

논문 2009-46SP-5-14

효과적인 동영상 처리를 위한 움직임 보상 기반 잡음 예측

(Motion-Compensated Noise Estimation for Effective Video Processing)

송 병 철*

(Byung Cheol Song)

요 약

일반적인 동영상 처리에서 효과적으로 잡음을 제거하기 위해서는, 입력 동영상의 잡음 세기나 잡음 분산을 정확하게 찾아낼 필요가 있다. 그러나, 일반적으로 잡음 정보를 정확히 파악하기는 힘들다. 본 논문은 인접 잡음 영상간 움직임 보상에 기반한 정확한 잡음 분산 예측기법을 제안한다. 먼저, 입력 잡음 영상 내 각 블록에 대해 움직임 추정을 수행하고 최적의 움직임 보상 블록의 잔여 분산을 계산한다. 그리고, 구해진 최적 분산값과 근사한 분산값들을 적용적으로 평균화하고 적당히 스케일링함으로써, 그 영상에 대한 잡음 분산 예측치가 얻어진다. 실험결과를 통해 제안하는 방법이 매우 정확하게 잡음 세기를 예측하고 안정적임을 보인다.

Abstract

For effective noise removal prior to video processing, noise power or noise variance of an input video sequence needs to be found exactly, but it is actually a very difficult process. This paper presents an accurate noise variance estimation algorithm based on motion compensation between two adjacent noisy pictures. Firstly, motion estimation is performed for each block in a picture, and the residue variance of the best motion-compensated block is calculated. Then, a noise variance estimate of the picture is obtained by adaptively averaging and properly scaling the variances close to the best variance. The simulation results show that the proposed noise estimation algorithm is very accurate and stable irrespective of noise level.

Keywords : 잡음 예측, 잡음 제거, 움직임 보상

I. 서 론

잡음은 영상 개선, 압축, 포맷 변환, 주사 변환 등 많은 동영상 처리의 효과에 지대한 영향을 끼치는 요소이다. 따라서, 입력 동영상 내 잡음 세기를 정확하게 예측하는 것은 상기 영상처리의 효과를 향상시키는데 매우 중요하다.

동영상 신호는 획득 단계나 전송 단계에서 잡음에 의

해 자주 왜곡된다. 예를 들면, CCD (Charge-Coupled Device) 카메라 센서에서 영상 신호를 획득할 때나 아날로그 채널을 통해 신호를 전송할 때 수신단에서 잡음이 발생하곤 한다. 잡음 신호는 일반적으로 i.i.d (independent, identically distributed) 부가 백색 가우시안 잡음 (Additive White Gaussian Noise: AWGN)⁽¹⁾이라고 가정한다:

$$I_n(i, j) = S_n(i, j) + \eta_n(i, j) \quad (1)$$

여기서, $S_n(i, j)$ 는 n 번째 프레임에서 (i, j) 에 위치한 원 영상 신호를, $I_n(i, j)$ 는 관측 영상 신호를, $\eta_n(i, j)$ 은 잡음 신호를 의미한다. 잡음 예측이란 입력 영상에 존재하는 $\eta_n(i, j)$ 의 표준편차 σ_η 혹은 분산을 예측하는 것을 말한다. 여기서 정확한 σ_η 예측치는 잡음제거 장

* 정회원, 인하대학교 전자공학부

(School of Electronic Engineering, Inha University)

※ 이 논문은 지식경제부와 한국산업기술진흥원의 전략기술인력양성사업으로 수행된 연구결과이며, 이 논문은 2009년 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2009-0071385).

접수일자: 2009년2월13일, 수정완료일: 2009년8월27일

치들과 연동되어 동작하는 동영상 처리 시스템들의 성능을 향상시키기 위해 반드시 필요하다^[2~5].

본 논문은 두 인접 잡음왜곡 영상들 간에 움직임 보상에 기반한 정확한 잡음 분산 예측 기법을 제안한다. 먼저, 영상 내 각 블록에 대해 움직임 추정을 수행하고, 최적 정합 블록의 잔여 (residue) 분산을 계산한다. 그런 다음, 최적 분산에 근사한 잔여 분산값들을 선택적으로 평균함으로써 영상 전체의 잡음 분산을 얻는다. 실험 결과를 통해 제안한 방법이 정확하고 안정적으로 잡음 세기를 예측함을 보인다.

