으로써 실제 아날로그 잡음이 순수한 AWGN 특성을 가지지 않음을 보인다. 특히 잡음의 연관성으로 인해 야기되는 기존의 잡음 추정 방법이 가지는 문제점을 보인다. 이러한 문제를 극복하면서 잡음의 연관성을 고려하기 위하여, 본 논문에서는 자기회귀 (auto-regressive: AR) 모델을 아날로그 잡음 모델로 설정한다<sup>[10~11]</sup>. AR 모델 파라미터를 예측하기 위해 먼저 사용자의 도움을 받아 시간방향으로 정적인 영역을 정한 후, ground truth 영상을 추정한다. 추정된 ground truth 영상과 잡음 영상으로부터 AR 모델을 예측한 후 이미 알고 있는 잡음이 없는 영상에 AR 잡음을 추가하여 합성된 잡음 영상을 만든다. 합성된 AR 잡음이 있는 영상과 원 영상으로부터 학습기반 최적 자승 필터링 방법에 기반을 둔 잡음 제거 방법을 적용한다.

본 논문은 다음과 같이 구성된다. II장에서는 전통적인 AWGN 기반 방식에 대비하여 제안된 학습기반 방식의 전체적인 구조를 설명한다. 또한 실제 녹화된 아날로그 방송 영상으로부터 잡음의 특성을 분석하는 방안을 설명하고, 기존의 잡음 분산 추정 방법이 가지는 문제점을 분석한다. III장에서는 AR 잡음을 추정하고 합성하는 방식과 함께, AR 잡음 모델을 활용하여 학습기반 잡음 제거를 수행하는 방안을 설명한다. IV장에서는 모의 실험결과를 보이고, V장에서 결론을 맺는다.

## II. 제안한 방식의 구조 및 잡음 특성 분석

### 1. 문제 제기 및 제안한 방식의 전체적 구조

그림 1은 AWGN 모델에 기반을 둔 전통적인 잡음 제거 방법의 기본 구조를 보여준다. 잡음이 없는 영상  $f$ 에 AWGN 모델의 잡음  $n$ 이 섞여 잡음이 있는 관측된 영상  $g$ 가 얻어졌다고 가정한다. 잡음 추정부에서는 영상  $g$ 로부터 잡음의 분산  $\hat{\sigma}$ 를 예측하고, 예측된  $\hat{\sigma}$ 를 바탕으로 필터링의 강도를 적절히 조절하여 잡음 제거를 수행한다. 그림 1과 같은 전통적인 구조를 이용해서 녹화된 아날로그 영상 잡음을 효과적으로 제거하려면 다음 두 가지 중요한 조건이 성립해야 한다. 잡음 모델이 AWGN 모델에 근접해야 하고, 잡음 추정 방법이 비교적 근접한 잡음 예측치를 제공해야 한다. 다시 말해 잡음의 특성이 가정한 모델과 유사하지 않거나, 잡음 분산의 예측이 적절하지 않다면 필터링 방식을 적절히 제어할 수 없기 때문에 충분한 화질 개선 효과를 기대할 수 없다.

AWGN의 잡음 모델을 바탕으로 한 기존의 잡음 추정 방법들은 AWGN 환경에서는 비교적 좋은 성능을 보여준다. 그림 2는 잡음 분산이 16인 AWGN을 삽입한

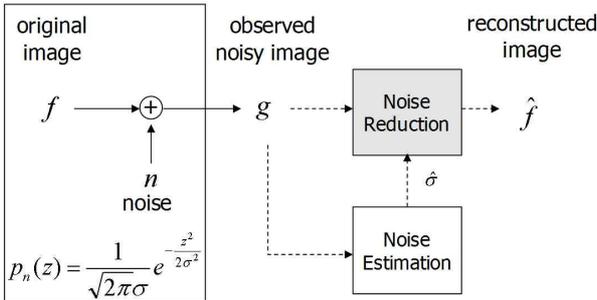


그림 1. AWGN 모델에 기반을 둔 전통적인 잡음 제거 방법의 기본 구조

Fig. 1. Basic structure for conventional noise reduction schemes based on AWGN model.

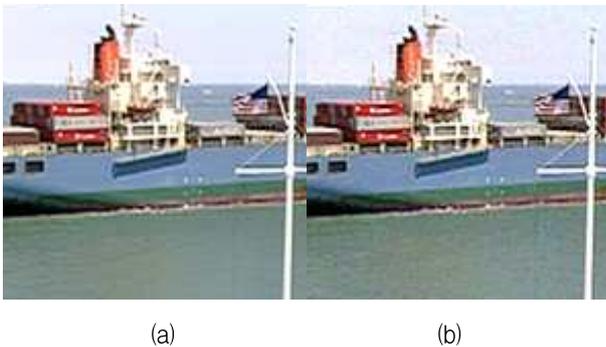


그림 2. (a) 잡음이 없는 영상 (b) AWGN (잡음분산=16)를 삽입한 영상

Fig. 2. (a) Original image without noise and (b) noisy image with AWGN (noise variance=16).

