

# 객체 기반 영상 분할을 위한 비등방성 확산 모델에 기반한 전처리 기법 연구

임재성

국립과학수사연구원 법공학부 디지털과

## Image pre-processing method based on non-isotropy diffusion model for object based image segmentation

Jae-Sung Lim

Department of Digital Analysis Division, National Forensic Service

**요약** 본 논문은 영상 분할 목적의 적응적 전처리 기법을 제안한다. 본 연구에서는 객체 단위의 군집화를 위해서 가능한 한 영역 간의 경계는 보존하면서 영역 내부는 평탄화 과정을 수행하여 군집화를 하는데 활용될 수 있는 적응적 전처리 기법의 설계를 목적으로 한다. 입력 영상이 주어졌을 때, 노이즈를 제거하고 영상 내 객체 정보와 세부적인 디테일들을 보존하려는 일반적인 디노이징 기법과는 다르게 제안하는 전처리 기법은 확산 모델을 기반으로 하여 영상 영역 간 경계 부 및 내부에 각각 확산 계수를 적응적으로 부여하며, 모든 영역에서의 획일화된 확산 과정을 수행하는 것이 아닌 영상 특성을 고려한 적응적 전처리를 수행한다. 따라서, 분할 단계 이전에 가능한 한 단순화 된 영상을 제공하는 것을 목적으로 하였다. 제안하는 전처리 기법의 영상 분할에 대한 효과를 평가하기 위해서, 본 연구에서는 PRI와 SC의 지표가 활용되었다. 실험 결과에서 제안된 전처리 기법을 적용한 영상 분할 결과에서 과분할이 억제된 결과와 함께 벤치마크 방법에 비해 PRI와 SC의 정량적인 지표값에서 최대 0.92%, 13.7%가 각각 개선된 것을 보여주었다.

**Abstract** This paper proposes an adaptive preprocessing method for image segmentation. This study aims to design a preprocessing method that performs a smoothing process intra-region and preserves the inter-region as much as possible for object-based clustering. When an input image is given, the proposed preprocessing method (which is different from the general denoising method that removes noise and preserves detailed information) based on the diffusion model adaptively assigns a diffusion coefficient between the intra and inter regions. The proposed preprocessing method considers image characteristics rather than performing a uniform diffusion process in all image regions. Therefore, the proposed method aims to provide a simplified image before the segmentation step. To evaluate the effect of the proposed preprocessing method on image segmentation, in this study, we utilized indicators such as probabilistic rand index (PRI) and segmentation covering (SC). The experimental results showed that the quantitative index values of PRI and SC were improved by up to 0.92% and 13.7%, respectively, compared to the benchmark method, with the result of suppressing over-segmentation in the image segmentation results by applying the proposed preprocessing method.

**Keywords** : Adaptive Diffusion Coefficient, Smoothing Intra Region, Preserving Inter Region, Image Characteristics, Object-Based Segmentation

---

본 논문은 국립과학수사연구원(2022-디지털-03) 연구과제로 수행되었음.

\*Corresponding Author : Jae-Sung Lim(National Forensic Service)

