

2D 텐서 보팅에 기반 한 손상된 텍스트 영상의 복원 및 분할

박 종 현^{*} · Nguyen Dinh Toan^{**} · 이 귀 상^{***}

요 약

본 논문에서는 잡음에 의해 손상된 텍스트 영상으로부터 복원 및 분할을 위한 새로운 접근 방법을 제안한다. 제안된 방법은 손상된 영역의 복원을 위하여 색상 및 비색상 성분을 2차 대칭 스틱 텐서로 표현하고 보팅 기반의 손상된 영역을 복원하였으며, 마지막으로 클러스터링 방법에 의해 분할을 수행한다. 먼저 우리는 제안된 색상 선택합수에 의해 잡음에 강건한 색상과 비색상 성분을 선택한다. 두 번째 단계에서는 각각의 선택된 특징 벡터들은 스틱 텐서로 표현하였으며 제한된 보팅 커널의 필드내에서 이웃하는 보티들과 통신을 통하여 새롭게 정의된다. 따라서 2차 보팅 후 각각의 스틱 텐서는 이웃하는 텐서와 같은 특성을 가지며 손상된 영역들을 복원할 수 있다. 마지막으로 복원된 영상의 성능을 평가하기 위하여 적응적 평균 이동 알고리즘과 클러스터링 알고리즘을 이용하여 영상 분할을 수행하였다. 실험에서 제안된 방법은 전체적인 처리과정을 자동적으로 수행 가능하였으며 배경 및 객체의 영역에서 효율적인 복원 및 분할을 수행할 수 있었다.

키워드 : 스틱 텐서, 보팅, 복원, 텍스트 영상 분할

Corrupted Region Restoration based on 2D Tensor Voting and Segmentation

Park, Jonghyun[†] · Nguyen Dinh Toan^{**} · Lee, Gueesang^{***}

ABSTRACT

A new approach is proposed for restoration of corrupted regions and segmentation in natural text images. The challenge is to fill in the corrupted regions on the basis of color feature analysis by second order symmetric stick tensor. It is show how feature analysis can benefit from analyzing features using tensor voting with chromatic and achromatic components. The proposed method is applied to text images corrupted by manifold types of various noises. Firstly, we decompose an image into chromatic and achromatic components to analyze images. Secondly, selected feature vectors are analyzed by second-order symmetric stick tensor. And tensors are redefined by voting information with neighbor voters, while restore the corrupted regions. Lastly, mode estimation and segmentation are performed by adaptive mean shift and separated clustering method respectively. This approach is automatically done, thereby allowing to easily fill-in corrupted regions containing completely different structures and surrounding backgrounds. Applications of proposed method include the restoration of damaged text images; removal of superimposed noises or streaks. We so can see that proposed approach is efficient and robust in terms of restoring and segmenting text images corrupted.

Key Words : Stick Tensor, Voting, Restoration, Text Image Segmentation

1. 서 론

정보통신 및 멀티미디어 그리고 소프트웨어와 하드웨어의 향상으로 영상처리 기술에 대한 수요는 과거의 어느 때보다 증가하고 있는 추세이다. 특히 기술의 발달로 다양한 영상 및 비디오 데이터를 획득할 수 있으며 이러한 데이터들은

전처리 과정 후 다양한 분야에 응용되어 지고 있다. 따라서 영상 및 비디오 처리를 기반으로 한 컴퓨터비전의 응용 분야는 매우 광범위하며 다양한 접근 방법을 통하여 관심 영역을 시스템화 하기 위하여 오랜 기간 동안 연구되어 왔다. 영상에 포함된 다양한 객체들은 관심 대상으로 패턴인식 및 컴퓨터비전 시스템에서 중요한 응용 분야로 전처리 과정에서 정확하고 신뢰할 수 있는 영역들로 분할할 수 있어야 한다. 그러나 영상에 포함된 객체들은 다양한 잡음에 의해 손상 가능하며 정확한 처리 과정 없는 결과는 객체 본래의 의미를 정확하게 전달하는데 어려움이 있어 컴퓨터비전 시스템의 성능 저하를 제공할 수 있다. 특히 인간의 시각 시스템의 성능 저하를 제공할 수 있다. 특히 인간의 시각 시스템의 성능 저하를 제공할 수 있다.