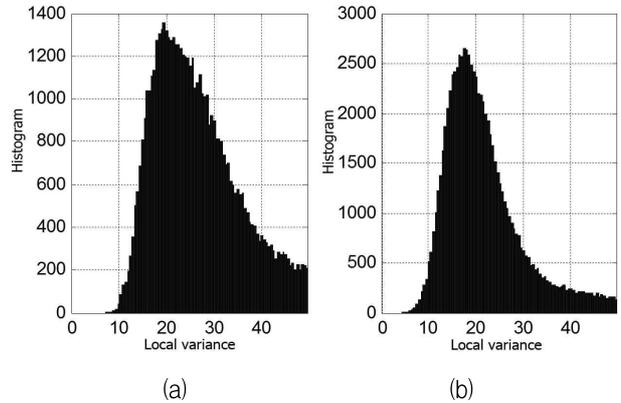


그림 3. AWGN 잡음 영상에 대한 히스토그램 기반의 잡음 예측 (a)  $g$ 의 국부 분산의 히스토그램 (b) Rank의 방법<sup>[7]</sup>에 의해 얻어진  $g$ 의 공간 영역 예측 오차의 국부 분산의 히스토그램

Fig. 3. Histogram-based noise estimation for noisy images with AWGN (a) histogram of local variance of  $g$  and (b) histogram of local variance of the spatial prediction error of  $g$  obtained by Rank's method<sup>[7]</sup>.

영상을 보여주며, 그림 3은 히스토그램 방법으로 잡음의 분산을 예측한 예를 보여준다. 그림 3(a)는 잡음이 있는 영상  $g$ 의 국부 분산의 히스토그램을 보여준 것으로 낮은 분산 영역에서 잡음의 분포와 유사한 특성을 확인할 수 있다. 그림 3(b)는 Rank의 방법<sup>[7]</sup>에 의해 얻어진  $g$ 의 공간 영역 예측 오차의 국부 분산의 히스토그램으로 그림 3(a)에 비해 잡음의 특성을 좀 더 잘 반영하는 것을 알 수 있다. 그림 3(b)에서 최대값을 가지는 지점이 잡음 분산 16에 비교적 근접하다는 점은 주목할 만하다.

실제 녹화된 아날로그 잡음 영상에 대해서도 국부 분산의 히스토그램을 활용하는 기존 접근 방법을 시도해 볼 수 있다. 그림 4는 녹화된 아날로그 잡음 영상의 일부이다. 잡음 추정의 정확성을 높이기 위해 그림 4(b)처럼 일부 영역만을 취하여 국부 분산의 히스토그램을 구한 것이 그림 5이다.

AWGN 환경에서 비교적 정확한 예측 값을 얻어낸 그림 3과는 달리 실제 녹화된 아날로그 잡음 영상에 대해서 그림 5는 특이한 결과를 보여준다. 그림 2(b)와 그림 4(b)는 잡음의 차이가 크지 않게 보이지만, 그림 5(a)는 그림 3(a)와는 다르게 국부 분산의 최소값이 상대적으로 매우 작으며, 공간 영역 오차의 국부 분산 히스토그램인 그림 5(b)는 그림 3(b)에 비해 극히 낮은 값에서 최대 피크를 보인다. 즉, 그림 5를 바탕으로 잡음



그림 4. (a) 녹화된 아날로그 잡음 영상  
(b) 시간적으로 정적인 일부 영역  
Fig. 4. (a) Recorded analog noisy video and  
(b) its sub-region without temporal movement.

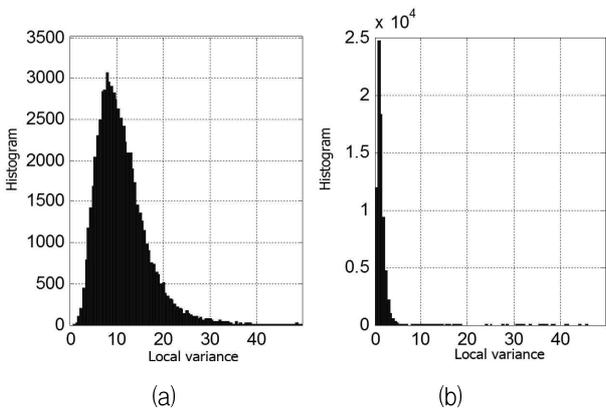


그림 5. 녹화된 아날로그 잡음 영상에 대한 히스토그램 기반의 잡음 예측 (a)  $g$ 의 국부 분산의 히스토그램 (b) Rank의 방법<sup>[7]</sup>에 의해 얻어진  $g$ 의 공간 영역 예측 오차의 국부 분산의 히스토그램  
Fig. 5. Histogram-based noise estimation for recorded analog noisy images (a) histogram of local variance of  $g$  and (b) histogram of local variance of the spatial prediction error of  $g$  obtained by Rank's method<sup>[7]</sup>.

의 정도를 추정하면 그림 4는 낮은 잡음 분산 추정 값을 얻게 된다. 잡음의 분산이 낮게 예측되면 필터링의 강도가 약해져 잡음이 제대로 제거되지 않은 상태로 필터링 결과가 얻어질 수 있다.