email: aceljs2000@naver.com

Received August 20, 2021

Revised September 24, 2021

Accepted January 7, 2022

Published January 31, 2022

## 1. 서론

영상 분할은 영상 내 장면 중에서 관심 있는 영역 혹은 의미 있는 영역으로 영상을 나누는 것을 말한다. 영상 분할 분야는 패턴 인식, 의료 영상 처리, 차량 자율 주행, 감시 시스템 등 다양한 영상 처리 분야에 활용 될 수 있으며, 또한, 위 분야 등에 접목하고자 하는 등 효율적인 영상 분할 기법의 고도화를 위해 수많은 연구들이 진행되고 있다. 하지만, 기존의 영상 분할 방법들은 공통적으로 처리하고자 하는 원 영상이 디지털 기기로부터 취득되는 과정에서 노이즈 성분에 의해 영상이 열화 되는 경우[1], 영상 획득 시 촬영 조건 예를 들면 저조도 환경에서 촬영되는 경우 영상 내 피사체의 계조가 부족할 가능성이 높아져 영상 내 객체 간 경계가 구분이 되지 않을 정도로 모호한 경우 등에서 영상 내 객체가 필요 이상으로 과분할 되는 문제를 겪고 있다. 따라서 위에서 언급된 과분할을 일으킬 수 있는 대표적 요소 중 하나인 노이즈 성분을 제거하고 영상 영역 간 선명도를 높이기 위해 디노이징 기법으로 개선시킬 수 있다. 먼저 다양한 기존 디노이징 방법이 존재하는데 이들 방법들이 동작하는 방식을 개략적으로 살펴보면 크게 두 가지 범주로 나누어 볼 수 있다. 첫 번째는 지역적 방법이며, 두 번째는 비지역적 방법이다. 지역적 방법[2-4]은 복원하려는 중심 픽셀을 기준으로 인접 픽셀 데이터들의 평균을 이용하는 방법이다. 반면에 비지역적 방법[5]은 복원하려는 픽셀과 비지역적인 즉 탐색윈도우 범위 내에서의 유사도에 따른 가중치를 활용하여 원본 데이터를 복원하는 방식이다. 위 두 가지 범주 중에 비지역적 방법은 탐색윈도우 크기에 따라서 복원하려는 부분의 가중치를 부여하는 과정에서 우수한 디노이징 결과를 제공하는 대신 탐색 윈도우 내에서 가중치를 구하는 과정에서 높은 계산비용이 요구되는 단점이 있다. 이 논문에서 객체 기반 영상 분할을 위해 디노이징 과정에서 비지역적 방식에서와는 달리 계산비용이 적으면서도, 주변 픽셀들의 평균의 방법으로 복원 픽셀을 결정하여 영역 간 경계가 모호해 지는 지역적 기법의 단점을 보완할 필요가 있다. 또한, 위 두 가지 방식의 디노이징 기법은 노이즈 제거와 영상 영역의 전체적인 세부 디테일 정보의 선명도를 높이는데 목적을 두고 있기 때문에 객체 기반의 영상 분할을 위해 기존 디노이징 방법은 영상 영역 내의 질감 정보를 제거하는 데 있어 부족한 측면이 존재한다.

따라서, 이 논문은 입력 영상 내의 같은 특성을 가지는 영역을 최대한 같은 군집으로 분할 결과를 도출하는

전처리 기법으로 적응적 확산 모델을 제안하고자 한다. 이를 위해서, 확산 모델[6] 내 영상의 평탄화 정도를 결정하는 파라미터  $K$ 값을 적응적으로 조정하여 앞서 언급된 기존의 지역적 방법에서의 영역 간 모호해지는 단점을 보완하고 영역 내의 컬러 영상 정보에서 질감 및 노이즈 성분을 억제하여 단순화된 영상으로 만들고자 한다. 또한, 전처리 과정 이후에는 단순화된 영상에 히스토그램 기반  $K$  군집화 알고리즘을 적용하여 객체형태의 분할된 결과를 생성하고자 한다.

이후의 본 논문은 다음과 같은 구성으로 진행된다. 2절 본론 중 2.1절에서는 제안하는 전처리 과정으로 적응적 확산 모델에 대한 내용을 기술하고, 분할 결과를 제공하는 히스토그램 기반  $K$  군집화 알고리즘에 대한 기법에 대한 내용을 간략하게 기술하고자 한다. 2.2절 실험결과에서는 영상 분할 결과를 타 벤치마크 방법과 비교하여 결과 영상과 함께 제시하고자 하며, 마지막으로 위 연구에 대한 내용을 결론부에서 정리하도록 하겠다.

## 2. 본론

### 2.1 제안하는 적응적 확산의 전처리 기법 및 영상 분할 과정

이번 절에서는 확산 모델에 대한 설명과 함께 제안하고자 하는 기법을 다루도록 하겠다. 먼저, 확산 방정식은 다음과 같이 정의된다.

$$\frac{dI}{dt} = (C \nabla I), I(0) = I_0, \quad (1)$$

$$C = g(\|\nabla I\|) = e^{-\left(\frac{\|\nabla I\|}{K_{Adaptive}}\right)^2}, \quad (2)$$

$$C = g(\|\nabla I\|) = \frac{1}{\left(1 + \left(\frac{\|\nabla I\|}{K_{Adaptive}}\right)^2\right)}, \quad (3)$$

여기에서  $C$ 는 확산 정도를 나타내는 확산계수,  $\|\nabla I\|$ 는 그래디언트의 크기 그리고  $g(\|\nabla I\|)$ 는 영상 내 경계부를 감지하는 함수로써 엣지 디텍터(edge-detector)로써의 역할을 수행한다. 이 논문에서는, 객체 영역의 내부(intra-region)에서는  $K$  값의 강도를 높게 부여하여 확산 계수를 큰 값으로 설정하도록 하며, 그에 반해 영상의 경계부(inter-region, edge)에서는  $K$  값의 강도를 낮게 부여함으로써 확산 계수를 작은 값으로 설정하도록 제안하는 방법에서 설계를 하도록 한다.