본 논문의 구성은 다음과 같다. II장에서는 기존의 방법들에 대해 소개하고, III장에서는 제안한 방법을 기술한다. IV장에서 제안한 방법의 성능을 평가하고, 마지막 V장에서 결론을 맺는다.

II. 기존 방법

잡음 예측 방법들은 크게 두 가지로 분류할 수 있다: 프레임 내 예측 (intra-frame estimation)과 프레임 간 예측 (inter-frame estimation)⁽⁶⁻⁹⁾. 프레임 내 예측 방법들은 다시 평활화 (smoothing)기반 방식과 블록 기반 방식들로 나눌 수 있다. 평활화 기반 방법들에서는 일반적으로 평활화된 즉, 저대역 필터링된 영상과 입력 영상과의 차를 잡음이라고 가정한다. 특정 문턱값보다 작은 기울기 (gradient)를 갖는 화소들에 대해서만 상기 차 값들을 이용하여 최종 잡음 예측치가 구해진다. 블록 기반 방법들은 영상 내 각 블록들에 대해 분산값들을 구하고, 작은 분산값들에 대해서만 평균을 취하는 방식으로 잡음 세기를 예측한다.

Olsen은 AWGN의 표준편차를 예측하기 위한 여러 방법들을 서로 다른 대조 (contrast), 에지, 텍스처를 갖는 다양한 영상들에 적용하여 실험 및 분석을 수행하였다⁽⁷⁾. 그는 잡음-왜곡 정도가 심한 영상들에 대해서는 평활화 기반 기법들이 좋은 성능을 보이는 반면, 많은 연산량이 요구되고 영상에 따른 상세 조율 등이 필요한 단점을 지님을 보였다. 더욱이, 평활화 기반 기법들은 복잡하고 섬세한 텍스처를 갖는 영상들에 대해서는 만족스러운 성능을 보이지 못함을 보였다. 한편, 블록 기반 기법들은 평활화 기반 기법들에 비해 적은 연산량을 갖지만, 평탄한 영역에 대해서는 과도한 잡음을, 복잡한 영역에 대해서는 과소한 잡음을 예측하는 경향이 있다.

Amer와 Dubois는 영상 내 가장 평활한 영역의 블록

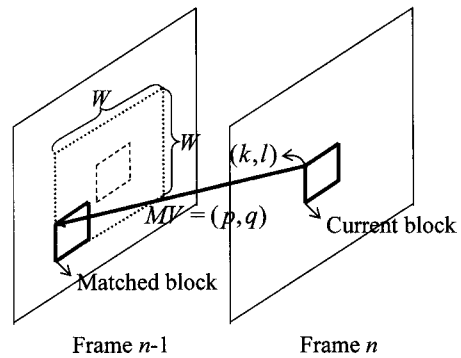


그림 1. 인접 프레임 간 움직임 보상

Fig. 1. Motion compensation between the adjacent frames.

들의 분산값들로부터 영상 전체의 잡음 분산을 예측하는 소위 블록기반 기법을 제안했다^[8]. 이 방법은 영상의 구조적 특성을 고려한 방식으로서, 영상 내 평활하지 못한 블록들을 선택하여 후보에서 제거하는 영상 특성 분석기를 이용한다. 그러나 전반적으로 텍스처가 분포한 영상에서는 예측 효율이 떨어지는 단점을 지닌다.

저자들은 프레임 내 기법들의 문제점을 개선한 프레임 간 기법을 제안한 바 있다^[9]. 이전 방법은 특정한 해상도 움직임 추정을 이용한 모델 기반 잡음 예측방식이다. 각 프레임에서 움직임 추정 오차가 작은 블록들을 이용하여 잡음 분산값을 예측하였다. 그러나 이전 방식의 잡음 예측 모델은 실제 잡음과 상당한 오차를 보이는 경우가 가끔 발생하였다. 본 논문에서는 이런 단점을 보완한 통계적 분석에 기반하여 보다 일반화된 움직임 보상기반 잡음 예측 기법을 제안한다.