^{*} 정 회 원 : 전남대학교 전자컴퓨터공학부 BK21 연구교수
^{**} 준 회 원 : 전남대학교 전자컴퓨터공학부 박사과정
^{***} 중신회원 : 전남대학교 전자컴퓨터공학부 교수
논문접수 : 2007년 11월 26일
수정일 : 2008년 1월 24일
심사완료 : 2008년 2월 11일

템은 다양한 잡음들에 의해 손상된 모양, 질감 그리고 색상 정보를 추론할 수 있는 능력을 가지고 있다. 그러한 인지 능력은 다양한 시스템들에 적용하고 있으며 객체의 국부적인 특징들로부터 전체적인 패턴을 자동적으로 추론하고자 하는 인지 시스템에 관심이 있다.

일반적으로 자연 영상은 PDA(personal digital assistant), 모바일 폰, 또는 비전 시스템들에 의해 쉽게 획득할 수 있다. 그리고 자연 영상이 가지고 있는 다양한 객체 정보들 중에서 텍스트 인식은 폭넓은 분야에서 응용성을 가지고 있다. 텍스트 정보의 인식 과정은 색상 분석, 잡음제거 및 복원, 텍스트 분할, 그리고 인식으로 분류할 수 있다. 자연 영상에서의 텍스트 정보는 중요한 의미를 가지고 있기 때문에 다양한 시스템 구현을 위하여 손상된 영역을 추론하여 복원하는 과정이 전처리 부분에서 수행되어야만 한다.

기존의 연구에서는 잡음에 의해 손상된 영상의 복원을 위해 기본적으로 denoising 방법이 적용되어지고 있다. Denoising 기술은 일반적인 영상처리 방법으로 영상에서의 잡음을 제거하여 원래의 신호를 잘 보존하는 효율적인 방법이다. 이러한 기술은 임펄스성 잡음이나 백색 잡음의 제거에 매우 효율적인 특징을 가지고 있으나 streaks, graffiti 등과 같은 영향으로부터 발생하는 손상 영역 및 잡음은 제거하는데 많은 어려움을 가지고 있다[1][2]. 또한 최근에는 사진이나 그림에서 손상된 영역이 잡음에 비해 매우 큰 영역으로 나타난 경우를 복원하기 위한 영상 인페인팅(image inpainting) 기술이 제안되었다[3][4]. 영상 인페인팅 기술은 손상된 사진 및 그림에서 손상된 부분을 주위의 질감 및 색상 정보를 이용하여 복원을 수행할 수 있으며, 선택되어진 영역의 제거 및 교체를 가능케 한다. 그러나 영상 인페인팅 기술은 복원하고자 하는 영역이 사용자에 의해 선택 되어야하며, 반복적인 처리 과정에 의해 많은 처리 시간이 요구된다. 따라서 텍스트 분할 문제점에 초점을 두어, 사용자가 사진에 어떠한 정의 없이 텍스트의 손상을 취급할 수 있는 방법을 개발하는데 목적이 있다.

본 논문에서 우리는 손상된 영상으로부터 영상의 객체가 가지고 있는 본래의 패턴을 추론하며 기존의 알고리즘에서 취급하기 힘든 잡음의 형태뿐만 아니라 강하게 손상된 영역까지 제어할 수 있는 새로운 접근 방법을 제안한다. 제안된 방법을 위하여 각각의 입력 데이터(또는 토큰(Token))들은 고유값(eigen value)과 고유벡터(eigen vector)를 가지고 2D 대칭 텐서(tensor)로 표현하였으며 주위의 토큰들과 보팅(voting)을 통하여 손상된 텍스트 영역의 복원을 수행한다. 텐서 보팅은 G.G. Medioni 등에 의해 제안되었으며 컴퓨터 비전을 위한 객체의 패턴 추론을 위하여 다양한 응용분야에 효율적으로 적용되고 있다[5][6].

제안된 텍스트 영상의 복원 및 분할 구현에서, 먼저 우리는 영상의 효율적인 분석 및 처리를 위하여 RGB 성분을 이용하여 HSI 공간에서 색상(hue)과 명도(intensity) 성분을 선택한다. 선택된 색상과 명도 성분을 이용하여 입력 데이터는 스틱 텐서로 표현하며, 제안된 보팅 영역에서 각각의 토큰은 이웃하는 보트들과 통신을 수행하여 새로운 텐서로

표현할 수 있다. 따라서 잡음 영역은 2차 보팅을 통하여 이웃하는 정보와 유사한 SS(stick saliency) 값들로 표현할 수 있으며 잡음 제거 및 정보의 교체가 가능하다. 또한 제안된 방법의 효율성 실험을 위해 클러스터링 알고리즘을 적용하여 성능의 비교 평가를 수행한다.

