

저조도 감시 카메라 비디오의 잡음 제거를 위한 적응적 시공간 평활화 파라미터 추정에 관한 연구

김대회*, 최재영*, 노용만*

*한국과학기술원 전기 및 전자공학과

e-mail : dhkim10@kaist.ac.kr, jygchoi@kaist.ac.kr, ymro@ee.kaist.ac.kr

Estimating parameter of adaptive spatio-temporal smoothing for noise reduction in low light surveillance video

Dae Hoe Kim*, Jae Young Choi*, Yong Man Ro*

*Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology

요 약

본 논문은 SNR 이 매우 낮은 저조도 영상의 잡음 제거를 위한 새로운 기술을 제안한다. 제안하는 기술은 입력 영상에서 파라미터를 자동/적응적 방식으로 추정하는 알고리즘을 특징으로 한다. 제안하는 기술의 효율성을 검증하기 위해 실질적인 환경에서 취득한 저조도 동영상들을 가지고 실험을 수행하였다. 실험을 통해 제안하는 기술을 활용하여 적응적으로 추정된 파라미터가 필터링(filtering) 성능을 잘 유지시킴을 검증하였다. 또한 기존 연구들과 비교할 때 저조도 동영상의 명암 대비 향상과 잡음 제거에 우수한 결과를 보임을 검증하였다.

1. 서론

최근 인적이 드문 곳에서 각종 범죄 발생이 급증함에 따라 감시 카메라(surveillance camera)의 중요성이 커지고 있다. 감시카메라는 지속적인 감시가 목적이기 때문에 환경 변화에 강인하게 촬영되는 것이 중요하다. 하지만 중저가 카메라의 경우 빛의 양이 매우 적은 밤에는 촬영된 물체(object)가 잘 보이지 않게 된다. 이렇게 사물이 거의 보이지 않을 정도로 어두운 상황에서 촬영된 동영상은 명암대비가 떨어지고 또한 낮은 신호대잡음비율(Signal-to-Noise Ratio, SNR)을 갖는 것이 특징이다. 따라서 일반적으로 사용하는 히스토그램(histogram)을 이용한 밝기 보정을 수행할 경우 잡음도 함께 증폭이 되어 영상의 질이 매우 떨어지는 문제점이 있다.

일반적인 동영상의 경우 인접 프레임(frame)간의 유사성이 크기 때문에 정지 영상과는 다르게 시간 축으로 사용할 수 있는 정보가 존재한다. 이런 시간적 정보를 활용하여 영상의 엣지(edge)는 보존하면서 잡음은 제거하기 위한 다양한 연구가 수행되었다. 우선 영상의 움직임의 궤적을 고려하는 움직임 보상 시공간 필터링(motion compensated spatio-temporal filtering)[7]이 연구되었다. 움직임 보상 필터링은 좋은 필터링 결과를 얻기 위해 움직임 추정(motion estimation)의 성능이 좋아야 한다. 하지만 잡음이 심한 경우 움직임 추정이 어렵다는 단점이 있다.

또한 3 차원 이방성 확산 방정식(anisotropic diffusion equation)[8]의 해를 구하여 비디오의 잡음을 제거하기

위한 연구도 진행 되었다. 하지만 확산 방정식의 해를 반복 연산(iteration)으로 구해야 하기 때문에 많은 연산이 필요하고 확산을 중지시키기 위한 기준을 정하기 어렵다는 단점이 존재한다.

저조도 동영상의 잡음 제거에 대한 연구로 양방향 적응적 시공간 축적(Bilateral Adaptive Spatio-Temporal Accumulation, ASTA)필터[9]가 있다. ASTA 필터는 움직임이 검출될 경우 공간적 필터를 더 적용하고 움직임이 없을 경우 시간적 필터를 더 적용하여 모션 블러(motion blur)를 방지한다. ASTA 필터는 움직이는 카메라에 대한 영상을 처리하기 위해 전처리로 전역 움직임 검출이 필요하다는 단점이 있다. 반면 적응적 시공간 평활화[1]는 정지한 카메라와 이동하는 카메라에 대해 일반화된 알고리즘을 적용하여 전처리가 필요하지 않고 구조가 단순해 그래픽 프로세싱 유닛(Graphic Processing Unit, GPU)을 통한 병렬 처리가 가능하다는 장점이 있다.