본 논문에서는 이러한 문제점을 극복하기 위해 영상의 잡음 특성 분석을 통해 얻어진 AR 모델을 활용하는 그림 6과 같은 학습기반 잡음제거 방식을 제안한다. 보

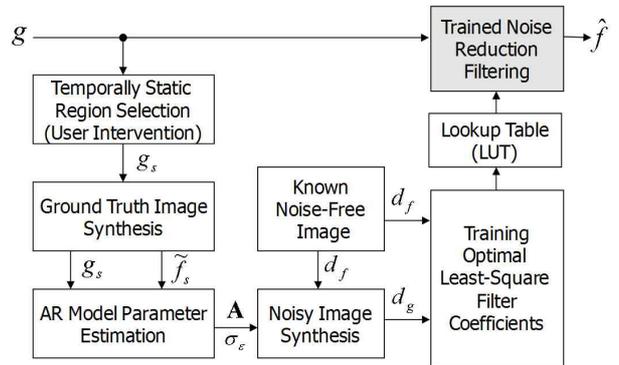


그림 6. 녹화된 아날로그 영상을 위한 제안된 학습기반 잡음제거 방법의 블록도  
Fig. 6. Block diagram of the proposed training-based noise reduction method for recorded analog video.

다 일반화된 잡음 모델인 AR 모델을 사용함으로써 잡음의 특성에 보다 강인하게 대처할 수 있으며, 학습을 통해 최적 자승 해의 필터 계수를 유도할 수 있다는 장점이 있다. 제안된 구조는 사용자에게 의해 시간방향으로 정적인 영역을 선택하는 단계, 선택된 영역에서 ground truth 영상을 얻는 단계, ground truth 영상과 관측된 영상으로부터 AR 모델 파라미터를 예측하는 단계, 예측된 AR 모델과 잡음이 없는 알려진 영상들을 활용하여 인위적으로 AR 잡음을 추가한 영상을 얻는 단계, 잡음이 없는 알려진 영상과 인위적으로 잡음을 합성한 영상을 바탕으로 최적 자승 필터 계수를 학습하여 look-up table (LUT)에 저장하는 단계, 학습되어 LUT에 저장된 필터 계수를 활용하여 실제 잡음 제거 필터링을 수행하는 단계로 구분된다.

## 2. Ground Truth 영상을 활용한 잡음 특성 분석

본 논문에서는 녹화된 아날로그 잡음의 특성을 파악하기 위하여 ground truth 영상 개념을 사용한다. 잡음의 특성을 제대로 분석하려면 잡음이 없는 원본 영상이 필요하다. 하지만, 녹화된 방송 영상의 경우 원본 영상을 확보하기 어려우므로, 원본 영상을 대신하여 ground truth 영상을 추정하여 사용했다. 식 (1)은 ground truth 영상  $\tilde{f}_s$ 를 얻는 과정을 보여준다.

$$\tilde{f}_s(t) = \frac{1}{N} \cdot \sum_{T=0}^{N-1} g_s(t+T) \quad (1)$$

여기서  $g_s$ 는 잡음이 있는 영상  $g$ 에서 시간 방향으로 움직임이 없는 영역을 취한 것이다. 정밀한 예측을 위



그림 7. 그림 4(b)에 대응되는 ground truth 영상  
Fig. 7. Ground truth image corresponding to Fig. 4(b).

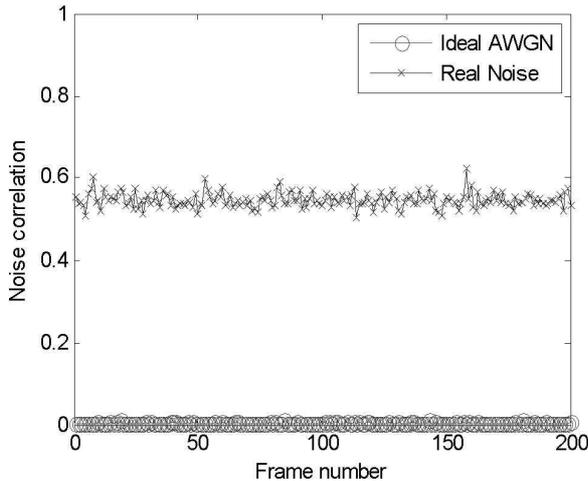


그림 8. 녹화된 아날로그 영상에서 볼 수 있는 인접한 잡음과의 correlation  
Fig. 8. Correlation with neighboring noise observed in recorded analog video.

해  $g$ 에서  $g_s$ 를 선택하는 것은 사용자에게 의해 선택된다. 식 (1)은  $\tilde{f}_s$ 가  $g_s$ 의 시간 방향 평균값으로 예측됨을 의미한다.  $N$ 은 사용된  $g_s$ 의 프레임수를 의미한다. 그림 7은 그림 4(b)에 대응되는 ground truth 영상의 예를 보인 것으로 200프레임의 영상을 사용하여 식 (1)에 의해 얻어진 것이다. 그림 7은 그림 4(b)와 비교할 때 잡음 성분이 대부분 제거된 원 영상에 준하는 ground truth 영상으로 사용하기 적합하다.

원 영상에 준하는 ground truth 영상을 얻은 후에는 잡음 영상에서 ground truth 영상을 빼 순수하게 잡음만 남은 영상  $n_s$ 을 식 (2)처럼 얻을 수 있다.

$$n_s = g_s - \tilde{f}_s \quad (2)$$

그림 8은 녹화된 아날로그 잡음  $n_s$ 의 공간적 연관성을 correlation관점에서 계산한 것이다. 이상적인 AWGN에서 잡음의 correlation은 0이지만, 녹화된 영상의 잡음에는 무시할 수 없을 정도의 correlation이 존재하는 것을 알 수 있다. 이것은 녹화된 아날로그 비디오의 잡음을 효과적으로 제거하기 위해 AWGN와는 다르게 잡음의 연관성을 고려하는 방안이 필요함을 보여준다.