본 논문은 다음과 같은 순서로 구성되어 있다. 먼저 2장에서 텐서 보팅에 대한 전반적인 내용에 대하여 설명하였다. 3장에서는 2차 스틱 텐서의 분석을 이용하여 손상된 영역의 복원 방법에 대하여 기술한다. 또한 4장에서는 복원된 텍스트 영상의 성능 평가를 위한 자동적인 영상분할 방법을 설명하였다. 마지막으로 5장에서는 실험 결과와 성능에 대하여 평가하였으며, 6장에서 결론을 제시한다.

2. 2차원 공간에서의 스틱 텐서

텐서는 벡터의 개념을 확장한 기하학적인 양으로 균일한 원(circle) 또는 구체(sphere)에 한 방향의 어떠한 힘을 작용시키면 변형하여 타원 또는 타원체의 형태로 표현할 수 있다. 따라서 텐서를 구성하는 각각의 파라미터들은 입력 데이터의 특성을 나타내며 정해진 필드내에서 보팅을 통하여 인접하는 정보와 교류 및 통신을 통하여 정교한 텐서로 재표현할 수 있다. 일반적으로 텐서 보팅은 두 성분에 기반하고 있다. 하나는 데이터 표현을 위한 텐서 계산, 그리고 데이터들과 통신을 위한 보팅이다. 입력 데이터는 각각의 위치에서 이웃하는 데이터들에 자신의 정보를 전파(propagation)한다. 그 정보는 텐서로 인코딩하며 보팅 필드에 의해 패턴 구조가 결정되어 진다. 따라서 각각의 위치에서 수집되어진 정보는 보팅 필드내의 모든 보트들에 캐스트(cast) 하여 각각의 특징 형태를 위하여 텐서의 고유값(eigenvalue)을 이용한 saliency 맵을 만들 수 있다. 그리고 관심 대상이 되는 특징에 대하여 분석 및 추론을 수행할 수 있다[5][6].

2.1 2차 대칭 텐서의 표현

입력 데이터들은 토큰으로 표현할 수 있으며 각각의 토큰들은 2D 텐서로 계산 가능하다. 따라서 텐서로부터 기하학적인 특징 추출에 의한 패턴을 추론할 수 있으며 텐서들은 이웃하는 정보와 통신을 통하여 새로운 텐서로 재표현할 수 있어 두 포인트 사이의 가장 적합한 경로를 만들 수 있다. 모든 토큰들은 이웃하는 토큰들로부터 보트(vote)들을 축적하며 수집된 정보는 2D 대칭 텐서의 합에 의해 다음 식 (1)과 같이 공분산 행렬로써 표현할 수 있다.

$$cov = \begin{bmatrix} \Sigma v_x^2 & \Sigma v_x v_y \\ \Sigma v_y v_x & \Sigma v_y^2 \end{bmatrix} \quad (1)$$

여기서 $[v_x, v_y]$ 는 주어진 데이터를 위하여 이웃하는 데이터에 의해 발생된 스틱 보트의 벡터이다. 주어진 고유 시스템은 두 개의 고유값 (λ_1, λ_2) 와, 두 개의 고유벡터는 (\hat{e}_1, \hat{e}_2) 로 구성되어 있으며 행렬 cov 는 다음 (2) 식과 같이 표현할 수 있다.

$$cov = [\hat{e}_1, \hat{e}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \hat{e}_1^T \\ \hat{e}_2^T \end{bmatrix} \quad (2)$$

$$cov_{2,d} = (\lambda_1 - \lambda_2) \hat{e}_1 \hat{e}_1^T + \lambda_2 (\hat{e}_1 \hat{e}_1^T + \hat{e}_2 \hat{e}_2^T) \quad (3)$$

여기서, $\hat{e}_1 \hat{e}_1^T$ 와 $\hat{e}_1 \hat{e}_1^T + \hat{e}_2 \hat{e}_2^T$ 는 2D에서 스틱 (stick) 텐서와 볼(ball) 텐서를 나타내며, 각각의 saliency는 $\lambda_1 - \lambda_2$ 와 λ_2 이다.

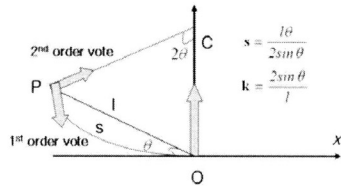
2.2 보팅 필드

보팅 필드는 2D 스틱 보팅 필드로부터 유도할 수 있으며 일반적으로 이것을 2D 스틱 커널이라 부른다. 보팅 필드에서 보트들은 접선 벡터나 법선 벡터에 의해 하나의 방향 성분으로 정의할 수 있다. 보트 필드내에서 보트들은 각각의 위치에서 접선 벡터 또는 법선 벡터에 의존한다. 그림 1에서, “공간에 주어진 O 와 P 를 지나는 곡선에 가장 부드러운 경로의 법선 벡터는 무엇인가?”에 대하여 생각해 볼 수 있다.