적응적 시공간 평활화가 기존 연구들에 비해 매우 어두운 영상의 필터링에 우수한 결과를 보였지만 사용자가 중요한 파라미터를 촬영하는 영상의 환경에 맞게 실험적으로 결정해야 하는 단점이 존재한다. 그렇기 때문에 적응적 시공간 평활화를 감시 카메라와 같이 밝기가 지속적으로 변화하는 환경에 적용하려면 파라미터가 수시로 갱신되어야 하고 그렇지 못할 경우 좋은 필터링 결과를 얻을 수 없어 실질적인 구현에 제약이 존재한다.

본 논문은 위에서 설명한 기존 기술에서 필터링 결과가 실험적으로 결정된 파라미터에 의존하는 한계를

해결하기 위해 입력 영상을 분석하여 파라미터를 자동/적응(automatic/adaptive)적으로 추정하는 방식을 제안한다. 본 논문에서 제안하는 방법은 파라미터의 추정을 통해 촬영 환경이 변하더라도 추가적인 파라미터 값을 설정할 필요 없이 우수한 품질의 출력 영상을 얻을 수 있는 것이 특징이다.

파라미터 추정의 필요성과 제안하는 방법의 성능 검증에 대한 실험을 수행하였다. 본 논문에서는 밝기가 다르게 촬영된 영상들을 고정된 파라미터와 입력 영상으로부터 추정된 파라미터로 각각 필터링 하도록 실험을 설계하였다. 실험 결과 고정된 파라미터를 사용한 경우에는 밝기가 변함에 따라 블러(blur)가 심하게 생기거나 잡음이 제거되지 않는 반면, 추정된 파라미터를 사용한 경우 밝기가 심하게 변하더라도 잡음은 잘 제거되고 엷지는 보존되는 것을 확인하였다.

본 논문은 2장에서 적응적 시공간 평활화[1]에 대해 설명하고, 3장에서 입력 영상에 적응적으로 파라미터를 정하는 방법을 제안하였다. 그리고 4장에서 제안한 방법에 대한 실험을 수행하였다.

2. 적응적 시공간 평활화

이 장에서는 제안하는 파라미터 추정 방법에 대한 이해를 돕기 위해 적응적 시공간 평활화[1] 과정에 대해 설명한다.

적응적 시공간 평활화는 구조 적응적 이방성 필터링[2]을 저조도 동영상에 맞게 수정한 것이다.

적응적 시공간 평활화가 대비를 향상시키지는 않기 때문에 대비를 향상시키는 명암 변환이 필요하다. 이런 명암 변환을 톤 매핑(tone mapping)이라고도 하는데 히스토그램 균등화가 간단한 예이다. 하지만 히스토그램 균등화는 밝은 부분에서 화소 값의 포화가 일어나 정보가 손실되는 단점이 있다.

밝은 영역의 명암 구조는 보존하면서 어두운 영역의 동적 범위는 늘려주기 위해 대비가 제한된 히스토그램 균등화를 사용하였다[4]. 대비가 제한된 히스토그램 균등화는 클리핑 한계(clipping limit)C [4]를 통해 사상함수(mapping function)의 기울기를 제한하여 어두운 부분의 대비를 향상시킨다.