무시할 수 없을 정도의 잡음의 correlation은 그림 5(b)에서 잡음의 분산이 기대치에 비해 낮게 나오는 것을 설명해 준다. 공간적 예측 오차를 활용하여 잡음의 정도를 추정하는 방법은 식 (3)과 같은 분산과 관련된 정의를 활용한다.

$$\text{var}(X - Y) = \text{var}(X) + \text{var}(Y) - 2\text{cov}(X, Y) \quad (3)$$

식 (3)에서  $\text{var}(\ )$ 와  $\text{cov}(\ )$ 는 각각 분산과 공분산을 나타내는 함수이며,  $X$ 와  $Y$ 는 각각 잡음이 있는 현재 픽셀과 인접한 픽셀을 나타내는 확률변수이다. 잡음 예측을 위해 평탄한 영역을 선택해 현재 픽셀과 인접한 픽셀 간의 예측 오차가 매우 미미하다고 가정하면, 예측 오차의 분산뿐만 아니라 예측 오차와 잡음과의 공분산 값들이 무시할 만한 값이 되므로 근사적으로 식 (3)은 식 (4)와 같아진다.

$$\text{var}(X - Y) \approx 2 \cdot (\text{var}(n) - \text{cov}(n_X, n_Y)) \quad (4)$$

식 (4)에서  $\text{var}(n)$ 는 잡음의 분산을 의미하며,  $n_X$ 와  $n_Y$ 는 각각  $X$ 와  $Y$ 위치에서 잡음 성분을 의미하는 확률변수이다. 이상적인 AWGN 환경에서는 잡음간의 연관성이 없으므로  $\text{cov}(n_X, n_Y) = 0$ 이 되어  $\text{var}(X - Y) \approx 2 \cdot \text{var}(n)$ 로 추정할 수 있으므로 정확한 예측이 가능했으나, 실제 녹화된 아날로그 영상에서는  $\text{cov}(n_X, n_Y)$ 값이 무시될 수 없으므로 식 (4)에 의해  $\text{cov}(n_X, n_Y) = 0$ 인 경우보다 더 작은 값이 얻어진다. 따라서 잡음의 연관성이 무시될 수 없는 실제 환경에서는 공간 영역의 예측 오차를 활용하는 방식이 잡음 분산의 추정에 관한 정확성을 향상시키는데 한계를 가진다. 본 논문에서는 이러한 한계를 극복하고, 실제 녹화된 아날로그 영상에 내재하는 잡음 특성을 충실히 고려하기 위해 AR 모델을 바탕으로 잡음의 특성을 추정하고, 잡음이 없는 알려진 영상에 AR 잡음을 추가한 후 학습기반의 필터링 방법을 적용하여 비디오 잡음을 제거하는 방안을 사용한다.

### III. AR 모델에 바탕을 둔 학습기반 잡음제거

#### 1. AR 모델을 이용한 잡음 모델링과 합성

잡음의 연관성을 고려하기 위해 AR 모델을 이용해 잡음의 특성을 모델링하고 합성하여 학습기반 필터링에

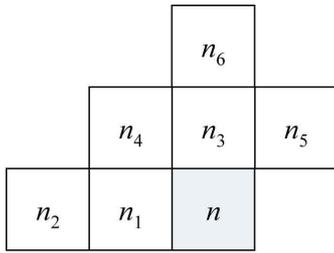


그림 9. 사용된 AR 모델에서의 인접 픽셀 위치들  
Fig. 9. Neighboring pixel positions used for the AR model.

활용한다. 그림 9는 사용된 AR 모델에 활용되는 인접 픽셀들의 위치를 보여준다.

식 (5)는 AR 잡음 모델을 식으로 나타낸 것으로,  $P$  개의 AR 파라미터  $a_i$ 와 백색잡음  $\varepsilon$ 로 잡음을 정의한다. 본 논문에서는  $P=6$ 을 사용한다.  $n^k$ 와  $n_i^k$ 는  $k$ 번째 샘플에서의 잡음 성분과 해당되는 인접한 잡음 성분을 각각 의미한다.

$$n^k = \sum_{i=1}^P a_i \cdot n_i^k + \varepsilon \quad (5)$$

식 (6)~(8)은 AR 파라미터의 추정을 설명하기 위해 백색잡음을 제외하고 식 (5)을 행렬 형태로 표현한 것이다.

$$n^k = \mathbf{X}^k \mathbf{A} \quad (6)$$

$$\mathbf{X}^k = [n_1^k \ n_2^k \ n_3^k \ n_4^k \ n_5^k \ n_6^k], \quad (7)$$

$$\mathbf{A} = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6]^T \quad (8)$$

AR 모델 파라미터는 식 (9)~(10)과 같은 covariance method을 통해 식 (11)~(12)와 같이 추정할 수 있다<sup>[10]</sup>. 여기서  $S$ 는 샘플의 수이다.

$$\bar{\mathbf{N}} = \bar{\mathbf{X}} \mathbf{A} \quad (9)$$

여기서

$$\bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X}^1 \\ \mathbf{X}^2 \\ \vdots \\ \mathbf{X}^k \\ \vdots \\ \mathbf{X}^S \end{bmatrix}, \quad \bar{\mathbf{N}} = \begin{bmatrix} n^1 \\ n^2 \\ \vdots \\ n^k \\ \vdots \\ n^S \end{bmatrix} \quad (10)$$

$$\mathbf{A} = (\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} \bar{\mathbf{X}}^T \bar{\mathbf{N}} \quad (11)$$



