

## 멀티레벨 홀로그래픽 저장장치를 위한 적응 EM 알고리즘

김진영\*, 이재진<sup>o</sup>Adaptive Threshold Detection Using  
Expectation-Maximization Algorithm for Multi-Level  
Holographic Data StorageJinyoung Kim\*, Jaejin Lee<sup>o</sup>

## 요약

본 논문은 멀티레벨을 가지는 홀로그래픽 저장 장치에서 EM (Expectation-maximization) 알고리즘을 이용한 적응 문턱전압검출기를 제안한다. 멀티레벨을 이용한 홀로그래픽 저장 장치의 경우 DC 오프셋의 정도에 따라 비적응 문턱전압검출기의 성능에 매우 심각한 영향을 미친다. EM 방법은 채널을 통과한 데이터를 이용해 Expectation step과 maximization step을 반복하면서 평균과 분산을 추정하는 방법이다. DC 오프셋이 있는 상황에서 제안된 방법을 적용하여 문턱값을 찾아내서 검출한 결과 일정한 한도 내의 DC 오프셋의 경우는 DC 오프셋이 없는 경우와 동일한 성능을 보였다.

**Key Words** : Adaptive threshold detector, DC offset, Expectation maximization, Misalignments, Multi-level holographic data storage

## ABSTRACT

We propose an adaptive threshold detector algorithm for multi-level holographic data storage based on the expectation-maximization (EM) method. In this paper, the signal intensities that are passed through the four-level holographic channel are modeled as a four Gaussian mixture with unknown DC offsets and the threshold levels are estimated based on the maximum likelihood criterion. We compare the bit error rate (BER) performance of the proposed algorithm with the non-adaptive threshold detection algorithm for various levels of DC offset and misalignments. Our proposed algorithm shows consistently acceptable performance when the DC offset variance is fixed or the misalignments are lower than 20%. When the DC offset varies with each page, the BER of the proposed method is acceptable when the misalignments are lower than 10% and DC offset variance is 0.001.

## 1. 서론

멀티레벨 홀로그래픽 데이터 저장 시스템은 차세대 고용량 저장장치로 주목받고 있다<sup>[1]</sup>. 한 픽셀에 1비트 보다 많은 데이터를 저장할 수 있으며, 이는

같은 수의 이진 디지털 데이터 이미지보다 많은 정보를 저장할 수 있는 멀티레벨 데이터를 담고 있다는 것이다. 홀로그래픽 저장장치 시스템에서는 저장장치 특성에 맞는 멀티레벨 코딩과 적절한 신호처리 방법이 필요하다<sup>[2,3]</sup>. 본 논문에서는 멀티레벨 홀

\* 본 연구는 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2011-0018016).

• 주저자 : 숭실대학교 정보통신전자공학부 정보저장 및 통신 연구실, kijiyou@ssu.ac.kr, 정회원

<sup>o</sup> 교신저자 : 숭실대학교 정보통신전자공학부 정보저장 및 통신 연구실 zlee@ssu.ac.kr, 종신회원

논문번호 : KICS2012-10-495, 접수일자 : 2012년 10월 16일, 최종논문접수일자 : 2012년 10월 23일

로그래픽 데이터 저장 시스템을 위해서 EM (expectation-maximization)을 기반으로 하는 적응형 threshold 검출 알고리즘을 제안한다.

EM 알고리즘은 반복적인 최대유사도 방법 중에 가장 많이 사용되는 방법이다<sup>4,6</sup>. EM 알고리즘은 멀티패스 채널의 추정, 혼합 심볼 검출과 타이밍 오프셋 판단 문제, 그리고 마르코프 모델의 수행에 적합하다<sup>7,9</sup>. EM 알고리즘은 expectation step (E-step)과 maximizing step (M-step)을 포함하는 반복적인 방법이다. E-step은 완벽한 데이터의 log likelihood 기대값은 관측값과 이전 반복 단계의 파라미터에 의해서 추정된다. 그리고 M-step은 밀도 파라미터의 추정된 완벽한 데이터를 기반으로 업데이트 된다. 완벽한 데이터는 밀도 파라미터의 대체값처럼 수행되고 그것은 추정 문제를 쉽게 한다.