여기서 가상의 연속하는 호(arc)를 따라 일정한 곡선을 유지하기 때문에 접촉 서클에 연속하는 O 와 P 는 가장 적당한 것 중에 하나이다. 가장 적당한 법선은 점 P 에서 곡면에 있는 법선에 의해 주어진다. 따라서 스틱 보트의 방향은 O 와 P 를 가장 부드럽게 연결하는 법선이 된다. P 에서 법선 벡터의 길이는 보트의 saliency를 나타내는 호의 길이 OP 와 circular arc의 곡면에 반비례한다. 따라서 모든 보팅 크기는 기본적인 2D 커널에 기반하고 있으며 다음 식 (4)과 같이 나타낼 수 있다.

$$SDF(s, k, \sigma) = e^{-\left(\frac{s^2 + ck^2}{\sigma^2}\right)} \quad (4)$$

s 는 OP 의 호 길이이며, k 는 곡률, c 는 높은 곡률을 가지고 제어하기 위한 상수 값이다. 그리고 σ 는 효율적인 이웃 영역에 대한 크기로 분석을 위한 스케일 값을 나타낸다. 이 함수에 대한 자세한 내용은 [5]을 참조할 수 있다. 따라서 각각의 입력 위치에서 정규 방향의 벡터를 얻을 수 있다면, 우리는 2D 공간에서 높은 saliency 값을 가지고 있는 토큰들을 연결하기 위한 부드러운 구조를 추론할 수 있다. 특징 추출은 모든 불연속인 위치에서 보팅에 의해 수행된다.



(그림 1) 스틱 텐서에 의한 1차 및 2차 보팅 필드

3. 손상된 영역의 복원을 위한 스틱 텐서의 분석

3.1 색상 성분의 선택

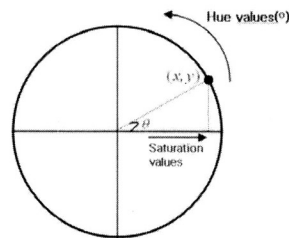
영상처리 과정에서 선택한 색상 모델 및 성분들은 잡음

및 조명의 변화에 매우 많은 영향을 받으며 결과에 중요한 영향을 제공한다. 따라서 보다 효율적인 색상 성분의 선택은 시스템의 성능 향상을 제공하며 신뢰할 수 있는 결과를 만들 수 있다. 일반적으로 색조(hue) 성분은 빛의 변화에 대해 강건한 특성을 가지고 있다. 만약 채도(saturation) 성분이 너무 낮은 경우나 명도가 작은 값이거나, 너무 높은 경우 채도 성분은 안정적이지 못하며 의미 없는 값으로 나타나는 것이 기본적인 이론이다. 이런 영역에서는 명도 성분이 영상의 객체를 구별하는데 유용한 정보가 된다. 따라서 우리는 영역 분할의 효율성 관점에서 색상 영역에 속하는지 또는 비색상 영역에 속하는지를 결정하는 선택 함수를 제안한다. Sural *et. al.*은 영상 화소의 색조나 명암 성분 중 어느 쪽이 더 우월한지를 채도의 특성에 의해 결정하여 칼라 영상의 분할을 수행하였다[12]. 그러나 채도에 대한 경계값은 조명에 불변한 값이 아니다. 또한 이 방법은 부정확한 경계값의 문제로 색상 영역과 비색상 영역으로 잘못 판단될 수도 있어 많은 고립 영역들이 발생할 수 있다. 제안된 선택 함수는 채도 값을 이용하지 않고 원래의 영상이 가지고 있는 R(red), G(green), B(blue) 상호간의 차를 이용하여 시각적 측면에서 우월한 성분을 선택한다. 차가 0에 가까울수록 완전한 무채색을 가진다는 특징을 이용하여 색상(chromatic) 성분과 비색상(achromatic) 성분을 구분하였다. 색상 성분의 결정을 위한 수식은 다음과 같다.

$$F(x, y) = \frac{|R(x, y) - G(x, y)| + |G(x, y) - B(x, y)| + |B(x, y) - R(x, y)|}{3} \quad (5)$$

만약 $F(x, y) > TH$ 면 색조 성분이 선택되고 그렇지 않으면 $F(x, y)$ 는 명암 성분으로 분류한다. 본 연구에서 TH 의 경계값 설정을 위하여 실험에 사용한 자연 영상들을 분석하였으며, 가장 적당한 실험치 값으로 $TH=20$ 으로 하였다. 그림 2는 HSI 공간에서 H와 S 성분의 특징을 그림으로 보여주고 있다. (x, y) 의 위치값은 θ 의 크기를 가지고 있는 색상값을 나타낸다.

