

Coiflet Wavelet과 LoG 연산자를 이용한 자연이미지에서의 텍스트 검출 알고리즘

*신 성, 백 영 현, 문 성 룡, 신 홍 규
원광대학교 전자공학과
e-mail : nova929@hotmail.com

Text Extraction Algorithm in Natural Image using LoG Operator and Coiflet Wavelet

*Seong Shin, Young-hyun Baek, Sung-ryong Moon, Hong-kyu Shin
Department of Electronic Engineering
Wonkwang University

Abstract

This paper is to be pre-processing that decides the text recognizability and quality contained in natural image. Differentiated with the existing studies, In this paper, it suggests the application of partially unified color models, Coiflet Wavelet and text extraction algorithm that uses the closed curve edge features of LoG (laplacian of gaussian)operator. The text image included in natural image such as signboard has the same hue, saturation and value, and there is a certain thickness as for their feature. Each color element is restructured into closed area by LoG operator, the 2nd differential operator. The text area is contracted by Hough Transform, logical AND-OR operator of each color model and Minimum-Distance classifier. This paper targets natural image into which text area is added regardless of the size and resolution of the image, and it is confirmed to have more excellent performance than other algorithms with many restrictions.

I. 서론

최근 과학 기술이 발전함에 따라 고도의 영상처리

기술이 현실화, 상용화 되어가고 있는 실정이며, 로봇이나 휴먼 시스템 등의 자동화 기술에 적용되어 인간 생활에 도움을 주고자 하는 노력들이 활발히 진행되고 있다. 디지털 카메라나 캠코더로 취득한 영상에 포함되어 있는 텍스트 정보는 그 영상에서 중요한 정보들을 내포하고 있으며, 이 텍스트 정보의 응용은 위의 친인간적인 자동화 기술의 범위를 늘리는데 기여할 수 있을 것이다. 따라서 최근 들어 이에 대한 많은 연구가 이루어지고 있는 실정이나, 사진과 같은 복잡한 이미지 내에 자연스럽게 포함된 텍스트 인식은 배경과의 구별 기준이 모호하며, 여러 종류의 다양한 잡음이 포함되어 있어 인식에 상당한 어려움을 겪고 있다.

따라서 자연 이미지에 내포된 텍스트의 인식 성능을 좌우하며 질을 결정하는 전처리 과정에 대한 연구가 절실히 필요하며, 기존 연구로는 밝기 변화를 이용한 텍스트 영역 검출 방법, 명도 이미지의 수직 수평분산을 이용한 방법, Differential top-hat 연산을 이용한 방법, 컬러의 연속성을 이용한 방법, RGB 3차원 공간의 색 변화 값을 이용한 방법 등이 있다.[3,4,5]

본 논문은 기존연구와 차별화된 방법으로 Coiflet Wavelet과 부분 통합적인 컬러모델의 적용 및 LoG 연산자의 폐곡선 에지 특징 이용한 텍스트 검출 알고리즘을 제안하고자 한다. 본 제안은 영상의 크기나 해상도, 기울어짐 등에 상관없이 자연스럽게 텍스트 영역이 첨가된 자연 이미지를 대상으로 하며, 동일 영상에 대하여 기존의 텍스트 추출 알고리즘과 비교함으로써 제안 알고리즘의 우수성을 확인하고자 한다.

II. 부분통합적인 컬러모델과 웨이블릿의 적용

2.1 공간 컬러 모델

공간 컬러 모델은 일반적으로 받아들여지는 표준적인 방법으로, 좌표 시스템의 하나이며 각각의 색을 한 점으로 나타내는 좌표시스템의 부분 집합이다. 컬러모델의 종류로는 RGB, CMY, YCbCr, YPbPr, HSV, HMMD등이 있으며, 본 논문은 RGB, HSV, YCbCr 컬러모델을 사용하여 구성한다. [1]

RGB컬러모델은 빨강, 초록, 파랑의 기본 스펙트럼 성분들로 나타나며, 디지털 영상을 얻기 위해 사용되는 대부분의 컬러 카메라들이 RGB 포맷을 사용한다.

