

컬러 영상 처리를 위한 Mean Shift 기법 개선

황영철* · 배정호* · 차의영*

*부산대학교

Modified Mean Shift for Color Image Processing

Young-chul Hwang* · Jung-ho Bae* · Eui-young Cha*

*Pusan National University

E-mail : sofe@pusan.ac.kr

요 약

본 논문에서는 개선된 mean shift를 이용한 컬러 영상 분할을 소개한다. Mean shift는 Yizong Cheng에 의해 재조명되고 Dorin Comaniciu 등에 의해 정리되어 영상 필터링(image filtering), 영상 분할(image segmentation), 물체 추적(object tracking) 등 여러 응용 분야에 널리 활용되고 있다. 커널을 이용해 밀도를 추정하고 밀도가 가장 높은 점으로 커널을 연속적으로 이동함으로써 지역적으로 주요한 위치로 데이터 값을 갱신시킨다. 그러나 영상에 포함된 모든 화소에 대해 mean shift를 수행해야하기 때문에 연산 시간이 많이 소요되는 단점이 있다. 본 논문에서는 mean shift 필터링 과정을 분석하고 참조수렴방법과 강제수렴방법을 이용해 소요 시간을 단축시켰다. 모든 점에 대해 mean shift를 수행하는 대신 특정 조건을 만족하는 픽셀은 이웃 픽셀의 수렴 값을 참조하고, mean shift 과정에 진동 또는 미미한 이동을 계속하는 픽셀은 강제 수렴을 실시하였다. 개선된 방법과 기존의 mean shift 방식을 적용하여 영상 필터링과 영상 분할에 적용한 실험에서 결과 영상에는 차이가 적고 기존의 방법에 비해 수행 시간이 24% 정도 소요됨을 확인하였다.

키워드

mean shift, mean shift filtering, image segmentation, nonparametric clustering, fast mean shift

1. 서 론

Mean shift를 이용한 데이터 분석은 1975년 Fukunaga와 Hostetler에 의해 제안되었지만[1] 주목을 받지 못하다가 1995년 Y. Cheng에 의해 재조명되고[2] 연구됨으로써 영상처리 분야에 응용되기 시작했다. 그 후 D. Comaniciu, P. Meer이 2002년 mean shift에 대해 자세하게 분석하고 관련 명제들을 증명함으로써 영상 처리, 데이터 분석 등 여러 응용 분야에 이용될 수 있는 기초를 마련하였고[3] 근래에는 객체 추적, 영상 필터링, 영상 분할, 클러스터링 및 스테레오 비전 등에 이용되고 있다. 특히 컬러 영상 처리 분야에서, 기존에 흑백 영상에 사용되던 방법을 단순히 차원을 확장한 개념으로 컬러 영상 처리에 사용하는 방법들의 단점을 벗어나 컬러 영상에 적합한 새로운 방법을 제시하면서 우수한 영상 처리 결과를 보여준다[3].

Mean shift를 이용한 영상 처리는 공간영역

(spatial domain)과 특징영역(range domain)이 결합된 결합영역(joint domain)에서 커널을 이용해 지역적으로 밀도가 가장 높은 점을 찾고 그 위치로 윈도우의 중심을 옮기는 과정을 반복한다. 윈도우가 한 점에 수렴하게 되면 반복을 멈추고 다음 데이터 위치로 이동해 같은 과정을 반복하게 된다.

Mean shift를 이용한 영상 분석은 좋은 성능을 보이지만 계산량이 많아 연산 시간이 많이 소요된다는 단점이 있다. 본 논문에서는 이런 단점을 개선할 수 있는 수정된 mean shift 알고리즘을 소개한다. 먼저 기존의 mean shift 알고리즘을 소개하고 연산 시간이 많이 소요되는 단계를 분석한다. 그리고 연산 시간을 단축시키는 새로운 방법을 소개하고 다양한 이미지에 대해 제안된 방법을 적용한 영상 필터링과 영상 분할을 실험함으로써 제안된 방법의 성능을 평가한다.

II. 본 론

2.1 Mean shift 알고리즘

Mean shift 알고리즘은 윈도우에 포함된 표본의 평균(mean)을 구하고 표본의 평균이 최대가 되는 위치로 윈도우의 중심을 이동(shift)하는 과정을 반복한다. 윈도우의 중심이 지역적으로 평균의 최대 위치(mode)에 도달하게 되면 이동을 멈춘다.

D-차원 특징 공간상의 임의의 데이터 집합 $\{x_i | i=1,2,\dots,n\}$ 에 대해, 커널 $K(\cdot)$ 와 반경 h 의 윈도우를 이용한 다변량 커널 밀도 추정 함수는 다음과 같다.

$$\hat{f}_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (1)$$

여기서 $k(\cdot)$ 는 커널 K 의 프로파일이고 MISE를 보장하는 커널은 Epanechnikov 커널로 알려져 있다[4].