여기서 색상 θ 값은 $0^\circ \sim 360^\circ$ 의 범위 값을 가진다. 그러나 색상 성분이 가지는 값의 특성상 0° 와 180° 사이의 색조 성분들의 각도에 의한 방향이 180° 와 360° 사이의 색상 성분들의



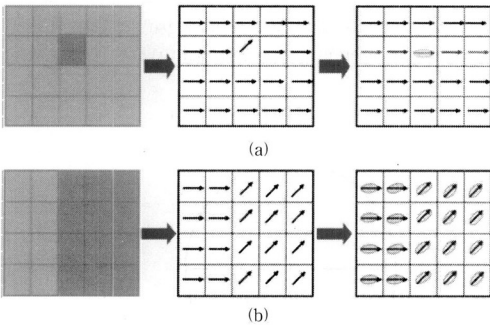
(그림 2) Hue와 Saturation에 의한 표현

방향과 동일하게 나타난다. 본 연구에서 우리는 실험을 위하여 $\theta = \theta/2$ 에 의해 색상 성분에 대한 중복의 문제를 해결하였다. 따라서 θ 는 $0^\circ \leq \theta \leq 180^\circ$ 범위의 값을 가진다. 결과적으로 색상 성분과 명암 성분은 0.0~1.0 값으로 정규화 하였으며 선택된 데이터들은 텐서 보팅 단계에서 입력 토큰으로 사용할 수 있다.

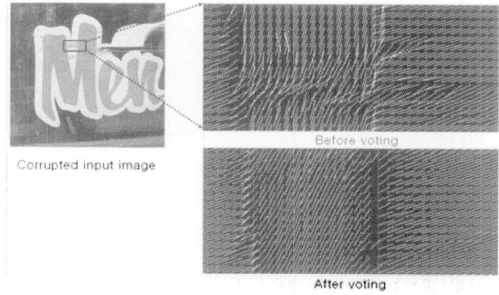
3.2 텐서 보팅에 의한 손상 영역 복원

입력 토큰은 식 (2)처럼 텐서로 표현 가능하며 주어진 보팅 공간내에서 이웃하는 정보들과 통신을 통하여 정보를 교환하여 새롭게 정의할 수 있다. 따라서 영상내 존재하는 잡음 및 손상된 영역을 효율적으로 제거할 수 있으며 이웃하는 영역의 패턴과 유사하게 복원할 수 있다.

(그림 3)은 텐서 보팅에 의한 잡음 제거 및 영역의 특징 보존을 보여주기 위한 가상 데이터 (synthetic data)로 입력 데이터에 대한 보팅 후 스틱 텐서의 방향에 대하여 보여주고 있다. (그림 3)에서 (a)는 첫 번째 R 영역에 잡음 B를 포함한 경우이다. (그림 3(a))의 가운데 영상은 입력 데이터의 특성을 스틱 텐서로 보여주고 있으며 스틱 텐서의 방향이 잡음 위치에서 다른 주변과 상이하게 나타난다. 따라서 스틱 텐서들은 제한된 보팅 필드에서 이웃하는 보터들과 정보 교환을 통하여 이웃하는 정보의 saliency 성분에 상응하는 스틱 텐서로 재표현할 수 있으며, 보팅 후 결과를 오른쪽에서 보여주고 있다. (그림 3)의 (b) 영상은 두 개의 영역을 가지고 있는 경우를 스틱 텐서의 특징(중간 영상)과 결과(오른쪽 영상)를 보여주고 있다. 두 영역의 경계부분에서 원래의 텐서 성분과 정확하게 일치하지는 않지만 각각의 색상 영역에서 유사한 패턴으로 묘사되어 특별한 변화 없이 각각의 색상 영역을 잘 보존하고 있음을 보여주고 있다. (그림 4)는 실제 영상이 어떤 물리적인 영향에 의해 손상된 경우 스틱 텐서를 기반한 보팅 전과 후를 왼쪽 그림에서 보여주고 있다. 왼쪽의 손상된 원래의 텍스트 영상에서 표시된 박스 내의 잡음 영역에 대하여 주위의 성분과 상응하는 텐서들로 표현됨을 확인할 수 있다.



(그림 3) Synthetic 색상 성분을 가지고 있는 데이터에서 텐서 보팅 결과(R:red 성분, B:blue 성분): (a) R 영역에 B 잡음을 포함하는 경우, (b) R과 B 두 개의 영역을 가지고 있는 경우.