(a)



(b)

그림 10. 확대된 잡음 영상 (a) 원본 잡음 영상  
(b) AR 모델로 합성된 잡음 영상

Fig. 10. Magnified noisy image (a) original noisy image and (b) the noisy image synthesized by the AR model.

$$\sigma_\varepsilon^2 = \frac{(\bar{\mathbf{N}} - \bar{\mathbf{X}} \mathbf{A})^T (\bar{\mathbf{N}} - \bar{\mathbf{X}} \mathbf{A})}{S} \quad (12)$$

본 논문에서는 예측된 AR 모델을 바탕으로 학습기반 필터링 구조를 취하므로, 잡음이 없는 알려진 학습 영상  $d_f$ 와 함께 AR 잡음이 추가된 영상  $d_g$ 가 필요하다.  $d_g$ 는 AR 모델에 의해 식 (5)로부터 얻어지는 AR 잡음성분  $n_{ar}$ 과  $d_f$ 의 합으로 식 (13)과 같이 얻어진다.

$$d_g = d_f + n_{ar} \quad (13)$$

실제 잡음과 합성된 잡음의 유사성을 보기 위해, 그림 10(a)는 원본 잡음 영상  $g_s$ 를, 그림 10(b)는 ground truth 영상에 합성된 AR 잡음을 추가한 영상  $\tilde{f}_s + n_{ar}$ 를 보여준다. 원본 잡음 영상과 비교할 때 비슷한 특성을 가지는 잡음 영상이 합성될 수 있음을 알 수 있다.

## 2. 학습기반 잡음 제거 방법의 적용

잡음이 없는 원 영상  $d_f$ 와 합성을 통해 생성된 잡음 영상  $d_g$ 를 활용하여 학습 기반 잡음 제거 필터링을 적용한다. 그림 11은 학습기반 필터링을 위한 학습과정과 잡음 제거 과정을 보여준다. 학습과정에서는 예측된 AR 모델 파라미터를 이용해서 잡음이 없는 알려진 영상  $d_f$ 로부터 잡음이 있는 영상  $d_g$ 를 합성한다. 또한 잡음이 있는 영상의 각 화소는 분류화기를 통해 각각의 분류 코드별로 구분된다. 분류화 방법으로는 식 (14)~

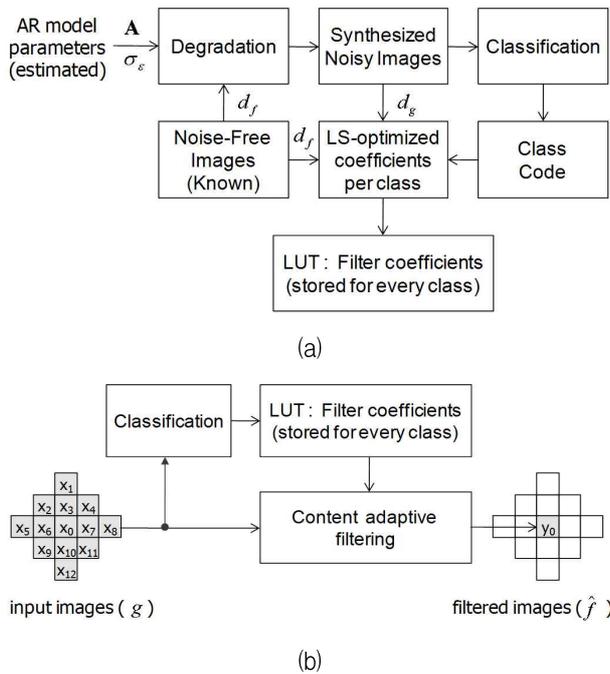


그림 11. 학습기반 잡음 제거 과정

(a) 학습과정 (b) 필터링 과정

Fig. 11. Training-based noise reduction procedure.

(a) training step and (b) filtering step.

(16)과 같은 adaptive dynamic range coding (ADRC)<sup>[12]</sup>를 기반으로 하는 13비트의 분류코드 (class code)를 사용한다.

$$c_i = \begin{cases} b_{12} \oplus b_i, & 0 \leq i \leq 11 \\ c_{12}, & i = 12 \end{cases} \quad (14)$$

여기서

$$b_i = \begin{cases} 1, & x_i \geq (x_{\max} + x_{\min})/2 \\ 0, & \text{otherwise} \end{cases}, \quad (15)$$

$$c_{12} = \begin{cases} 1, & x_{\max} - x_{\min} \geq T_{DR} \\ 0, & \text{otherwise} \end{cases}. \quad (16)$$

식 (14)에서  $c_i$ 는 13비트로 구성된 분류코드의  $i$ 번째 비트를 의미하며,  $\oplus$ 는 exclusive-OR (XOR) 논리 연산을 의미한다. 식 (15)에서  $x_i$ 는 그림 11(b)에서 보이는 것과 같이 필터링의 대상이 되는 픽셀 및 인접 픽셀을 의미한다.  $x_{\max}$ 와  $x_{\min}$ 은 인접 픽셀들의 최대값과 최소값을 각각 의미한다. 식 (16)에서  $T_{DR}$ 은 실험적으로 정해지는 상수값이다.