밀도의 변화를 나타내는 프로파일 $g(x) = -K'(x)$ 가 존재할 때, 커널 $G(x)$ 는 커널 $K(x)$ 의 섀도우(Shadow) 커널이라 한다. 밀도의 모드를 찾는 것은 평균 이동(mean shift) 벡터의 수렴점을 찾는 과정이고 아래 식과 같다.

$$m_{h,G}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (2)$$

x 는 현재의 커널 중심이고, x_i 는 윈도우 범위 안의 데이터이다. 그러므로 $(j+1)$ 시간에 커널의 중심 y_{j+1} 은 식 (3)과 같다.

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}, j=1,2,\dots \quad (3)$$

즉, mean shift 알고리즘은 커널 $G(\cdot)$ 를 사용해서 윈도우 내의 데이터들에 대해 커널 $G(\cdot)$ 를 가중치로 하는 가중치 평균을 구하고 커널의 중심과 새로운 중심의 차가 0이 되는 수렴점을 찾는 과정이다.

Mean shift 알고리즘을 영상에 적용하는 방법은 다음과 같다. 먼저 공간영역과 특징영역을 결합한 결합영역 x_i 를 생성한다. 보통 공간영역은 x-y 2차원이고 특징영역은 색상을 표현하는 3차

원이므로 결합영역은 5차원 데이터가 된다. x_i 와 z_i 를 d-차원의 결합영역의 원본 영상 데이터와 필터링된 영상 데이터라고 하자. 영상 내 모든 화소에 대해 아래의 과정을 수행한다.

1. $j = 1, y_{i,1} = x_i$ 로 초기화
2. 식 (3)을 이용해 수렴할 때까지 $y_{i,j+1}$ 구하고 $y = y_{i,c}$ 로 설정
3. $z_i = (x_i^s, y_{i,c}^r)$ 로 설정

위 첨자 s 와 r 은 각각 공간(spatial)과 특징(range)를 나타내고 3번 과정은 위치 s 에서 필터링된 데이터가 y_c^r 에서 수렴함을 의미한다.

2.2 Mean shift 알고리즘 분석

Mean shift 알고리즘을 분석하면 수행시간 t_{MSF} 는 아래와 같이 나타낼 수 있다.

$$t_{MSF} = \alpha \times W \times H \times I \times T \quad (4)$$

여기서 W 와 H 는 각각 영상의 가로, 세로 크기이고 I 는 수렴하기 까지 커널을 이동하는 평균수이다. T 는 윈도우를 탐색해서 밀도를 추정하는 시간이다. Mean shift 필터링은 영상에 포함된 모든 화소에 대해 윈도우의 중심을 이동하면서 윈도우에 속하는 화소들의 밀도를 구하는 과정이므로 식 (4)와 같이 표현할 수 있다.

밀도를 추정하는 시간은 윈도우의 크기에 비례하기 때문에 윈도우의 크기를 줄이면 수행 시간을 단축할 수 있다. 그러나 윈도우의 크기를 줄이면 영상 필터링과 영상 분할 시에 영상이 과분할되어 적절한 결과를 보장할 수 없기 때문에 윈도우의 크기를 임의적으로 설정할 수 없다.

W, H 와 관련된 영상의 크기는 입력 영상에 따라 결정된다. 영상이 커질수록 W 와 H 의 곱이 증가하기 때문에 mean shift를 수행하는 시간도 함께 증가하게 된다. 수행 시간 단축을 위해 실제 영상의 크기를 축소할 수는 없지만 mean shift를 수행하는 화소의 수를 줄임으로써 전체 수행 시간을 단축시킬 수 있다. 본 논문에서는 이런 목적으로 참조수렴방법을 제안한다.

I 는 영상에 따라서 차이가 나고 실험적으로 특징 값이 유사한 화소들이 많이 분포한 영상에서 커널 이동 횟수가 증가하는 특징이 있다. 특히, 지역적으로 특징 값이 유사한 화소들이 밀집되어 있을 경우 커널이 아주 천천히 이동하거나 진동하는 등의 경우가 발생해 전체 수행 시간이 증가하게 된다. 본 논문에서는 커널 중심 값(mean)의 변화를 분석해서 진동 또는 미미한 움직임을 반복할 경우 강제로 수렴시키는 방법을 제안한다.