(그림 4) 손상된 실제 영상의 스틱 텐서의 보팅 결과

4. 분리된 K-means 알고리즘을 이용한 영상 분할

손상된 영상에서 잡음 제거 및 복원을 수행한 후 영상은 알려지지 않은 분리 가능한 균질 영역들로 구성되게 된다. 이들 영역을 분할하기 위해 분리된 K-means 알고리즘을 사용한다. 일반적으로 K-means 알고리즘은 수동적으로 선택한 초기 값들에 의해 정확한 결과를 얻기에 많은 어려움이 있다. 그러므로 본 논문에서는 평균 이동에 기반한 모드 검출(mean shift-based mode detection) 알고리즘을 이용하여 클러스터의 수를 추정하고 K-means 알고리즘에서 초기 값으로 사용한다. 본 연구에서는 [8][9]에서 제안한 평균 이동 기반의 모드 검출 방법을 적용하였으며 효율적으로 객체의 모드를 검출할 수 있었다. 따라서 선택된 모드들의 개수나 중심점들은 K-means 알고리즘의 초기 값으로 사용할 수 있다. 제안된 분할 알고리즘은 두개의 다른 K-means 알고리즘을 수행해야 한다. 전처리 과정에서 생성된 색상 성분에서 명도 성분은 선형이지만 색상 성분은 비선형적인 구조를 가지기 때문이다. 따라서 먼저 intensity 값에 대해서는 주어진 초기값을 이용하여 분할을 수행한다. 다음으로는, [0.6-1.0]으로 정규화한 hue 값에 대한 분할을 수행하며 비선형적인 값은 [7]에서 제안한 접근 방법을 사용하여 비선형 계산을 수행하였다. 다음 알고리즘 I는 제안한 텍스트 영상 분할에 대하여 보여주고 있다.

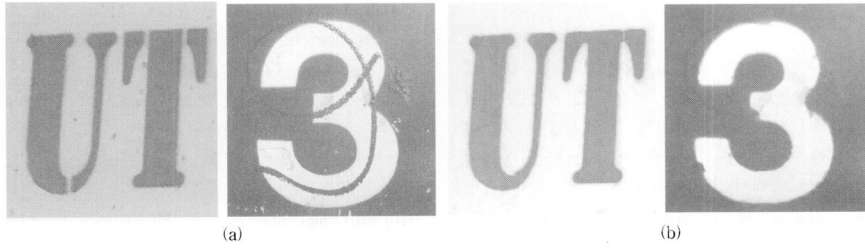
ALGORITHM I: 제안된 영역 분할 알고리즘

STEP 1: 입력->Improved Image Data 값($I(x,y)$);
(클러스터의 수와 클러스터의 중앙값은 mean-shift 알고리즘을 이용하여 추정함)

STEP 2: CLUSTERING USING K-MEANS ALGORITHM
IF $I(x,y) > T$, then Clustering hue components by K-means algorithm considering the nonlinear property;

Else Clustering intensity components by K-means clustering with Euclidean distance;

STEP 3: 출력 -> Segmented Image

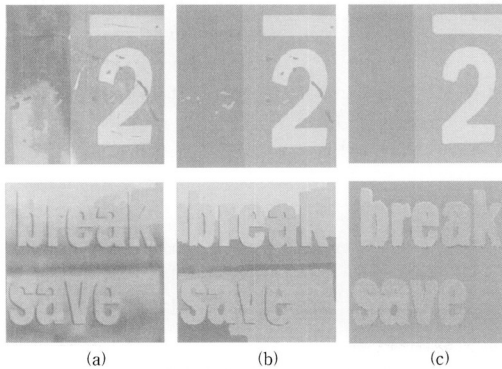


(그림 5) 제안된 방법에 의한 손상된 영역의 복원 결과:(a) 손상된 텍스트 영상, (b) 복원된 결과 영상.