III. 제안한 방법

1. 기본 개념

움직임 예측은 현재 프레임 (프레임 n) 내 각 블록에 적용되고, 이전 프레임 (프레임 n-1)의 적당한 탐색영역에서 현재 블록에 대한 최적 정합 블록을 찾는다 (그림 1). 그림 1에서, 현재 블록의 좌측 최상단 화소 위치는 (k, l) 이며, 탐색 영역 Ω 은 $W \times W$ 로 정의한다. 일반적인 정합 오차 측정자 SAD (Sum of Absolute Difference)가 최소가 되는 블록을 최적 정합 블록으로 선택한다 (식 (2)와 (3) 참조).

$$SAD(p, q) = \sum_{x, y} |I_n(k+x, l+y) - I_{n-1}(k+x+p, l+y+q)| \quad (2)$$

$$SAD_{\min} = \min_{(p', q') \in \Omega} SAD(p', q') \quad (3)$$

또한, 움직임 보상 잔여 블록의 (i, j) 에 위치한 화소를 다음과 같이 정의할 수 있다.

$$\begin{aligned} \delta(i, j) &= I_n(i, j) - I_{n-1}(i+p, j+q) \\ &= S_n(i, j) - S_{n-1}(i+p, j+q) + \eta_n(i, j) - \eta_{n-1}(i+p, j+q) \\ &\equiv \delta_S(i, j) + \delta_\eta(i, j). \end{aligned} \quad (4)$$

식 (4)에서 $\delta_S(i, j)$ 와 $\delta_\eta(i, j)$ 은 각각 원 신호와 잡음 성분의 잔여들이다. 현재 블록이 움직임이 전혀 없거나, 움직임 보상이 완벽하다면, $\delta_S(i, j)$ 은 0일 것이다. 따라서 AWGN을 가정하면 $\delta_\eta(i, j)$ 은 정규 분포를 갖는 두 독립 확률 변수의 선형 조합이라고 할 수 있다. 이후부터는 설명의 편의를 위해 각 성분의 좌표는 표기하지 않는다. 이럴 경우, 확률 변수 δ_η 는 두 확률변수 η_n 과 η_{n-1} 의 차 즉, $\eta_n - \eta_{n-1}$ 으로 정의할 수 있다. 따라서 η_n 과 η_{n-1} 은 모두 평균은 0이면서, 각각 $\sigma_{\eta_n}^2$ 과 $\sigma_{\eta_{n-1}}^2$ 의 분산값을 갖는다고 할 수 있다.

한편, 확률변수 이론에 다음 정리가 존재한다⁽¹⁰⁾:

정리:

X_1, X_2, \dots, X_K 가 독립된 정규분포의 확률 변수이고, X_k 는 평균 μ_k 와 표준편차 σ_k 를 갖는다고 하면, 그 선형 조합 $Y, a_1X_1 + a_2X_2 + \dots + a_KX_K$ 은 평균 $\mu_Y = a_1\mu_1 + a_2\mu_2 + \dots + a_K\mu_K$ 와 분산 $\sigma_Y^2 = (a_1)^2\sigma_1^2 + (a_2)^2\sigma_2^2 + \dots + (a_K)^2\sigma_K^2$ 를 갖는 정규 분포의 확률 변수이다.

상기 정리에 따르면, δ_η 의 평균은 0이고,

$$\sigma_{\delta_\eta}^2 = \sigma_{\eta_n}^2 + \sigma_{\eta_{n-1}}^2. \quad (5)$$

임을 알 수 있다. 다시 말해,

$$\sigma_{\eta_n}^2 = \sigma_{\delta_\eta}^2 - \sigma_{\eta_{n-1}}^2 \quad (6)$$

현재 블록이 움직임이 없거나 움직임 보상 자체가 완벽할 경우, 식 (4)와 식 (6)으로부터 다음 식 (7)이 얻어진다.

$$\sigma_{\eta_n}^2 = \sigma_{\delta}^2 - \sigma_{\eta_{n-1}}^2 \quad (7)$$

식 (7)을 이용하면, 우리는 현재 블록의 지역 잡음 분산 (local noise variance)을 유도할 수 있다.

그러나 통상적인 움직임 추정 기법들에 대해 완벽한 움직임 보상을 기대할 수는 없다. 그래서 잡음을 정확하게 예측하기 위해서는 비교적 움직임이 보상된 블록들이나 움직임 벡터가 $(0,0)$ 인 블록들의 잔여 분산값 (σ_δ^2)들만을 고려할 필요가 있다. 그럴 경우, 식 (7)에 근거하여 믿을만한 각 입력 프레임 전역 잡음 분산 (global noise variance)을 유추할 수 있다.