홀로그래픽 저장장치에서 문턱전압 검출기 알고리즘이 사용될 때, DC 오프셋은 검출기 에러를 발생시킨다. 홀로그래픽 저장장치의 DC 오프셋은 빛의 세기와 초점 거리에 의해서 발생할 수 있다. 이 경우에 EM 기반의 적응 문턱전압 검출기 알고리즘은 검출 알고리즘의 성능을 향상시킬 수 있다. 싱글 레이어 홀로그래픽 저장장치를 위한 문턱전압 검출기 알고리즘에 대한 연구는 참고문헌 10, 11에 있다. 본 논문에서 우리는 다양한 DC 오프셋과 어긋남에 대한 환경에서 EM 기반의 적응 문턱전압 검출기 알고리즘의 성능을 비교하였다.

## II. 홀로그래픽 채널

홀로그래픽 채널은 CCD (charge-coupled device)를 통과한 형상으로 모델링 하였다<sup>12-14</sup>. 데이터가 기준빔을 통과한 후 나온 회절된 신호를 CCD를 이용하여 검출한다. 연속적인 PSF (point spread function)은 다음과 같이 모델링 된다.

$$h(x,y) = \frac{1}{\sigma_b^2} \text{sinc}^2\left(\frac{x}{\sigma_b}, \frac{y}{\sigma_b}\right)$$

여기서  $\sigma_b$ 는 빛이 블러 (blur) 가 되는 정도이다. 본 논문에서는 이 블러값을 1.85로 설정하였다.  $\text{sinc}(x,y)$ 는 다음과 같이 정의한다.

$$\text{sinc}(x,y) = \left( \frac{\sin(\pi x)}{\pi x} \frac{\sin(\pi y)}{\pi y} \right)$$

$(p, q)$  지점에서의 이산 PSF  $h[p, q]$  는 다음과 같다.

$$h[p, q] = \int_{q-\frac{1}{2}}^{q+\frac{1}{2}} \int_{p-\frac{1}{2}}^{p+\frac{1}{2}} h(x,y) dx dy$$

여기서 CCD 배열이 제대로 된 위치에 있지 않고, 어긋난 형태로 위치하게 되는 것을 어긋남 이라 하고,  $x$  축의 어긋남을  $m_x$ ,  $y$  축의 어긋남을  $m_y$ 로 정의할 때 연속적인 PSF는 다음과 같다.

$$h(x) = \frac{1}{\sigma_b^2} \text{sinc}^2\left(\frac{x+m_x}{\sigma_b}, \frac{y+m_y}{\sigma_b}\right)$$

입력 데이터  $d[p, q]$ 는 4레벨의 심볼을 가지고 있고, 채널을 통과하여 수신된 데이터  $r[p, q]$ 는 다음과 같이 정의된다.

$$r[p, q] = d[p, q] \otimes h[p, q] + n[p, q] + c_{DC}$$

여기서,  $n[p, q]$ 는 AWGN (additive white Gaussian noise) 이고,  $c_{DC}$ 는 DC 오프셋이다. 홀로그래픽 채널에서, 빛은 투과하여 그 세기를 측정하기 때문에, 이 빛의 세기가 페이지마다 달라질 것을 가정하여, DC 오프셋은 각 페이지에 일정하게 하였다.

본 논문에서는 알려지지 않은 DC 오프셋이 들어 있는 4개의 Gaussian 혼합 시그널로 모델링된 4레벨 홀로그래픽 채널을 통과한 데이터로 가정하였고, 최대유사도 방법을 이용해 문턱전압값을 추정하였다. 기본 문턱전압값은 DC 오프셋이 없는 데이터를 이용하여 만들었다.