각 화소는 식 (14)에 의해 분류된 후, 각 분류코드별로 최적 자승기반의 필터계수를 유도한 후 look-up

table (LUT)에 저장하는 것으로 학습과정을 마친다. 최적 자승기반의 필터계수를 유도하는 것은 Hu의 방식과 동일하다<sup>[13]</sup>. 필터링 과정에서는 입력 영상의 각 화소를 학습과정에서 사용된 동일한 분류화기를 통해 분류한 후, 해당되는 분류코드에 대응되는 필터 계수를 LUT에서 읽어서 적응적 필터링을 수행한다.

#### IV. 모의 실험

인위적으로 잡음이 추가된 영상과 실제 녹화된 영상을 바탕으로 제안된 방법을 평가하였으며, *container ship, mother and daughter, foreman, coast guard, silent voice*로 구성된 총 1500프레임의 CIF영상을 학습 영상으로 사용했다. 실제 녹화된 영상은 영상녹화장비를 통해 아날로그 케이블 방송을 녹화한 것으로, 실제 녹화된 그림 4(a)의 일부 영역에서 얻은  $150 \times 100$ 크기의 그림 12(a)를 바탕으로 ground truth 영상인 그림 12(b)를 식 (1)의 방법으로 구한 후 잡음 모델 파라미터를 얻었다. Ground truth 영상을 얻기 위해 사용된 프레임 수는  $N=200$ 이며, 실제 구해진 잡음의 분산은 약 10.8이다. 식 (9)~(11)을 통해 구해진 AR 모델 파라미터는  $a_1 = 0.7393$ ,  $a_2 = -0.2760$ ,  $a_3 = 0.4164$ ,  $a_4 = -0.2594$ ,  $a_5 = 0.0343$ ,  $a_6 = -0.0387$ ,  $\sigma_\epsilon^2 = 4.9140$ 이다. 분류화 과정에서 사용되는  $T_{DR}$ 값으로는 30을 사용했다.

AR 모델에 기반을 둔 학습기반 잡음 제거방법의 성능을 비교하기 위해 적응적 Wiener필터링<sup>[14]</sup>과 함께 동일한 잡음 분산을 가지는 AWGN 모델 기반 학습기반 필터링 방법과 비교하였다. 여기서 AWGN 모델 기반 학습기반 필터링 방법이란 식 (13)과는 다르게 원 영상에 AWGN성분  $n_{awgn}$ 을 추가하여 식 (17)과 같이 잡음

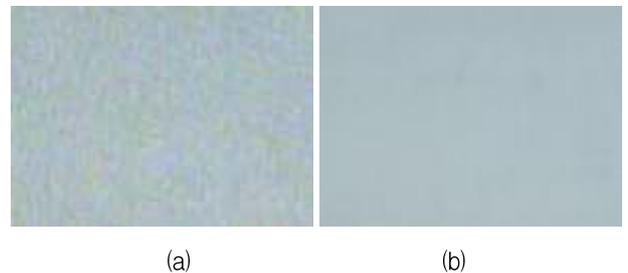
그림 12. 잡음 추정을 위해 선택된 일부 영역 (a) 잡음이 있는 영상  $g_s$  (b) ground truth 영상  $\tilde{f}_s$ .Fig. 12. Selected region for noise estimation (a) noisy image  $g_s$  and (b) its ground truth image  $\tilde{f}_s$ .

표 1. 잡음 분산 10.8을 가지는 AWGN이 삽입된 학습 영상에 필터링을 수행한 결과

Table 1. Filtering results for training sequences corrupted with AWGN of noise variance 10.8.

영상	평균 PSNR_Y [dB]
Noisy	37.8
Adaptive Wiener	39.2
Trained Filter (AWGN)	40.3

표 2. AR 잡음으로 합성된 학습 영상에 대한 필터링 결과

Table 2. Filtering results for training sequences corrupted with the AR noise.

영상	평균 PSNR_Y [dB]
Noisy	37.8
Adaptive Wiener	38.8
Trained Filter (AWGN)	38.7
Trained Filter (AR)	39.2

이 있는 영상  $d_g^*$ 를 만들어 학습 기반 필터링을 수행하는 방법을 의미한다. 이 때 사용되는 AWGN의 분산은 AR 모델을 사용한 경우 얻어지는 잡음 분산과 동일하도록 설정한다.

$$d_g^* = d_f + n_{awgn} \quad (17)$$

그림 13은 추정된 AR 모델을 기반으로 합성된 잡음 영상  $d_g$ 와 이와 동일한 잡음 분산을 가지도록 AWGN를 추가한 잡음 영상  $d_g^*$ 를 보여준다. 그림 13(a)는 AR 모델 기반의 제안된 학습기반 필터링을 위해 사용되는 영상인 반면, 그림 13(b)는 잡음 분산의 이상적인 추정을 가정하고 AWGN 모델을 기반으로 학습을 수행하기 위한 영상이다. 동일한 잡음 분산에 대해서 시각적으로 느껴지는 잡음의 정도가 다소 차이가 있다.