톤 매핑을 통해 대비가 향상된 비디오를 이용해 적응적 시공간 평활화를 수행한다. 적응적 시공간 평활화는 각각 화소 주위의 명암 구조에 맞게 회전 및 크기 변환된 가우시안 커널(gaussian kernel)과 입력 영상의 가중치 평균으로 이루어진다. 임의의 시공간 점 $\mathbf{x}_0 = (x'_0, y'_0, t'_0)$ 에서 적응적 시공간 평활화를 통해 계산된 출력화소 값 $f_{out}(\mathbf{x}_0)$ 은 다음과 같이 정의 된다.

$$f_{out}(\mathbf{x}_0) = \frac{1}{\mu(\mathbf{x}_0)} \iiint_{\Omega} k(\mathbf{x}_0, \mathbf{x}) f_i(\mathbf{x}) d\mathbf{x}, \quad (1)$$

여기서 $\mu(\mathbf{x}_0)$ 는 커널의 합이 1 이 되게 하여 영상의 평균 값이 변하지 않게 하는 정규화 인자이고 Ω 는 \mathbf{x}_0 의 일정 범위 내에 있는 이웃 화소들이다.

엷지에서의 블러를 막기 위해서는 각 화소에 대해 주위의 명암도가 균일한 방향으로 폭이 넓고 엷지가 있는 방향으로 폭이 좁은 커널이 생성되어야 한다.

각 화소 주위의 명암도 정보를 얻기 위해 구조 텐서(structure tensor)를 이용한다[6].

구조 텐서의 고유 값(eigenvalue)과 고유 벡터(eigenvector)는 해당 화소 주위의 명암도 구조 정보를 나타낸다. 각 고유 값의 크기는 대응되는 고유 벡터 방향으로의 경사도(gradient)의 크기를 나타낸다.

커널 $k(\mathbf{x}_0, \mathbf{x})$ 는 아래의 수식을 활용해 각 화소 \mathbf{x}_0 에 대해 생성된다.

$$k(\mathbf{x}_0, \mathbf{x}) = e^{-\frac{1}{2} \sum_{i=1}^n \left(\frac{(\mathbf{x}-\mathbf{x}_0) \cdot \mathbf{v}_i}{\sigma(\lambda_i)} \right)^2}, \quad (2)$$

여기서 \mathbf{v}_i 는 구조 텐서의 고유 벡터로 커널의 축의 방향을 결정하고, λ_i 는 \mathbf{v}_i 와 대응되는 고유 값이다.

함수 $\sigma(\lambda_i)$ 는 커널의 \mathbf{v}_i 방향으로의 폭을 결정하는 감소 함수로 다음과 같이 정의된다.

$$\sigma(\lambda_i) = \begin{cases} (\sigma_{\max} - \sigma_{\min}) e^{-\frac{\lambda_i + 2}{d} + \sigma_{\min}}, & \lambda_i > 2d/5 \\ \sigma_{\max}, & \lambda_i \leq 2d/5, \end{cases} \quad (3)$$

여기서 σ_{\max} 과 σ_{\min} 은 커널의 최대 폭과 최소 폭을 각각 결정하고, d 는 \mathbf{v}_i 방향으로 커널의 폭을 결정하는 파라미터이다.

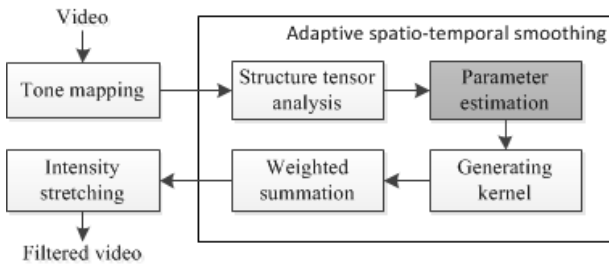
적응적 시공간 평활화가 명암 대비를 축소하는 경향이 있기 때문에 일반적으로 평활화 후에 명암대비 스트레칭을 수행한다.

3. 제안하는 파라미터 추정 방법

적응적 시공간 평활화[1]가 저조도 환경에서 촬영된 영상의 필터링에 우수한 성능을 나타내지만 촬영 환경이 변할 때마다 적절한 파라미터를 수동으로 결정해야 한다는 단점이 존재한다. 수동으로 설정해야 하는 파라미터는 감시 카메라와 같이 환경이 지속적으로 변화하고 사용자가 파라미터를 계속 조절하기 어려운 상황에 적응적 시공간 평활화를 적용하기 어렵게 만든다. 따라서 변화하는 환경에서도 필터링이 잘 수행되도록 입력 영상에 적응적으로 파라미터를 자동으로 추정하는 방법이 필요하다.