## III. Expectation-Maximization 방법

우선  $Y$ 를 4레벨 홀로그래픽 신호의 세기로 정의하였고, 4개의 Gaussian 혼합 신호로 가정하였다. 이를 랜덤한 값으로 본다면 다음과 같다.

$$Y = \sum_{j=1}^4 \Delta_j Y_j, \quad \sum_{j=1}^4 \Delta_j = 1$$

여기서  $\Delta_j$ 는  $\{0,1\}$ 에 속해있고,  $j = 1,2,3,4$ 이다. 각각의  $Y_j \sim N(\mu_j, \sigma_j^2)$ 이다.  $\Delta_j$ 는 혼합확률이고,  $\mu_j$ 와  $\sigma_j^2$ 은 각각 Gaussian 분포의 평균과 분산이다.  $\hat{\mu}_j, \hat{\sigma}_j^2, \hat{\pi}_j$ 를 각각 Gaussian 분포의 평균과 분산, 그리고 혼합 확률의 추정치로 정의하였다. 샘플  $y_i, i = 1, \dots, N$ 가 주어졌을 때, 평균과 분산, 그리고 혼합 확률을 추정하는 EM 알고리즘은 다음과 같다.

**0번 과정 : 초기화 (Initialization)**

$n = 0$ 으로, 비 감소 정렬된 샘플  $y_1, y_2, \dots, y_N$ 을 4개의 서브셋으로 동등하게 나누고, 각각의 서브셋의 평균과 분산을 계산한다.  $\hat{\mu}_j$ 와  $\hat{\sigma}_j^2$ 을  $j$ 번째 서브셋의 평균과 분산으로 정의하고,  $\pi_j$ 는 0.25,  $n = n + 1$ 로 한다. 그 다음, 1번 과정으로 넘어간다.

**1번 과정 : Expectation 과정 (E-step)**

이전 단계의  $\pi_j, \mu_j, \sigma_j^2$ 을 이용하여, 신뢰성 정보  $\hat{\gamma}_{ij}$ 를 다음과 같이 계산한다.

$$\hat{\gamma}_{ij} = \frac{\pi_j \phi_{\mu_j, \sigma_j^2}(y_i)}{\sum_{j=1}^4 \pi_j \phi_{\mu_j, \sigma_j^2}(y_i)},$$

여기서  $i = 1, \dots, N, j = 1, \dots, 4$ 이고,  $\phi_{\mu_j, \sigma_j^2}$ 은  $(\mu_j, \sigma_j^2)$ 을 가지는 Gaussian 밀도이다.

**2번 과정 : Maximization 과정 (M-step)**

갱신된 신뢰도  $\hat{\gamma}_{ij}$ 를 이용하여 새로운 평균, 분산, 혼합분포를 다음과 같이 계산한다.

$$\hat{\mu}_j = \frac{\sum_{i=1}^N \hat{\gamma}_{ij} y_i}{\sum_{i=1}^N \hat{\gamma}_{ij}},$$

$$\hat{\sigma}_j^2 = \frac{\sum_{i=1}^N \hat{\gamma}_{ij} (y_i - \hat{\mu}_j)^2}{\sum_{i=1}^N \hat{\gamma}_{ij}},$$

$$\hat{\pi}_j = \frac{\sum_{i=1}^N \hat{\gamma}_{ij}}{N}$$

**3번 과정 : Convergence 과정**

아래의 식에 수렴할 때까지 1번 과정과 2번 과정을 반복한다. 본 논문에서 사용한 정지 기준은 이전 반복 단계의 평균과 갱신된 평균의 상대적 차이값의 합을 이용하였다.

$$\sum_{j=1}^4 \left( \frac{\mu_j - \hat{\mu}_j}{\hat{\mu}_j} \right)^2 < 10^{-5}$$

위의 식을 만족하면 갱신된 값을 최종값으로 설정하고 반복을 멈추고, 아니면  $\mu_j = \hat{\mu}_j, \sigma_j^2 = \hat{\sigma}_j^2, \pi_j = \hat{\pi}_j, n = n + 1$ 으로 설정하고 1번 단계로 간다.

**IV. 실험 결과**

실험을 위해서, 4개의 레벨을 가지는 멀티레벨 입력을 가지는 홀로그래픽 채널에 DC 오프셋이 존재하고, 홀로그래픽 채널의 블러  $\sigma_b$ 를 0.5로 하였다.  $128 \times 128$  크기의 페이지를 이용하였고, SNR (signal to noise ratio)은  $20 \log_{10} 1/\sigma_w$ 로 정의하였다. EM알고리즘을 이용한 적응문턱전압을 이용한 결과와 비적응 문턱전압을 이용한 결과를 비교하였다.