5. 실험 및 결과

제안된 방법의 성능 평가를 위하여 256×256 크기의 텍스트가 포함된 칼라 자연 영상을 사용하였다. 실험에 사용된 영상은 ICDAR 영상 DB [10][11]와 실제 디지털 카메라를 이용하여 획득한 것을 사용하였다. 실험을 위하여 본 연구에서는 획득된 영상으로부터 텍스트가 포함된 영역을 임의로 검출한 후 제안한 알고리즘을 적용하였다. 또한 손상된 영역의 효율적인 분할과 성능 평가를 위하여 기존의 클러스터링 알고리즘을 적용하여 비교 분석을 수행하였다. 본 연구에서 실험 결과는 텍스트 영상에 기반하고 있으므로 일반적인 색상 및 질감 정보의 복원 보다 텍스트 획의 패턴 복원을 보여주고자 한다. (그림 5)는 제안된 스틱 텐서에 의한 복원 결과를 보여주고 있다. 일반적으로 텍스트 영상에서 텍스트는 (a)의 왼쪽 그림처럼 인공적인 방법으로 글자의 모양에 변형을 주거나, (a)의 오른쪽 그림처럼 다양한 형태의 물리적인 영향에 의해 손상될 수 있다. 제안된 방법에 의해 텍스트의 끊김 및 손상 영역을 정확한 복원 결과를 보여주고 있다.

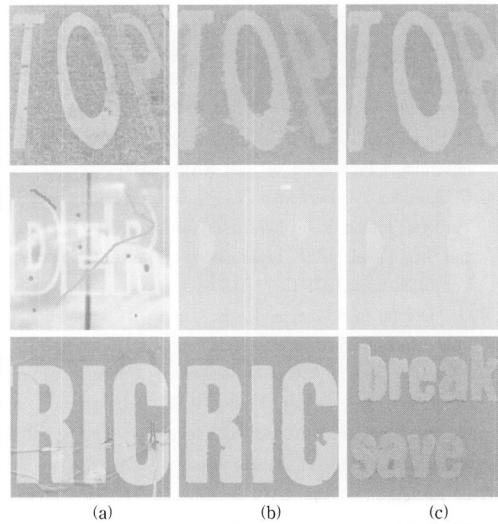
(그림 6)은 임의로 그래픽 툴을 이용해 손상된 텍스트 영상과 명암의 변화를 포함하고 있는 텍스트 영상을 보여주고 있다. 일반적으로 실외에서 획득한 영상은 다양한 조명의 변화로 복잡한 구조를 제공할 수 있다. 따라서 조명의 영향은 (그림 6)에서와 같이 복잡한 배경을 만들 수 있으며 복잡한 결과



(그림 6) 조명의 반사에 의해 복잡한 배경을 포함하고 있는 텍스트 영상: (a) 입력 영상, (b) EDISON 툴을 이용한 분할 결과, (c) 제안한 방법의 결과

를 제공한다. (그림 6(b))는 [9]에 의해 제안된 EDISON 툴을 이용하여 분할한 결과를 보여주고 있다. 분할 결과 배경 영역에서 텍스트 영역보다 많은 영역들로 분할이 됨을 확인할 수 있다. (그림 6(c))는 제안된 방법에 의한 결과로 텍스트 영역과 배경 영역을 보다 더 단순하게 분할하였음을 보여주고 있다. 그리고 그림 7은 다양한 형태의 손상 영역을 가지고 있는 텍스트에 대한 결과를 보여주고 있다.

(그림 7)의 (b)는 1st 보팅 후 결과로 텍스트 영역과 배경 영역의 경계부분에서 많은 고립 영역들이 불규칙하게 나타남을 볼 수 있다. 그 이유는 배경의 색상 성분에서 나타나는 색상 성분이 불규칙한 hue 성분을 포함하고 있기 때문이다. 따라서 이러한 불규칙 잡음과 같은 영역을 제거하기 위하여 2nd 보팅을 수행할 수 있다. (그림 6(c))에 두 번의 보팅이 끝난 후 결과 영상을 보여주고 있으며, 2nd 보팅을 통하여 효율적으로 잡음의 제거 및 손상된 영역의 복원이 가능함을 볼 수 있다. 잡음에 의해 손상된 많은 텍스트 영상의 실험 결과의 비교 분석을 통하여 제안된 방법에 의한 손상된 영역의 복원 및 잡음 제거의 효율성을 확인할 수 있으며 분할 결과에서도 보다 더 좋은 결과를 확인할 수 있다.



(그림 7) 반복적인 보팅에 의한 손상된 텍스트 영역의 복원 및 분할 결과: (a) 입력 영상, (b) 1st 보팅 후 결과, (c) 2nd 보팅 후 결과 영상

6. 결 론

본 논문에서는 2D 텐서의 특성을 이용한 잡음 제거 및 손상된 영역의 복원을 위한 새로운 접근 방법을 제안하였다. 주어진 칼라 영상은 HSI 공간에서 각도 값을 이용하여 스틱 텐서로 표현 가능하며 보팅 공간에서 보팅을 통하여 이웃 텐서와 상응하는 텐서로 재표현할 수 있다. 제안된 방법은 3단계로 구성되었으며, 첫 번째 단계에서는 RGB 색상 성분으로부터 색상 성분과 비색상 성분을 분리하였다. 그리고 두 번째 단계에서는 분리된 색상 성분을 텐서로 표현하였으며 보팅을 통하여 잡음 제거와 손상된 영역을 복원 하였다. 마지막으로 복원된 영상의 성능평가를 위해 분리된 K-means 알고리즘을 적용하여 영상을 분할하였다. 기존의 방법과 달리 제안된 방법은 특별한 파라미터 및 사전 정보 없이 스틱 텐서의 보팅을 통하여 다양한 잡음들을 제거할 수 있었으며 객체가 가지는 원래의 패턴을 복원할 수 있었다.