식(3)에서 파라미터 값 d 가 작을 경우 필터링이 되지 않고 지나치게 클 경우 가우시안 커널의 폭이 커져 평활화가 많이 되기 때문에 필터링 결과에 가장 큰 영향을 끼친다. 또한 영상의 특성에 따라 최적의 값이 민감하게 변화하기 때문에 적절한 파라미터 값 d 를 찾는 것이 전체 필터링 성능 향상에 중요한 역할을 한다고 볼 수 있다. 식(3)에서 나타낸 것과 같이 파라미터 값 d 는 각 고유 값들과 크기를 직접 비교하기 때문에 본 논문에서는 기존 필터링 방법에 추가적으로 구조 텐서의 고유 값 분석을 통해 d 를 추정하는 방법을 제안한다.

(그림 1)은 제안하는 필터링 시스템의 흐름도를 나타낸다. 영상의 임의의 화소에서의 경사도가 크면 구조 텐서의 고유 값도 함께 커진다. 그렇기 때문에 경사도의 크기가 다른, 엷지 부분의 화소와 명암도가 균일한 영역의 화소에 대한 고유 값은 크기의 분포가 다르다. 두 집단 사이의 경향을 분석하기 위해 아래



(그림 1) 제안하는 필터링 시스템 흐름도

의 수식을 통해 적응적 엣지 검출[5]을 적용하고 엣지가 아닌 화소들을 분류한다.

$$P_{non_edge} = \{ \mathbf{x} \mid \|\nabla f(\mathbf{x})\| < G_{th} \}, \quad (4)$$

$$G_{th} = percentile(\|\nabla f\|, \alpha), \quad (5)$$

여기서 ∇f 는 영상 f 의 경사도를, G_{th} 는 경사도의 크기에 대한 α 백분위수(percentile)를 나타내며 $\|\cdot\|$ 은 벡터의 크기를 의미한다. 또한 함수 $percentile$ 은 집합으로 이루어진 첫 번째 인자의 원소들에 대해 어떤 값보다 크기가 작은 원소 수의 전체 원소 수에 대한 비율이 α %인 원소를 구하는 함수이다. 백분위수를 통해 α 가 같더라도 영상마다 다른 임계 값을 가질 수 있게 한다. P_{non_edge} 는 경사도의 크기가 임계 값 G_{th} 보다 작아 엣지가 아닌 것으로 분류된 화소들의 집합을 의미한다. 또한 집합 λ_{non_edge} 를 아래와 같이 정의한다.

$$\lambda_{non_edge} = \{ \lambda_i \mid \lambda_i \in P_{non_edge} \}, \quad (6)$$

여기서 λ_{non_edge} 는 P_{non_edge} 에 포함된 화소들의 고유 값들을 원소로 포함하는 집합이다.

λ_{non_edge} 의 하위 50%에 해당하는 고유 값을 가지는 화소들은 실제로 균일한 영역에 있고, 나머지 화소들은 고유 값이 커질수록 엣지일 확률이 높다고 가정하였다. 이 가정하에 식(3)이 하위 50%의 고유 값에 대해 σ_{max} 가 되고 50%의 고유 값보다 커질수록 점차 σ_{min} 으로 수렴하도록 식(3)을 사용하는 경우에 파라미터 d 를 아래와 같이 구할 수 있다.

$$d = \frac{5}{2} median(\lambda_{non_edge}), \quad (7)$$

여기서 함수 $median$ 은 입력으로 주어진 집합에 포함된 원소들의 중간값($median$)을 구하는 함수이다.