그림 1은 DC 오프셋이 0.01로 일정하게 발생할 때, 각 심볼의 히스토그램을 나타낸 것이다. EM 알고리즘을 이용한 적응 문턱전압검출기의 성능은 초기값에 의해 변한다. 따라서, 본 실험에서는 DC 오프셋값이 없을 때, 초기 평균값을 초기값으로 설정하였고, 분산은 0.1로 하였다. 그림 2부터 5까지의 결과에서, "IDLE" 상태는 어긋남이 없고, DC 오프셋값이 0이다.

그림 2는 DC 오프셋값이 일정할 때의 비적응 문턱전압검출기와 적응 문턱전압검출기의 BER 성능을 보여준다. 여기서는 SNR이 30일 때를 이용해서 실험하였다. DC 오프셋값이 0.03과 0.07일 때, 적응 문턱전압검출기의 성능은 IDLE 상태와 비슷하게 나온다. 반면, 비적응 문턱전압검출기의 성능은

매우 나빠진다. DC 오프셋이 0.1일 때 적응 문턱전압검출기의 성능이 조금 나빠지지만, 0.2가 되면 검출이 실패되는 정도의 성능이 나온다.

그림 3에는 DC 오프셋이 각 분산을 가지고 변할 때의 BER 성능이 나와 있다. 각 분산을 가지고 Gaussian 분포를 가지고 변하는 상황을 가정하여 실험하였다. 분산이 변할 때, 성능은 더 나빠진다. 여기서도 SNR이 30일 때 실험하였다. 적응 문턱전압검출기의 성능은 분산이 0.03을 가지고 변할 때만 좋은 성능을 보여준다. 분산이 0.03보다 커지면 비적응 문턱전압검출기의 성능도 나빠지만, 적응 문턱전압검출기의 성능도 매우 나빠진다.

그림 4에는 DC 오프셋이 0.07로 일정할 때, 여러 크기의 어긋남에 대하여 두 검출기의 성능을 비교하였다. x축과 y축의 어긋남이 각각 10%씩 들어가 있을 때, 적응 문턱전압검출기의 성능은 IDLE 상황과 비슷한 성능을 보여준다. 어긋남이 (20%, 20%) 있을 때는, BER 성능이 기준보다 약 2.5 dB 나빠졌고, (30%, 30%) 있을 때는, 검출 성능이 매우 나빠진다.

그림 5에는 DC 오프셋이 0.001의 분산을 가지고 변할 때, 여러 크기의 어긋남에 따른 두 검출기의 성능을 비교하였다. 적응 문턱전압검출기의 성능은 DC 오프셋이 일정하게 생길때와 비슷한 성능을 보여주는 반면, 문턱전압검출기의 성능은 매우 나빠진다. 적응 문턱전압검출기의 경우 DC 오프셋이 아주 크게 발생하지 않는다면, IDLE 상황과 비슷하게, 혹은 약간 성능이 나빠지지만, 비적응 문턱전압검출기의 경우 약간의 DC 오프셋에도 성능 열화가 어느 정도 크게 나오는 것을 볼 수 있다.