그러나 앞으로의 연구에서는 더 복잡하고 다양한 형태의 잡음에 대하여 많은 연구가 수행되어야 할 것이다. 그러나 알고리즘의 한계로 인한 복잡한 영상에서의 손상된 영역의 복원에 어려움을 가지고 있으며, 영상내 작은 폰트의 글자를 가지고 있는 경우 보팅 스케일 σ 값에 의해 작은 글자들이 잡음으로 고려되는 경우가 발생하였다. 따라서 차후 연구에서는 더 복잡하고 다양한 형태의 글자를 포함하고 있는 텍스트 영상에서 좀 더 많은 연구가 수행 되어야 할 것이다.

참 고 문 헌

[1] H. Kaiqu, W. Zhenyang and W. Qiao, "Image Enhancement based on the Statistics of Visual Representation," *Image and Vision Computing*, Vol.23, pp.51-57, 2005.

[2] M. Motwani, M. Gadiya, R. Motwani and F.C. Harris, "A Survey of Image Denoising Techniques," *International Signal Processing Conference of GSPx*, 2004.

[3] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, "Image Inpainting," *Proceeding of SIGGRAPH*, 2000.

[4] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro and J. Verdera, "Filling-in by Joint Interpolation of Vector Fields and Gray Level," *IEEE Transactions on Image Processing*, Vol.10, No.8, pp.1200-1211, 2001.

[5] G. Medioni, M.S. Lee, and C.K. Tang, *A Computational Framework for Segmentation and Grouping*, Elsevier, 2000.

[6] J. Jia, C.K. Tang, "Inference of Segmented Color and Texture Description by Tensor Voting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.6, pp.771-786, 2004.

[7] C. Zhang, P. Wang, "A New Method of Color Image Segmentation based on Intensity and Hue Clustering," *IEEE International Conference on Pattern Recognition*, Vol.3, pp.3617-3621, 2000.

[8] D. Comaniciu and P. Meer, "Mean Shift Analysis and Application," *IEEE International Conference on Computer Vision*, pp.1197-1203, 1999.

[9] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol.24, No.5, pp.603-619, 2002.

[10] <http://algoval.essex.ac.uk/icdar/Competitions.html>

[11] S.M. Lucas, A. Panaretos, L.Sosa, A. Tang, S. Wong and R. Young, "ICDAR 2003 Robust Reading Competitions," *IEEE International Conference on Document Analysis and Recognition*, pp.682-687, 2003.

[12] S. Sural, G. Qian, S. Pramanik, "Segmentation and Histogram Generation using the HSV Color Space for Image Retrieval," *IEEE International Conference on Image Processing*, Vol.2, pp.589-592, 2000.



박 종 현

e-mail : jhpark@chonnam.ac.kr
 1994년 목포대학교 전자공학과 학사
 1997년 목포대학교 전자공학과 석사
 2001년 목포대학교 전자공학과 박사
 2002년~2004년 전북대학교 Post-Doc.
 2004년~2006년 University of Southern California, USA, Post-Doc.

2006년~현 재 전남대학교 BK21 연구교수
 관심분야 : 영상처리, 패턴인식, 컴퓨터비전, 의료영상처리



Nguyen Dinh Toan

e-mail : toan@chonnam.ac.kr
 2005년 Received the B.S. degree in Information Technology from Vietnam National University Ho Chi Minh City, University of Natural Science

현 재 M.S. degree in computer science at Chonnam National University in Korea

관심분야 : Video coding, Ubiquitous Computing



이 귀 상

e-mail : gslee@chonnam.ac.kr
 1980년 서울대학교 전기공학과 학사
 1982년 서울대학교 전기계산기공학과 석사
 1991년 Pennsylvania 주립대학 전산학 박사
 1984년~현 재 전남대학교 전자컴퓨터 공학과 교수

관심분야 : 멀티미디어통신, 영상처리 및 컴퓨터비전, 임베디드 시스템