본 논문에서는 커널의 폭이 고유 값이 커질수록 빠르게 감소하여 엣지 성분을 더 보존하게 하기 위해 식(3)을 아래와 같이 적용하였다[3].

$$\sigma(\lambda_i) = \begin{cases} (\sigma_{max} - \sigma_{min}) e^{-\frac{\lambda_i}{d} + 1} + \sigma_{min}, & \lambda_i > d \\ \sigma_{max}, & \lambda_i \leq d. \end{cases} \quad (8)$$

또한 식(8)을 사용할 때에 파라미터 d 는 아래와 같이 추정한다.

$$d = median(\lambda_{non_edge}). \quad (9)$$

그리고 적응적 시공간 평활화[1]는 처리하고자 하는 영상들은 매우 어둡다는 것을 전제로 한다. 그렇

기 때문에 경사도의 크기가 작아 고유 값 또한 작은 쪽으로 치우치는 경향이 있다. 이를 보완하기 위해 영상 전체 화소 수에 대한 영상의 가장 어두운 화소 수의 비율만큼 λ_{non_edge} 에서 작은 고유 값을 제외한 λ_{high} 에서 파라미터 d 를 아래의 식을 통해 추정하였다.

$$d = median(\lambda_{high}), \quad (10)$$

여기서

$$\lambda_{high} = \{ \lambda \mid \lambda > percentile(\lambda_{non_edge}, r) \}, \quad r = \frac{N_{min}}{N_{total}} \times 100(\%),$$

그리고 N_{min} 은 영상의 최소 명암 값을 갖는 화소 수이고 N_{total} 은 영상의 전체 화소 수이다.

위에서 제시한 파라미터 추정 방법은 적응적 시공간 평활화[1] 과정에서 계산되는 값들을 사용하기 때문에 효율적이다.

4. 실험

실험은 밝기가 변화하는 환경에서 파라미터가 고정된 적응적 시공간 평활화[1]와 입력 영상에서 파라미터를 추정하여 적응적 시공간 평활화를 각각 수행하여 필터링 결과의 차이를 비교하였다.

필터의 성능을 측정하기 위해 다음과 같이 화소 (x,y) 에서의 국소 대비(local contrast)를 정의한다[10].

$$c(x,y) = \frac{I_{max}(x,y) - I_{min}(x,y)}{I_{max}(x,y) + I_{min}(x,y)}, \quad (11)$$

여기서 I_{max} 와 I_{min} 은 (x,y) 에서 일정 범위 내에 있는 화소들의 최대, 최소값을 나타낸다.

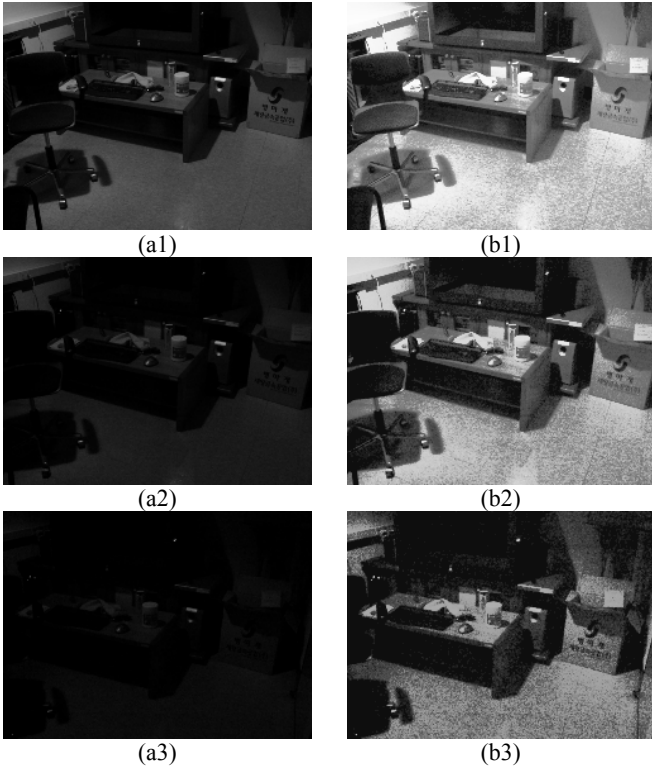
균일한 영역의 평활화가 잘 될수록 대비가 낮으므로 다음과 같이 일정 영역의 평균 대비를 정의한다.