2. 제안한 방법

전역 잡음 분산 σ_η^2 을 예측하기 위해, 현재 프레임 내 움직임 보상이 가장 잘 된 상위 M 개의 블록들의 σ_δ^2 들을 평균하여 잡음 분산 예측치 σ_e^2 를 얻는다. 식 (7)로부터, 다음 식 (8)이 유도될 수 있다.

$$\sigma_e^2 = \frac{\sum (\sigma_{\delta[m]}^2 - \sigma_{\eta_{n-1}}^2)}{M} \quad (8)$$

식 (8)에서, $\sigma_{\delta[m]}^2$ 은 m 번째로 작은 SAD를 갖는 블록의 σ_δ^2 를 의미하며, $\sigma_{\eta_{n-1}}^2$ 은 이전 프레임 (프레임 $n-1$)의 σ_η^2 이며 이미 계산된 잡음 분산 예측값이다. 일반적으로 잡음은 stationary하기 때문에, 식 (6), (7), (8)에 따라 움직임 보상이 잘된 블록들로부터 얻어진 σ_e^2 는 전체 프레임을 대표하는 전역 잡음이라고 볼 수 있다. 그러나 파라미터 M 이 식 (8)의 과정에서 항상 고정될 경우, 변동이 심한 $\sigma_{\delta[m]}^2$ 값들 때문에 σ_e^2 가 과대하게 예측될 수도 있다.

이런 과대예측 문제를 방지하기 위해서는 비슷한 σ_δ^2 을 갖는 블록들만 식 (8)의 연산 과정에 포함시키는 것이 바람직하다. 다시 말해, 다음 식 (9)의 조건을 만족하는 σ_δ^2 만이 식 (8)의 연산 과정에 포함된다.

$$|\sigma_\delta - \sigma_{\delta_r}| < \tau \quad (9)$$

식 (9)에서 $\sigma_{\delta_r}^2$ 은 가장 작은 SAD를 갖는 블록의 σ_δ^2 이다. 이런 과정을 통해 식 (9)의 연산 과정이 안정화될 수 있다. 본 논문에서 τ 는 입력 영상과 무관하게 2로 설정된다. 조건 (9)에 기반한 다음 식 (10)을 통해 결과적으로 매우 정확하고 안정적인 분산 예측치를 계산할

수 있다.

$$\sigma_e^2 = \frac{\sum_{\sigma_{\delta(k,l)}^2 \in \Pi} (\sigma_{\delta(k,l)}^2 - \sigma_{\eta_{n-1}}^2)}{M'} \quad (10)$$

식 (10)에서 $\sigma_{\delta(k,l)}^2$ 은 (k,l) 위치에 있는 블록의 σ_{δ}^2 를 의미하며, Π 는 조건 (9)를 만족하는 $\sigma_{\delta(k,l)}^2$ 들의 집합이다. 또한, M' 는 Π 에 속한 원소들의 수이다.

식 (10)에 기반한 제안한 방법을 요약하면 다음과 같다:

- 1) 현재 프레임 (프레임 n)의 각 블록에 대해, 이전 프레임 (프레임 n-1)의 탐색 영역 ($W \times W$)에서 최소의 SAD를 갖는 최적 정합 블록을 찾는다.
 - 1-1) 각 블록의 $\sigma_{\delta(k,l)}^2$ 를 계산한다.
 - 1-2) 모든 $\sigma_{\delta(k,l)}^2$ 들 중에서, 최소 SAD에 대응하는 $\sigma_{\delta(k,l)}^2$ 를 찾는다.
- 2) 조건 (9)를 만족하는 $\sigma_{\delta(k,l)}^2$ 들을 고른다.
- 3) 식 (10)에 따라 현재 프레임의 σ_e^2 를 구한다.

단계 2)의 조건 (9)에서 현재 프레임에 대한 σ_{δ}^2 을 사용할 경우 프레임 지연문제가 발생하기 때문에, 본 논문에서는 이전 프레임의 σ_{δ}^2 를 대신 사용한다. 실험 결과 이에 따른 성능 저하는 아주 미미한 것을 확인하였다.

IV. 실험

성능평가를 위해 다양한 움직임들을 갖는 총 6개의 CIF (352x288) 동영상들을 사용하였다; *containership, football, foreman, mobile, news, stefan*. 각 동영상의 처음 50프레임들만 이용하였다. 객관적인 평가를 위해 총 6가지 잡음 표준 편차 (0, 3, 6, 9, 12, 15)를 갖는 AWGN을 원 영상에 삽입하였다.