따라서 확산 과정은 확산 모델을 기반으로 하여 지역적인 영상 영역(a local image region)에서 중심 픽셀(또는 영역)과 주변 인접 픽셀(혹은 영역)의 그레이 레벨(gray level)의 차이에 의해 발생 되는 지역적 변화(local variation)를 제거하면서 영상의 경계부의 정보는 유지하는 역할을 수행한다. 따라서, 제안하는 방법은 모든 영상 영역에서 확립화 된 확산 과정을 수행하는 것이 아닌 영상 영역별로 평탄화 파라미터인 K 값을 영상의 특성을 고려하여 부여 되도록 함으로써 적응적인 영역 평탄화가 수행되도록 설계하였다. 위 언급된 내용은 다음 식에 의해 결정된다.

$$W_{Factor}(x, y) = \sum_{a=-WZN}^{WZN} \sum_{b=-WZN}^{WZN} \frac{P_B(x+a, y+b)}{(S)^2 \cdot (N)} \quad (4)$$

$$K_{Adaptive}(x, y) = AW_{Factor}(x, y), \quad (5)$$

예를 들어 중심 픽셀(또는 중심 영역)을  $(x, y)$ 라 할 때, 그 중심 픽셀 보다 인접한 여덟 개의 픽셀 간을 비교하여 중심 픽셀 보다 인접 픽셀의 값이 크면 1을 부여하고 그렇지 않으면 0을 부여하도록 하여 (4)식을 통해 가중치를 픽셀별로 결정하도록 한다.  $S$ 와  $N_{Color}$ 는  $3 \times 3$  윈도우 사이즈와 영상의 컬러 공간의 수를 각각 의미한다.

$WZN$ 는  $S$ 값의 절반 사이즈를 의미하며,  $A$ 는 노이즈 레벨에 따라 강도가 결정되는 계수이다. 따라서,  $K_{Adaptive}$ 는 위의 식 (5)에 의해서 결정된다.

윈도우의 중심픽셀을 기준으로 모든 방향으로 그래디언트의 차 정보를 이용하여 윈도우 내의 픽셀 영역에 대한 (6)번 식에 의해 확산 과정이 이루어지게 된다.

$$I_{t+1}(x, y) = I_t(x, y) + \lambda [C_N^t \partial_N + C_S^t \partial_S + C_W^t \partial_W + C_E^t \partial_E] I_t(x, y) + \frac{\lambda}{\sqrt{2}} [C_{NW}^t \partial_{NW} + C_{SE}^t \partial_{SE} + C_{NW}^t \partial_{NW} + C_{SW}^t \partial_{SW}] I_t(x, y) \quad (6)$$

여기에서  $N, S, W, E$ 는 중심픽셀을 기준으로 상, 하, 좌, 우 방향을 각각 의미한다.  $C_x$ 는 각 방향으로의 확산 계수를 의미하며,  $\partial_x$ 는 여덟 방향에 대한 일차 미분 값을 의미한다. 또한,  $t$ 는 반복 횟수를 의미한다.  $\lambda$ 와  $\lambda/\sqrt{2}$ 는 상, 하, 좌, 우의 네 방향과 위 네 방향에서 파생되는 대각 방향으로의 평탄화를 조절하는 결정하는 계수, 스텝사이즈가 된다.



Fig. 1. Original noisy-textured image(a), De-textured result using proposed preprocessing method(b)

Fig. 1에서 볼 수 있듯이, Fig. 1의 (a)에 비해 (b)의 결과 영상에서 질감 및 노이즈 성분이 제거된 것이 확인되며, 위 (1)~(6)식을 통해 적응적 확산 과정이 이루어지게 된다. 위의 결과를 바탕으로 분할 과정을 수행하게 되는데 앞선 전처리 과정이 K 군집화 알고리즘을 통한 분할 결과를 도출 시 과분할을 일으킬 수 있는 요소를 억제하는데 기여할 것으로 관찰된다. 위 전처리 과정에 이어 분할 결과가 도출되는 히스토그램 기반 K 군집화 방법은 양자화 된 히스토그램 빈상에서 최대-최소 알고리즘[7]을 통해 군집의 수를 적응적으로 정하게 된다. 위 군집의 수를 정함과 동시에 라벨링이 수행된다. 따라서, 계산비용이 적으면서도 정확한 영상 분할 결과를 획득할 수 있게 된다.