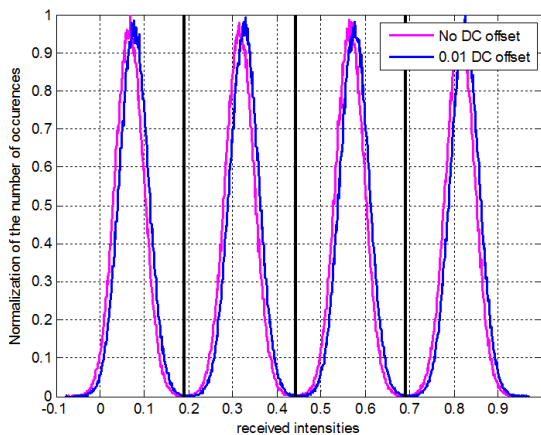


그림 1. 심볼의 히스토그램  
Fig. 1. Histogram of each symbol

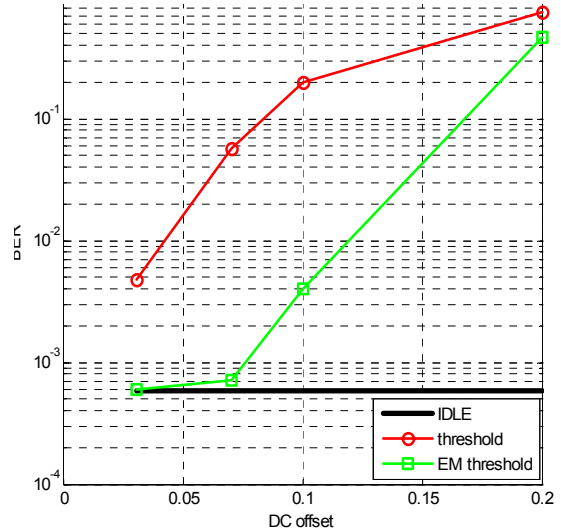


그림 2. DC 오프셋이 일정할 때 BER 성능  
Fig. 2. BER performance when the DC offset is constant

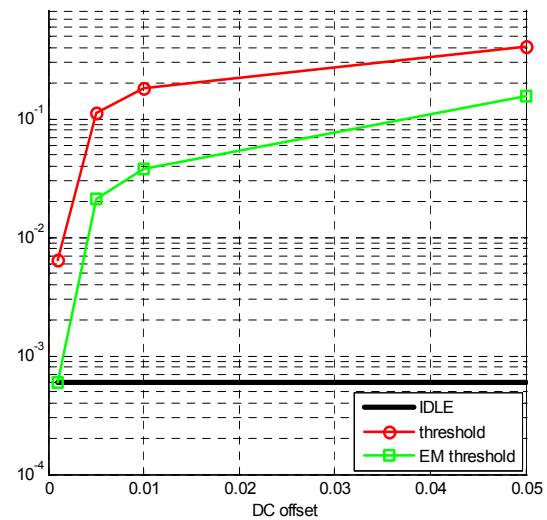


그림 3. DC 오프셋이 각 분산값을 가지고 변할 때 BER 성능  
Fig. 3. BER performance when the DC offset varies with some variance

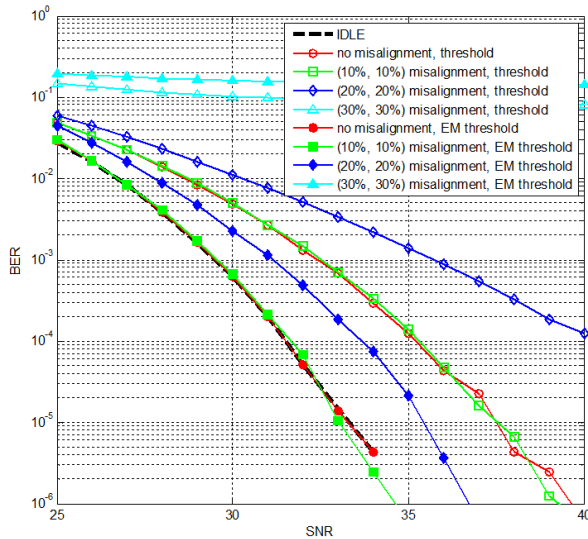


그림 4. DC 오프셋이 0.03일 때, 어긋남에 따른 BER 성능  
Fig. 4. Mis-alignment performance when DC offset is 0.03

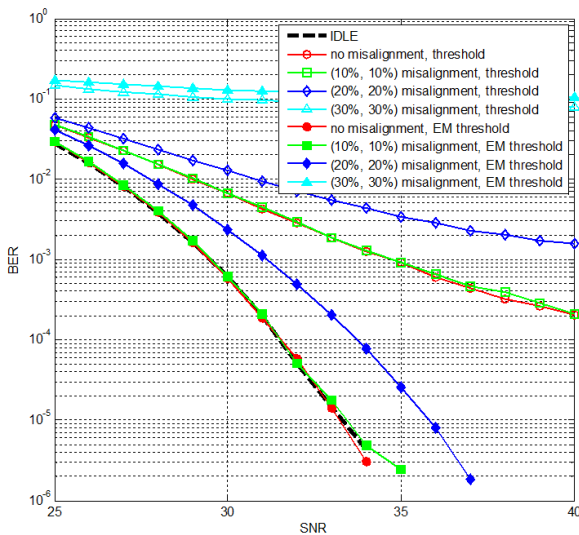


그림 5. DC 오프셋이 0.001의 분산을 가지고 변할 때, 어긋남에 따른 BER 성능  
Fig. 5. Mis-alignment performance when DC offset varies with the variance of 0.001