2.3 개선된 mean shift 알고리즘

우선 참조수렴방식을 살펴보면 영상에서 mean shift를 수행하는 화소 순서가 왼쪽에서 오른쪽, 위쪽에서 아래쪽 방향이다. 현재 화소에 대해 mean shift 과정을 수행하기 전에 윈도우에서 2사분면을 검색한다. 2사분면에 존재하는 화소들 중 현재 화소와의 거리가 윈도우 반경 이내인 참조후보 화소들을 찾는다. 참조후보 화소들 중 현재 화소와의 거리가 최소인 화소의 mean shift 수렴값을 현재 화소의 수렴값으로 대체한다.

$$\begin{aligned} \hat{s} &= \{x_i \mid \|x - x_i\| < h, x_i \in S\} \\ x_{ref} &= \operatorname{argmin}_i \|x - x_i\|, x_i \in \hat{s} \\ z_i &= z_{x_{ref}} \end{aligned} \quad (5)$$

식 (5)에서 \hat{s} 은 참조후보 집합이고, h 는 윈도우 반경이다. 공간영역의 윈도우 반경과 특징영역의 윈도우 반경이 서로 다를 수 있지만 공간영역과 특징영역을 결합영역으로 결합할 때 정규화 과정을 거치므로 상수 h 만으로 공간영역과 특징영역의 거리를 모두 표현할 수 있다. 영상에서 화소 값들이 급격히 변하기보다는 비슷한 색상들이 이웃하므로 참조수렴방법을 적용할 경우 mean shift를 수행해야하는 화소수를 크게 줄일 수 있고 결과적으로 수행 시간이 단축된다.

강제수렴방법은 mean shift vector 값의 변화를 살펴보고 값의 변화($\Delta m_{h,G}(\cdot)$)가 기준값(ϵ) 이하이면 mean shift 과정을 중단하고 현재 중심값을 수렴값으로 설정한다.

$$\begin{aligned} \text{if } |\Delta m_{h,G}(i)| < \epsilon, \text{ then } z_i &= (x_i^s, y_i^r) \\ \text{else } y(i) = y(i-1) + \Delta m_{h,G}(i) \end{aligned} \quad (6)$$

강제수렴방법을 이용할 경우 mean shift 과정에서 윈도우의 평균 이동 횟수가 감소해 수행 시간이 짧아진다.

2.4 영상 필터링과 영상 분할

개선된 mean shift 알고리즘을 이용한 영상 필터링 방법은 기본적인 mean shift 필터링과 유사하다. 먼저 결합공간영역 x_i 를 생성한다. 다음 단계에서 바로 mean shift 알고리즘을 수행하지 않고 참조수렴이 가능한지 검사한다. 윈도우의 2사분면에 속하는 x_j 들에 대해 참조가능 여부를 검사하고 참조 가능한 화소가 있으면 참조수렴방식으로 수렴값을 설정한다. 참조수렴방식의 조건을 만족하지 않는 화소일 경우 강제수렴방법을 이용해 mean shift 알고리즘을 수행한다.

영상 분할은 먼저 영상 필터링 과정을 수행한다. 필터링 완료 후 필터링된 화소에 대해 이웃한

z_i, z_j 의 거리가 윈도우 반경 보다 작으면 같은 세그먼트로 할당한다. 마지막으로, M보다 작은 세그먼트의 경우 선택적으로 이웃한 세그먼트로 병합해 미소 영역을 제거할 수 있다.

III. 실험 결과 및 분석

개선된 mean shift 알고리즘의 성능을 평가하기 위해 여러 가지 크기의 다양한 영상을 이용해 기본적인 mean shift 방법과의 수행 시간을 비교한다. 실험을 수행한 시스템의 환경은 Core2 CPU 1.80GHz에 2.0GB RAM, Windows XP이다.

먼저 강제수렴방법의 성능을 살펴보기 위해 기본적인 mean shift 방법과 제안한 방법을 이용한 영상 필터링에서 영상에 따른 윈도우의 평균 이동 횟수를 [표 1]에 나타내었다.

표 1. 알고리즘에 따른 윈도우의 평균 이동 횟수 (단위:번)

영상 크기	기본 MS	강제수렴
256x192	6.1	3.1
256x256	8.1	3.8
448x358	4.5	3.1

영상의 크기와 윈도우의 평균 이동 횟수는 관련이 없다. 기본적인 mean shift 방법과 비교해 볼 때 제안한 강제수렴방법의 윈도우 이동 횟수가 많이 줄었음을 확인할 수 있다.

기본적인 방법으로 mean shift 필터링 과정에서 수렴점을 찾아가는 과정(mean shift vector)을 [표 2]에 나타내었다. 실험을 CIE L*u*v* 색공간에서 실행하였기 때문에 특징값은 L,u,v 값을 나타낸다. 표에서 알 수 있듯이 윈도우의 중심이 아주 조금씩 이동함을 알 수 있다. 이와 같은 이동을 반복할 경우 수행 시간이 많이 걸리므로 강제수렴방법을 적용한다.