## 2.2 실험결과

위 2.1절의 제안하는 전처리 방법의 효용성을 검증하기 위해서, 실험에서는 버클리 분할 데이터 세트(BSDS)[8] 활용하였다. 또한, 제안하는 전처리 기법을 검증하기 위해서 두 가지 벤치마크 방법을 활용하여 비교한다. 그 벤치마크 방법은 히스토그램 기반 영상 분할 기법인 HKMC[9]와 HTFCM[10]으로 정하여 비교해보고자 한다. 본 실험에서 분할 결과를 평가하기 위해 제안하는 전처리 과정을 적용한 HKMC와 전처리 과정이 적용되지 않은 위 두 가지 벤치마크 방법 간에 분할 결과를 도출하며, 이를 정량적으로 평가하기 위해 확률적 랜드 인덱스(PRI)와 분할 커버링(SC)[11]을 활용하여 정량평가를 수행하고자 한다.

PRI는 다음의 식의 형태로 정의 된다.

$$PRI(S, G_k) = \frac{1}{T} \sum_{i < j} [c_{ij} p_{ij} + (1 - c_{ij})(1 - p_{ij})], \quad (7)$$

여기에서  $S$ 는 분할 결과이며  $\{G_k\}$ 는 실험에 활용되는 데이터 세트 중 영상 내에 판독 가능한 객체의 라벨링이 표시된 데이터 세트(Ground Truth, GT)를 의미한다.  $c_{ij}$ 는 픽셀  $i, j$ 가 같은 라벨을 갖은 사건을 말하며,  $p_{ij}$ 는  $c_{ij}$ 가 발생할 확률을 나타낸다. 즉 PRI는 입력 영상에서 가능한 모든 케이스의 픽셀 페어를 만들었을 때 이들 픽셀 페어 간에 동일 분할 영역의 여부가  $G_k$ 의 데이터와 알고리즘을 통해 도출된 분할 결과와 얼마나 일치 하는 것인지를 보여주는 지표라 할 수 있다.

다음으로 SC는 아래의 식으로 표현 될 수 있다.

$$o(R, R') = \frac{|R \cap R'|}{|R \cup R'|},$$

$$C(S' - S) = \frac{1}{N} \sum_{R \in S} |R| \max_{R' \in S'} o(R, R'),$$

(8)

여기에서  $S$ 는 GT에서 제공되는 분할 결과이고,  $S'$ 은 제안된 방법에서 도출된 분할 결과를 의미한다.  $R$ 은  $S$ 내의 분할 영역,  $R'$ 은  $S'$ 내의 분할 영역,  $N$ 은 전체 픽셀 개수를 나타낸다. 즉, GT에서 하나의 분할 영역을 기준으로 분할 결과의 영상 영역 중 둘의 분할 영역의 합에 대비하여 공통된 영역이 가장 큰 분할 영역을 찾아내어, 합을 구하는 형태로 GT의 모든 영역에 대해 반복된 작업을 수행하여 결과를 도출한다.

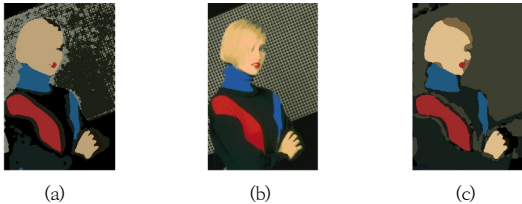


Fig. 2. Segmentation results: (a)HKMC, (b)HTFCM (c) Proposed method

Fig. 2에서 확인 할 수 있는 것처럼, Fig. 2(a)에서는 전경의 인물은 그 영역의 분할이 잘 이루어진데 비해 배경의 노이즈 성분과 질감성분의 영향으로 과분할 된 결과가 관찰되며, Fig. 2(b)도 배경 쪽의 질감의 세부적인 정보의 디테일 요소가 존재하여 과분할 된 것을 결과 영상을 통해 관찰할 수 있다. 또한, 그에 비해 Fig. 2(c)는 앞서 두 개의 영상에서 노이즈와 질감 성분으로 발생하게 된 불필요한 분할 결과들이 제안하는 적응적 확산 기법의 전처리 과정을 통해 제거됨으로써 전경과 배경 단위로 균집화가 이루어진 것을 결과 영상을 확인 할 수 있다. 뿐만 아니라, BSDS를 대상으로 PRI와 SC 지표 측면

에서도 전처리 과정이 포함되지 않은 벤치마크 방법에 비해서 Table 1에서 볼 수 있듯이 높은 수치를 보여주고 있다.