제안한 기법과 기존의 방법 중 구조 기반 잡음 예측 방법 [8]을 비교하였다. 논문 [8]과 동일하게 두 파라미터 W 와 s 는 각각 5로 설정하였다. 제안한 방법에서 움직임 추정은 전역탐색을 사용하였으며, 탐색 영역은 32x32로 설정하였다. 움직임 탐색 블록 크기는 16x16이다.

잡음 예측 정확도를 판단하기 위해 본 논문에서는 예

측값과 참값의 차 즉, 예측오차 $E_n = |\sigma_{e_n} - \sigma_{\eta}|$ 를 계산한다. 그 예측 오차의 평균과 분산을 각각 식 (11)과 (12)처럼 계산하고, 이를 통해 제안한 기법의 성능을 평가한다.

$$\mu_E = \frac{\sum_{n=1}^N E_n}{N} \quad (11)$$

$$\sigma_E^2 = \frac{\sum_{n=1}^N (E_n - \mu_E)^2}{N} \quad (12)$$

여기서, N 은 해당 동영상의 총 프레임 수이다.

표 1은 6개 영상들에 대해 평균 μ_E 와 σ_E 를 구한 것이다. 표 2는 6개 잡음 세기들에 대한 평균값들이다. 이 표들을 보면, 제안한 방법이 [8]보다 작은 μ_E 와 σ_E 를 가짐을 알 수 있다. 또한, 제안한 방법은 잡음 세기는 물론 영상들의 특성과 무관하게 고른 성능을 보인다.

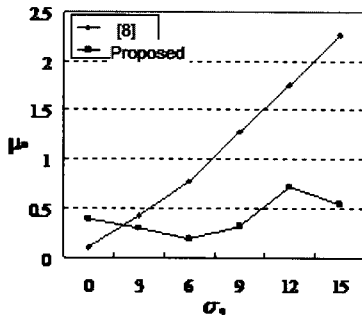
그림 2 (a)는 σ_{η}^2 이 0인 경우를 제외하고는 제안한 방법이 입력 잡음과 상관없이 [8]보다 정확하게 예측치를

표 1. 영상들에 대한 평균 μ_E 와 σ_E
Table 1. The averaged μ_E and σ_E for 6 sequences.

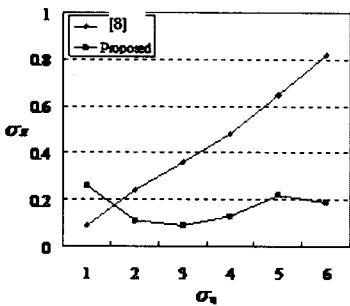
	[8]		Proposed	
	μ_E	σ_E	μ_E	σ_E
0	0.11	0.09	0.39	0.26
3	0.43	0.24	0.30	0.11
6	0.78	0.36	0.19	0.09
9	1.28	0.48	0.32	0.13
12	1.76	0.65	0.72	0.22
15	2.27	0.82	0.55	0.19
Average	1.10	0.44	0.55	0.19

표 2. 잡음 세기에 대한 평균 μ_E 와 σ_E
Table 2. The averaged μ_E and σ_E for 6 noise levels.

	[8]		Proposed	
	μ_E	σ_E	μ_E	σ_E
0	1.15	0.42	0.64	0.21
3	1.09	0.52	0.77	0.31
6	1.07	0.45	0.61	0.28
9	0.98	0.40	0.42	0.16
12	1.26	0.40	0.46	0.06
15	1.06	0.45	0.43	0.12
Average	1.10	0.44	0.55	0.19



(a)



(b)

그림 2. 제안한 방법과 [8]과의 비교 와 (a) μ_E vs. σ_n .
(b) σ_E vs. σ_n .
Fig. 2. Comparison of the proposed algorithm and [8]
(a) μ_E vs. σ_n . (b) σ_E vs. σ_n .

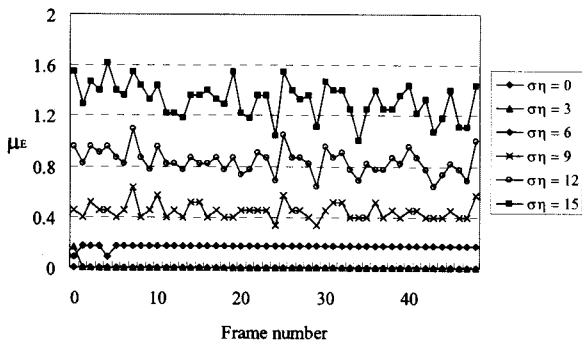


그림 3. 다양한 잡음 분산값들에 따른 평균 예측오차
Fig. 3. Average error over time for various noise variances.