표 2 mean shift vector

순서	특징값	순서	특징값
1	43.574, 28.097, 40.097	6	43.456, 21.722, 35.529
2	43.940, 24.302, 37.255	7	43.413, 21.718, 35.478
3	43.593, 22.476, 36.266	8	43.319, 21.713, 35.286
4	43.510, 22.062, 35.921	9	43.325, 21.775, 35.218
5	43.494, 21.817, 35.702	10	43.325, 21.775, 35.218

강제수렴방법을 적용하면 윈도우의 이동 과정이 [표 2]의 4 번째 과정에서 수렴하게 된다. 강제 수렴할 경우의 수렴값과 기본 방법에 의한 수렴값과의 차이는 작다. 즉, 필터링 된 결과 차이

는 크지 않으면서 수행 시간은 크게 단축할 수 있음을 알 수 있다.

참조수렴방법을 적용할 경우, 영상의 크기에 따라 mean shift 과정을 수행하는 픽셀 수를 기본 mean shift 방법과 비교한 결과를 [표 3]에 나타내었다. 기본방법을 사용할 경우 영상의 모든 픽셀이 mean shift를 수행하므로 영상의 크기와 mean shift를 수행하는 픽셀수가 같고 D/R은 기본방법 대비 참조수렴방법을 나타낸다.

표 3. 알고리즘에 따른 mean shift를 수행하는 픽셀 수

영상크기	기본방법	참조수렴방법	D/R
557x384	213888	64303	30.1%
448x358	160384	43574	27.2%
481x321	154401	41217	26.7%
256x256	65536	18854	28.8%
256x192	49152	16033	32.6%

영상의 크기와 영상의 내용에 따라 차이가 있지만 참조수렴방법은 기본방법의 약 30%의 픽셀만이 mean shift 과정을 수행하였다. BSD 이미지 300장을 포함한 많은 수의 영상을 이용한 실험 결과 평균적으로 영상 전체 픽셀 중 29.06%의 픽셀이 mean shift 과정을 수행하였고 그에 따라 기본방법에 비해 70% 정도의 수행 시간 감소 효과를 나타내었다.

[표 4]에는 전체적으로 기본적인 mean shift 방법과 본 논문에서 제안한 방법의 속도차를 필터링에 소요된 시간으로 비교한다. 실험에 사용된 사진 중 대표 크기의 사진에 대한 결과를 나타내고 있다. [그림 1]은 기본방법과 제안한 방법을 이용한 필터링 결과 영상이다. 그림에서 알 수 있듯이 기본적인 방법과 제안한 방법의 결과 영상이 큰 차이가 나지 않는다. 반면, 실험 결과 제안한 방법은 평균적으로 기본방법의 mean shift를 이용한 필터링의 23.7%의 시간이 소요되어 수행 시간이 크게 단축됨을 확인할 수 있었다.

표 4. 알고리즘에 따른 필터링 시간 (단위:초)

영상크기	기본방법	참조수렴방법	비율
557x384	6.66	1.33	20.0%
448x358	2.84	0.91	32.0%
481x321	3.89	0.94	24.2%
256x256	1.95	0.42	21.5%
256x192	1.50	0.31	20.7%

IV. 결론

본 논문에서는 mean shift 알고리즘의 수행 시간을 단축시킬 수 있는 방법을 제안하였다. 참조수렴방법을 이용하여 mean shift 과정을 수행하는 픽셀 수를 줄이고, 윈도우의 이동이 진동하거나 또는 미미한 경우 강제수렴방법을 이용하여 윈도우의 이동을 강제 종료하였다. 개선된 mean shift 방법을 다양하고 많은 수의 영상을 대상으로, 필터링과 분할 실험을 수행한 결과 개선된 mean shift 알고리즘은 결과 영상의 차이는 크지 않고 소요 시간은 기본 방법에 비해 23.7%의 시간이 소요되어 수행시간이 크게 단축되었다.



그림 1 mean shift 필터링 결과 영상. 위에서부터 원본영상, 기본방법, 제안한방법

참고문헌

- [1] K. Fukunaga and L. Hostetler, "The Estimation of the Gradient of a Density Function with Applications in Pattern Recognition", IEEE Tran. IT., Vol. 21, pp 32-40, 1975.
- [2] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering", IEEE Trans. PAMI, Vol. 17, pp 790-799, 1995.
- [3] D. Comaniciu and P. Meer, "Mean Shift : Robust Approach toward Feature Space Analysis", IEEE Trans. PAMI., Vol. 24, pp 603-619, 2002.
- [4] B.W. Silverman, *Density Estimation for statistics and Data Analysis*, New York : Chapman and Hall, 1986.