Table 1. Average PRIs and SCs for the benchmark and proposed method on BSDS

BSDS(500)	HKMC	HTFCM	Proposed method
Avg.PRI	0.756	0.758	0.763
Avg.SC	0.460	0.415	0.472

### 3. 결론

본 연구에서, 적응적 확산 기법을 활용한 영상 분할 목적의 적응적 전처리 방법을 제안하였다. 제안한 전처리 기법은 영상 내의 객체 정보를 구성하는 영역간의 경계인 구조적인 정보를 보존하면서 노이즈 성분을 제거할 뿐만 아니라 객체 기반의 영상 분할을 위해서 앞서 서론 부에서 소개한 기존의 디노이징 기법에서 부족했던 측면인 영역 내의 질감 성분을 제안하는 전처리 기법을 통해 제거함으로써 영역 단위의 단순화된 영상을 제공할 수 있었다. 또한, 제안하는 전처리 방법은 노이즈 레벨에 따라서 안정성에 관여하는 스텝 사이즈( $\rho$ )의 파라미터를 변화시켜가며 적용하지 않아도 되는 장점이 있을 뿐만 아니라 영상의 경계와 내부를 고려하여 적응적으로 평탄화 파라미터( $K$ )를 부여하기 때문에 영상의 경계를 보존하면서 영상 내 질감정보를 제거 할 수 있었다. 실험결과 측면에서도 제안된 전처리 기법이 적용된 영상 분할 결과에서 평균 PRI와 SC의 정량적인 지표면에서 개선되는 것을 확인할 수 있었으며, 영상 분할 결과에서 영역 간 구조적 정보는 유지하되, 과분할을 일으키는 요소인 노이즈 성분과 질감 정보를 제거하여, 전경과 배경에 해당되는 객체 영역 형태의 분할 결과로 구분되는 점 등으로 보아 제안된 전처리 기법의 효과성을 확인할 수 있었다.

### References

[1] Tsin.Y., "Statistical calibration of CCD imaging process," Computer Vision, 2001. ICCV 2001. Proceedings.  
DOI: <https://doi.org/10.1109/ICCV.2001.937555>

[2] M. Lindenbaum, M. Fishcher, and A. Bruckstein, "On

- gabor contribution to image enhancement,” Pattern Recognition, vol. 27, no. 1, pp. 1-8, Jan. 1994.  
DOI: [https://doi.org/10.1016/0031-3203\(94\)90013-2](https://doi.org/10.1016/0031-3203(94)90013-2)
- [3] L. Alvarez, P. L. Lions, and J. M. Morel, “Image selective smoothing and edge detection by nonlinear diffusion,” SIAM Journal on Numerical Analysis, vol.29, no. 1, pp.182-193, Feb. 1992.  
DOI: <https://doi.org/10.1137/0729012>
- [4] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” The Sixth International Conference Computer Vision, pp. 839-846, 1998.  
DOI: <https://doi.org/10.1109/ICCV.1998.710815>
- [5] A. Buades, B. Coll, J. M. Morel, “A non-local algorithm for image denoising,” IEEE Computer Vision and Pattern Recognition 2005 (CVPR 2005), vol. 2, pp. 60-65, Jun. 2005.  
DOI: <https://doi.org/10.1109/CVPR.2005.38>
- [6] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion.” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.12, no. 7, pp.629-639, July 1990.  
DOI: <https://doi.org/10.1109/34.56205>
- [7] V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, “Still image segmentation tools for content-based multimedia applications,” International Journal of pattern recognition and artificial intelligence, vol. 18, no. 4, pp. 701-725, Jun. 2004.  
DOI: <https://doi.org/10.1142/S0218001404003393>
- [8] C. C. Fowlkes, D. R. Martin, and J. Malik, “Local figure-ground cues are valid for natural images,” Journal of Vision, vol.7, no.8, pp. 1-9, Jun. 2007.  
DOI: <https://doi.org/10.1167/7.8.2>
- [9] T. W. Chen, Y. L. Chen, and S. Y. Chien, “Fast Image Segmentation Based on K-Means Clustering with Histogram in HSV Color Space,” Multimedia Signal Processing, pp. 322-325, Oct. 2008.  
DOI: <https://doi.org/10.1109/MMSP.2008.4665097>
- [10] K. S. Tan, and N. A. A. M. Isa, “Color image segmentation using histogram thresholding - Fuzzy C-means hybrid approach,” Pattern Recognition, vol. 44, no.1, pp.1-15, Jan. 2011.  
DOI: <https://doi.org/10.1016/i.patcog.2010.07.013>
- [11] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour Detection and Hierarchical Image Segmentation.” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 5, pp.898-916, May. 2011.  
DOI: <https://doi.org/10.1109/TPAMI.2010.161>

임 재 성(Jae-Sung Lim)

[정회원]



• 2018년 11월 ~ 현재 : 국립과학수사연구원 법공학부 디지털과

<관심분야>

영상처리, 영상분석, 3D 모델링