표 1은 잡음의 연관성이 없는 순수한 AWGN이 삽입된 학습 영상  $d_g^*$ 에 대해 적응적 Wiener 필터<sup>[14]</sup>와 학습 기반 필터링 방법의 필터링 성능을 PSNR로 비교한 것이다. 공정한 비교를 위해 적응적 Wiener 필터링에서 잡음 분산은 정확히 안다고 가정하였다. 이상적인 AWGN환경에서는 학습기반 필터링 방법이 PSNR측면에서 우수한 성능을 보이는 것을 확인할 수 있다. 이런 경향이 잡음 연관성이 있는 환경에서 어떻게 변하는지 살펴본 것이 표 2이다. 표 2는 AR 모델 파라미터를 기반으로 합성된 잡음 영상  $d_g$ 에 관한 필터링 결과를 보여준다. 표 2에서 잡음의 분산이 동일한 AWGN 모델을

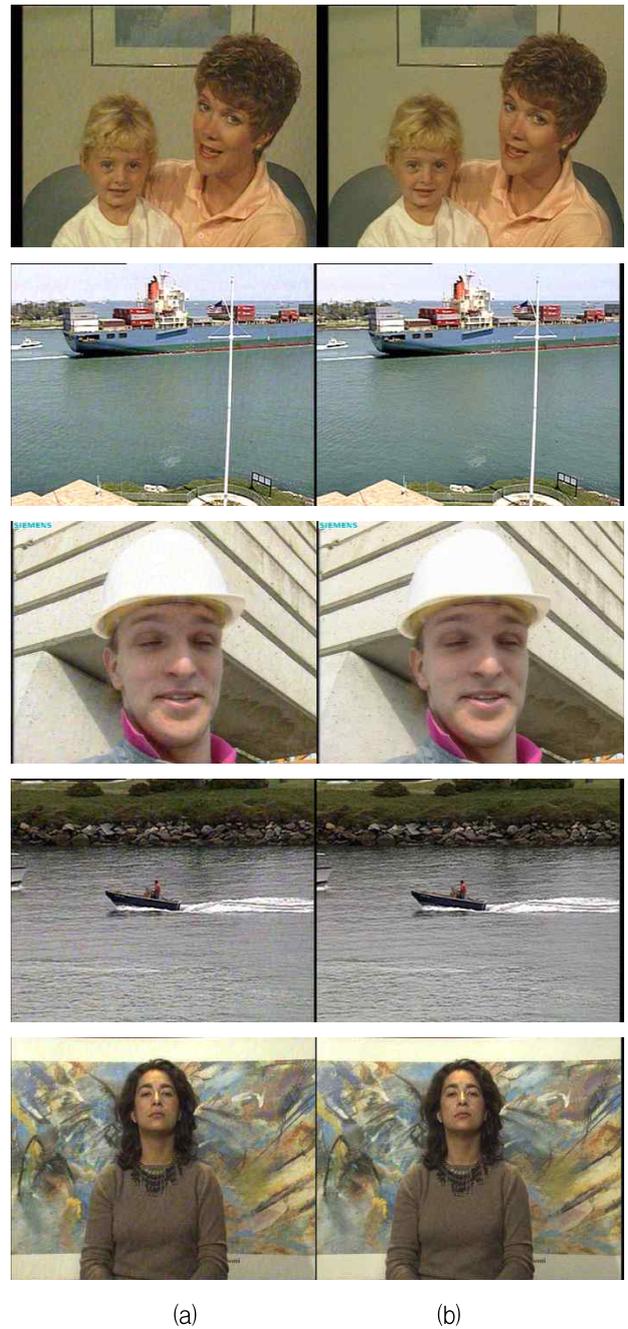


그림 13. (a) 좌측 열은 추정된 AR 모델을 기반으로 합성된 잡음 영상  $d_g$ , (b) 우측 열은 좌측열의 잡음 분산과 동일하도록 AWGN을 추가한 잡음 영상  $d_g^*$

Fig. 13. (a) The left column is for synthesized noisy images  $d_g$  based on the estimated AR model and (b) the right column is for noisy images  $d_g^*$  corrupted with AWGN of the same noise variance observed in the left column.

기반으로 학습을 수행한 방법과 적응적 Wiener 필터링 방법의 결과의 차이가 표 1에 비해 확연히 줄어든 반면,



(a)



(b)



(c)



(d)

그림 14. (a) 알려진 원 영상 (b) AR 잡음을 추가한 영상 (c) 적응적 Wiener필터링 결과 (d) AR기반 학습 필터 결과

Fig. 14. (a) Known original image, (b) noisy image corrupted with AR noise, (c) adaptive Wiener filtering result, and (d) AR-based trained filtering result.



(a)



(b)



(c)



(d)

그림 15. (a) 잡음이 있는 녹화된 아날로그 영상 (b) 적응적 Wiener 필터결과 (c) AWGN기반 학습 필터 결과 (d) AR기반 학습 필터 결과

Fig. 15. (a) Noisy recorded analog video, (b) adaptive Wiener filtering result, (c) AWGN-based trained filtering result, and (d) AR-based trained filtering result.

AR 모델을 기반으로 합성된 영상을 이용한 경우 개선된 PSNR을 얻을 수 있었다. 이는 효과적인 잡음 제거를 위해서는  $d_g$ 의 잡음 분산을 정확히 예측하는 것만으로는 충분하지 않으며, 잡음의 연관성을 고려하는 것이 필요함을 보여주는 결과이다.

그림 14는  $d_g$  영상에 적응적 Wiener 필터링<sup>[14]</sup>을 수행한 결과 영상과 AR 모델을 기반으로 학습기반 최적 자승 필터링을 수행한 결과를 보여준다. AR기반의 제안된 학습기반 필터가 배경에 있는 세밀한 부분들을 더 잘 보존하면서 잡음을 제거하는 것을 알 수 있다.