$$C = \frac{\sum_{(x,y) \in Q} c(x,y)}{N}, \quad (12)$$

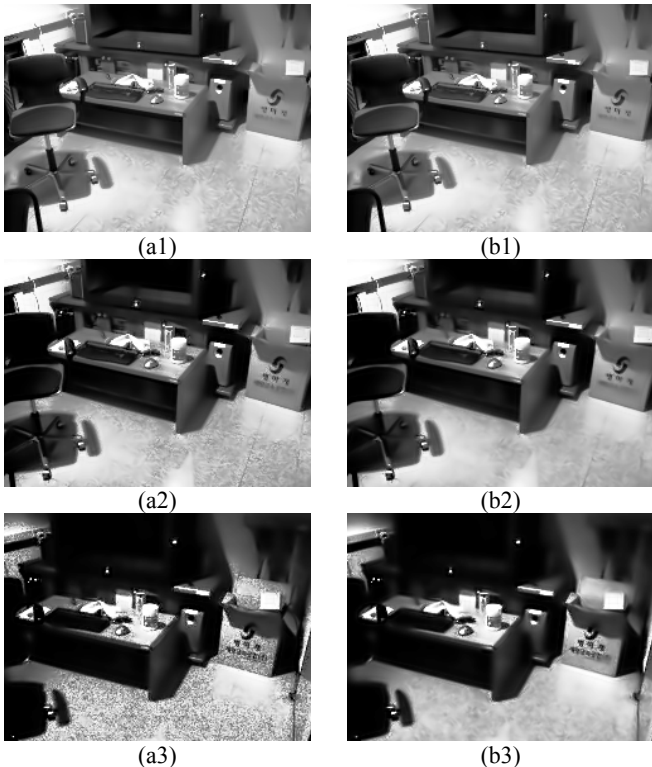
여기서 Q 는 영상의 영역을, N 은 영역 Q 에서의 화소 수를 의미한다. 필터의 성능이 우수할수록 균일한 영역의 평균 대비가 낮아야 한다.

실험에 사용한 카메라는 삼성 VLUU NV20 이다. 영상의 SNR 이 매우 낮은 특징으로 인해 손실 압축 방식으로 저장되는 동영상을 사용할 경우 데이터가 많이 손실되는 문제가 있었다. 이런 손실을 방지하기 위해 1024x768 해상도의 사진을 연속 촬영한 후 1/4로 축소하였다. 식 (8)에서 σ_{max} 와 σ_{min} 은 각각 2.5 와 0.1 을 사용하였다. 구조 텐서에 사용된 가우시안 커널 크기는 7x7x7 화소이고, 적응적 시공간 평활화에 사용한 커널의 크기는 13x13x13 화소이다.

(그림 2)는 실험에 사용한 영상과 톤 매핑된 영상이다. 보는 바와 같이 영상이 어두워질수록 잡음이 많아지는 것을 확인할 수 있다. (그림 3)은 밝기가 변하는 환경에서 고정된 파라미터 d 와 영상에서 추정된 파라미터 d 로 적응적 시공간 평활화[1]를 각각 수행한 결과이다. 고정된 파라미터는 비교적 밝은 영상에 대해 본 논문이 제안하는 파라미터 추정 방식으로 추정한 파라미터를 사용하였다. 실험 결과 적응적 시공간 평활화가 잡음 제거에 우수하지만 영상이 어두워져



(그림 2) 실험에 사용한 영상. (a1)~(a3): 밝기가 다르게 촬영된 입력 영상. (b1)~(b3): 톤 매핑된 영상.



(그림 3) 필터링 결과. (a1)~(a3): 고정된 파라미터 사용. (b1)~(b3): 영상에서 추정된 파라미터 사용.