### V. 결론

본 논문에서는 멀티레벨 홀로그래픽 저장 장치에서 EM 알고리즘을 이용하여 적응 문턱전압검출기를 알아보고 비적응 문턱전압검출기와 성능을 비교하였다. 적응 문턱전압검출기는 EM 알고리즘을 이용하여 수신된 신호에서 변화되는 DC 오프셋값을 추정하여, 문턱값을 변화시켜준다. 따라서, 여러 DC 오프셋값과, 어긋남에 대하여 적응 문턱전압검출기와 비적응 문턱전압검출기의 BER 성능 비교해 보았다. 적응 문턱전압검출기의 경우 DC 오프셋이 일정하게 생성되고, 어긋남이 (20%, 20%) 이내로 발

생할 경우 충분히 검출이 이루어지는 것을 확인하였다. 페이지마다 DC 오프셋값이 변할 경우, 어긋남이 (10%, 10%) 이내로 생기고, 분산이 0.001 이내로 DC 오프셋이 변할 때, 충분히 사용가능한 검출이 이루어진 것을 확인하였다.

### 참고 문헌

- [1] S. G. Srinivasa, O. Momtahan, A. Karbaschi, S. W. McLaughlin, A. Adibi, and F. Fekri, "M-ary, binary, and space-volume multiplexing trade-offs for holographic channels," *Proc. IEEE Globecom* 2006, pp. 1-5, 2006.
- [2] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Info. Th.*, vol. 45, no. 4, pp. 1361-1391, Apr. 1999.
- [3] J. Kim, J. Wee, and J. Lee, "Error correcting 4/6 modulation codes for holographic data storage," *Jpn. J. Appl. Phys.*, vol. 49, no. 8, pp. 08KB04, Aug. 2010.
- [4] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc.*, vol. 39, no. 1, pp. 1-38, Jan. 1977.
- [5] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *Ann. Math. Stat.*, vol. 41, no. 2, pp. 164-171, Feb. 1970.
- [6] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406-2413, July 1994.
- [7] C. N. Georghiades and D. L. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection," *IEEE Trans. Commun.*, vol. 39, no. 1, pp. 54-61, Jan. 1991.

- [8] M. Feder and A. Catipovic, "Algorithms for joint channel estimation and data recovery-application to equalization in underwater communications," *IEEE J. Oceanic Eng.*, vol. 16, no. 1, pp. 42-55, Jan. 1991.
- [9] R. L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. the IEEE*, vol. 77, no. 2, pp. 257-286, Feb. 1989.
- [10] T.-H. Chao, U.S. Patent 20070211322A1, 2007.
- [11] X. A. Shen, A.-D. Nguyen, J. W. Perry, D. L. Huestis, and R. Kachru, "Time-domain holographic digital memory," *Science*, vol. 278, no. 1, pp. 96-100, Jan. 1997.
- [12] M. Keskinoz and B. V. K. V. Kumar, "Efficient modeling of volume holographic storage channels (VHSC)," *Proc. SPIE*, vol. 4090, no. 1, pp. 205-210, Jan. 2000.
- [13] D. E. Pansatiankul and A. A. Sawchuk, "Multi-dimensional modulation codes and error correction for page-oriented optical data storage," *Proc. SPIE*, vol. 4342, pp. 393-400, 2002.
- [14] G. Yang, J. Kim, and J. Lee, "Mis-alignment channel performance of error correcting 4/6 modulation codes for holographic data storage," *J. KICS*, vol. 35, no. 12, pp. 971-976, Dec. 2010.

김진영 (Jinyoung Kim)



2007년 2월 숭실대학교 기계  
공학과 학사

2009년 2월 숭실대학교 정보  
통신전자공학부 석사

2012년 8월 숭실대학교 정보  
통신전자공학부 박사

<관심분야> 스토리지 시스템,  
반복 복호 부호, 채널코딩

이재진 (Jaejin Lee)

(2010년 12월호 참조)