연고 있음을 보인다. σ_n^2 이 0인 경우도 0.3미만의 미미한 예측 오차를 보이므로 무시할 수 있다. 그림 2 (b)는 제안한 방법이 [8]에 비해 아주 안정적임을 보인다. 그 예로서, 그림 3은 news 영상에서 안정적인 μ_E 의 시간 축 변이를 보이고 있다.

제안한 방법은 기본적으로 움직임 보상에 기반한 프레임 간 잡음 예측방식이기 때문에, [8]보다 많은 연산량을 필요로 하는 게 사실이다. 그러나 three-step search나

diamond search같은 고속 움직임 탐색을 사용하면 연산량을 크게 줄일 수 있다.

또한 제안한 방법이 동영상 부호기와 함께 사용될 경우 부호화를 위해 이미 예측된 블록 단위 움직임 벡터들을 재사용할 수 있으므로, 별도의 연산없이 움직임 추정에 의한 연산량을 상쇄할 수 있어 결과적으로 [8]보다도 연산량이 작을 수 있다.

V. 결 론

본 논문은 움직임 보상 기반의 잡음 예측 기법을 제안한다. 먼저, 영상 내 각 블록에 대해 움직임 추정을 수행하고, 최적 정합 블록의 잔여 분산을 계산한다. 그리고 나서, 최적 블록 분산에 근접한 잔여 분산들을 적용적으로 평균함으로써 해당 영상의 전역 잡음 분산을 얻을 수 있다. 실험을 통해 제안한 방법이 매우 안정적으로 잡음을 예측하고 기존의 방법보다 우수한 예측 성능을 보임을 확인하였다. 제안한 방법은 디지털 TV나 PVR같은 저장 매체에서 아날로그 입력 영상이나 잡음에 노출된 영상들을 압축/저장하거나 후처리를 수행하기 전에 효과적으로 잡음을 예측하여 이어지는 동영상 처리의 효과를 높일 수 있다.

참 고 문 헌

- [1] G. De Haan, *Video processing for multimedia systems*, University Press Eindhoven, Eindhoven, The Netherlands, 2003.
- [2] K. J. Boo and N. K. Bose, "A motion-compensated spatio-temporal filter for image sequences with signal-dependent noise," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 3, pp. 287-298, 1998.
- [3] S. D. Kim and J. B. Ra, "Efficient block-based video encoder embedding a Wiener filter for noisy video sequences," *Journal of Visual Communication and Image Representation*, vol. 14, no. 1, pp. 22-40, 2003.
- [4] 엄일규, "웨이블릿 영역에서 이변수 가우스 모델을 이용한 영상 잡음 제거," *전자공학회논문지-SP*, vol. 45, no. 6, pp. 57-63, 2008.
- [5] 엄일규, 김유신, "가우스 혼합모델과 형태학적 필터를 이용한 잡음 제거," *전자공학회논문지-SP*, vol. 41, no. 1, pp. 29-36, 2004.
- [6] K. Rank, M. Lendi, and R. Ubenhauer,

- "Estimation of image noise variance," *Proc. IEE Vis. Image Signal Process.*, vol. 146, no. 2, pp. 80-84, 1999.
- [7] S. I. Olsen, "Estimation of noise in images: an evaluation," *Graphical Models and Image Processing*, vol. 55, no. 4, pp. 319-323, 1993.
- [8] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 113-118, 2005.
- [9] B. C. Song and K. W. Chun, "Motion-compensated noise estimation for efficient pre-filtering in a video encoder," *IEEE International Conference on Image Processing*, 211-314, Barcelona, Spain, 2003.
- [10] E. R. Dougherty, *Probability and statistics for the engineering, computing, and physical sciences*, Chapter 5, pp. 229-230, Prentice-Hall, 1990.

 저 자 소 개



송 병 철(정회원)

1994년 한국과학기술원 전기 및 전자공학과 졸업 (학사).

1996년 한국과학기술원 전기 및 전자공학과 졸업 (석사).

2001년 한국과학기술원 전기 및 전자공학과 졸업 (공학박사).

2001년~2008년 삼성전자 디지털미디어연구소 책임연구원

2008년~현재 인하대학교 전자공학부 교수

<주관심분야: 영상 신호처리, 영상시스템/SoC>