<표 1> 균일한 영역에서의 평균 대비 측정값 비교

방법 \ 조도	높음	중간	낮음
톤 매핑	0.1634	0.1793	0.2373
기존 필터링	0.0565	0.0813	0.1927
제안하는 필터링	0.0565	0.0327	0.0571

잡음이 강해질 경우 필터링이 잘 되지 않았다. 반면 파라미터를 추정하여 평활화를 하였을 때에는 밝기가 변하더라도 필터링이 잘 되는 것을 확인할 수 있다.

<표 1>은 톤 매핑 후 필터링을 하기 전 영상, 고정된 파라미터로 필터링한 영상, 추정된 파라미터로 필터링한 영상에 대해 균일한 영역의 평균 대비를 측정 한 결과를 보여준다. <표 1>에서 조도가 높은 경우, 중간인 경우, 낮은 경우는 (그림 2)와 (그림 3)에서 각각 (a1)(b1), (a2)(b2), (a3)(b3)에 해당한다. <표 1>을 통해 제안한 알고리즘이 균일한 영역의 평활화를 가장 잘 수행함을 확인할 수 있다.

5. 결론

본 논문에서는 적응적 시공간 평활화의 파라미터를 입력 영상에서 추정하는 방법을 제안하였다. 제안하는 파라미터 추정 방식을 이용하여 변화하는 환경에서도 적응적으로 파라미터를 찾아 안정된 필터링 결과를 얻을 수 있었다. 실제 감시카메라에 적용하기 위해서는 처리 시간이 짧아야 한다. 적응적 시공간 평활화는 구조 텐서를 구하여 커널을 생성하고 실제 필터링을 하는데 매우 많은 연산이 필요하다. 하지만 앞의 과정들은 서로 독립적으로 연산이 가능하기 때문에 GPU 를 이용[1]하거나 전용 하드웨어를 설계해 구현할 경우 실시간 처리가 가능할 것이다.

참고문헌

- [1] H. Malm, M. Oskarsson, E. Warrant, P. Clarberg, J. Hasselgren, and C. Lejdfors, "Adaptive enhancement and noise reduction in very low light-level video," in *Proc. IEEE Int. Conf. Computer Vision*, pp. 1-8, 2007.
- [2] GZ. Yang, P. Burger, DN. Firmin, and SR. Underwood, "Structure adaptive anisotropic image filtering," *Image and Vision Computing*, vol. 14, no. 2, pp. 135-145, Mar. 1996.
- [3] H. Malm, and E. Warrant, "Motion dependent spatiotemporal smoothing for noise reduction in very dim light image sequences," in *Proc. Int. Conf. Pattern Recognition*, pp. 954-959, 2006.
- [4] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Geer, B. teer Haar Romeny, J. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," in *Computer Vision, Graphics and Image Processing*, vol. 39, pp. 355-368, 1987.
- [5] SC. Tai, and SM. Yang, "A fast method for image noise estimation using Laplacian operator and adaptive edge detection," in *Int. Symp. On Comm. Cont. and Signal Proc*, pp. 1077-1081, 2008.
- [6] B. Jähne. "Spatio-temporal image processing," Springer, 1993.
- [7] D. Kalivas and A. Sawchuk, "Motion compensated enhancement of noisy image sequences," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Proc.*, pp. 2121-2124, 1990.
- [8] S. Lee and M. Kang, "Spatio-temporal video filtering algorithm based on 3-d anisotropic diffusion equation," in *Proc. Int. Conf. Image Processing*, vol. 2, pp. 447-450, 1998.
- [9] E. Bennett and L. McMillan, "Video enhancement using per-pixel virtual exposures," in *Proc. SIGGRAPH*, pp. 845-852, 2005.
- [10] Q. Sun, JA. Hossack, J. Tang and ST. Acton, "Speckle reducing anisotropic diffusion for 3D ultrasound images," in *Comput. Med. Imaging and Graphics*, vol. 28, no. 2, pp. 100-110, 2008.