그림 15는 실제 녹화된 영상에 제안된 학습기반 방법을 적용한 결과를 보여준다. 그림 15(c)와 그림 15(d)는 동일한 학습기반 필터이지만, 학습 영상으로  $d_g^*$ 와  $d_g$ 을 사용한 경우에 따라 각각 달라진다. AR 잡음을 활용하는  $d_g$ 를 사용한 경우가 배경의 잡음을 훨씬 더 효과적으로 제거하면서 세밀한 부분은 잘 보존하는 것을 볼 수 있다. 이는 제안된 방식이 잡음의 연관성을 고려하는 AR 모델에 기반을 둬으로써 잡음의 특성에 보다 건설하게 대응하면서, 학습을 통해 최적 자승 기반의 필터링을 수행할 수 있기 때문이다.

## V. 결 론

본 논문에서는 녹화된 아날로그 영상에 내재하는 잡음을 제거하기 위해서 먼저 실제 녹화된 영상의 잡음 특성을 분석하였다. Ground truth 영상을 활용한 영상의 잡음 특성 분석을 통해 잡음의 연관성을 고려한 잡음 제거 방법의 필요성을 파악했으며, 이를 위해 AR 모델에 기반을 둔 학습기반 필터링 방법을 제안하였다. 알려진 원 영상과 예측된 AR 모델에 바탕을 둔 합성된 잡음 영상을 활용하여 최적 자승 기반 필터링 계수를 학습함으로써 잡음 연관성을 고려한 필터링 제거 방식을 적용할 수 있었다. 제안된 방법은 무시할 수 없을 정도의 잡음 연관성을 가진 실제 녹화된 아날로그 영상의 잡음 제거에 효과적으로 활용될 수 있을 것으로 기대된다.

## 참 고 문 헌

[1] G. de Haan, T. G. Kwaaitaal-Spassova, and O. A. Ojo, "Automatic 2-D and 3-D noise filtering for

television receivers," Proc. Int. Workshop HDTV, Oct. 1994.

- [2] G. de Haan, T. G. Kwaaitaal-Spassova, M. Larragy, and O. A. Ojo, "Memory integrated noise reduction IC for television," IEEE Trans. Consumer Electronics, vol. 42, no. 2, pp. 175-181, May 1996.
- [3] Y. Huang and L. Hui, "An adaptive spatial filter for additive Gaussian and impulse noise reduction in video signals," Proc. Int. Conf. Information, Commun. & Signal Processing and Pacific-Rim Conf. Multimedia, pp. 523-526, Dec. 2003.
- [4] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," IEEE Trans. Circuits and Systems for Video Technol., vol. 15, no. 1, pp. 113-118, Jan. 2005.
- [5] D.-H. Shin, R.-H. Park, S. Yang, and J.-H. Jung, "Block-based noise estimation using adaptive Gaussian filtering," IEEE Trans. Consumer Electronics, vol. 51, no. 1, pp. 218-226, Feb. 2005.
- [6] M. Ghazal, A. Amer, and A. Ghayeb, "Structure-oriented spatio-temporal video noise estimation," Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, vol. 2, pp. 845-848, May 2006.
- [7] K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," IEE Proceedings Vision, Image and Signal Processing, vol. 146, no. 2, pp. 80-84, Aug. 1999.
- [8] P. Q. Tuan, "Histogram-based noise estimation," Appendix A of his Ph.D. Thesis, Delft University of Technology, 2006.
- [9] J. Sijbers, D. Poot, A. J. Dekker, and W. Pintjens, "Automatic estimation of the noise variance from the histogram of a magnetic resonance image," Phys. Med. Biol., vol. 52, pp. 1335-1348, 2007.
- [10] C. W. Therrien, Discrete random signals and statistical signal processing, Prentice-Hall, pp. 535-541, 1992.
- [11] B. T. Oh, C.-C. J. Kuo, S. Sun, and S. Lei, "Film grain noise modeling in advanced video coding," Proc. Visual Commun. and Image Processing, 2007.
- [12] L. Shao, H. Zhang, and G. de Haan, "An overview and performance evaluation of classification-based least squares trained filters," IEEE Trans. Image Processing, vol. 17, no. 10, pp. 1772-1782, Oct. 2008.
- [13] H. Hu and G. de Haan, "Trained bilateral filters and applications to coding artifacts reduction," Proc. IEEE Int. Conf. Image Processing, pp. 1325-328, Sept. 2007.

- [14] J. S. Lim, Two-dimensional signal and image processing, Prentice-Hall, 1990.

---

— 저 자 소 개 —

---



김 성 득(정회원)

1994년 경북대학교 전자공학과  
학사 졸업

1996년 한국과학기술원 전기 및  
전자공학과 석사 졸업

2000년 한국과학기술원 전기 및  
전자공학과 박사 졸업

2000년~2003년 LG전자 차세대 단말 연구소  
2003년~현재 안동대학교 정보전자공학교육과  
부교수

<주관심분야 : 영상신호처리, 영상시스템>



임 경 원(정회원)

1990년 고려대학교 전자공학과  
학사 졸업

1992년 한국과학기술원 전기 및  
전자공학과 석사 졸업

1998년 한국과학기술원 전기 및  
전자공학과 박사 졸업

1997년~2001년 LG전자 종합기술원 선임 연구원

2001년~2006년 Conexant systems Inc., USA,

Principle engineer

2006년~현재 LG전자 수석 연구원

<주관심분야 : 영상신호처리, 영상 시스템, 반도체>