HSV컬러모델은 인간의 시각시스템과 가장 흡사한 컬러모델로 색상(Hue), 채도(Saturation), 명도(Value)의 세 가지 특성들이 컬러를 설명하는데 사용되며, YCbCr컬러모델은 컬러 정보로부터 광도를 분리하는 또 하나의 컬러공간으로 광도는 Y(휘도)로 기호화 되고 푸른 정보와 붉은 정보는 Cb와 Cr로 기호화 된다.[2][7]

2.1 부분통합적인 컬러모델

자연이미지의 텍스트 영역은 대부분 주변과의 차별을 위해 두드러진 특징 (배경과의 임계차가 큰 동일한 색상과 채도, 명도)을 가지고 있다. 본 논문은 이에 착안해서 공간컬러모델 중 RGB컬러모델로부터 HSV컬러모델의 S (채도, Saturation), V (명도, Value) 값과 YCbCr컬러모델의 Cb(푸른 정보), Cr(붉은 정보)를 분리하여 사용한다. HSV컬러모델의 색상(H)정보는 색평면의 각도에 의해 나타나는 값으로 사람의 눈에 비슷한 계열의 색일지라도 색평면에서는 미세한 명암 값에도 민감하게 반응하게 되어 텍스트 영역에서 색상의 동일성 문제 적용시 문제가 생기게 된다. 따라서 텍스트 영역의 색상의 동일성 문제는 색상정보의 변화 값이 덜 민감하게 반응하는 YCbCr컬러모델의 CbCr값을 사용한다.

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$V = \frac{1}{3} (R+G+B) \tag{1}$$

$$C_b = -0.16874R - 0.33126G + 0.50000B$$

$$C_r = 0.50000R - 0.41869G - 0.08131B$$

2.2 웨이블릿 변환

웨이블릿 변환은 푸리에 변환과 같이 기저 함수들의 집합으로 신호를 분해하는 하나의 방법이나 푸리에 변환이 모든 주파수 대역에 대하여 동일한 크기의 필터 윈도우를 사용하는 반면 웨이블릿 변환은 고주파 대역에서는 폭이 좁은 윈도우를, 저주파 대역에서는 폭이 넓은 윈도우를 사용한다. 따라서 웨이블릿 해석은 constant relative bandwidth analysis이라고도 일컬어지며, 주파수 대역의 변화 폭은 항상 주파수 값에 비례한다. 또한 웨이블릿 변환은 입력 신호를 특정 기저 함수의 집합으로 분리하는 과정으로도 이해할 수 있으며 웨이블릿 변환에 사용되는 기저 함수의 집합은 하나의 모 웨이블릿 기저 함수(mother wavelet basis function)에 대한 시간축 방향으로의 확대 및 축소 그리고 평행 이동을 통해 얻어진다. 즉 기본 웨이블릿 기저 함수는 특별한 형태의 band-pass필터로 생각할 수 있으며, 웨이블릿 변환의 상대 대역폭 불변성은 기본 웨이블릿 기저에 대한 시간축 방향 축소 및 확대에 의해 충족되어진다. 이에 따라 웨이블릿 변환에서는 주파수 대역이라는 용어 대신 스케일(scale)이라는 용어를 주로 사용하며, 다른 말로 원신호의 시간, 스케일 공간표현이라 일컫는다. 웨이블릿의 가장 일반적인 원형 웨이블릿은 다음과 같다.[6]

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \tag{2}$$

여기서 a는 웨이블릿의 크기를 결정하는 압축계수이며, b는 시간 축으로의 이동에 관계되는 전이계수이다.

III. LoG연산자의 폐곡선 에지 특징

3.1 LoG 연산자를 이용한 에지검출

영상의 에지는 입력 영상에 대해 물체의 모양이나 크기, 텍스트 상태 등의 많은 정보들을 가지고 있으며 화소값이 낮은 값에서 높은 값으로 또는 높은 값에서 낮은 값으로 변하는 지점에 존재한다[1].

에지를 검출하기 위한 연산자로는 크게 1차 미분 연산자와 2차 미분 연산자로 나눌 수 있는데, 1차 미분 연산자로는 소벨, 로버츠, 프리윗 연산자 등이 있으며 2차 미분 연산자로는 라플라시안 연산자, LoG 연산자, 영교차 연산자 등이 있다. 각각의 연산자는 에지가 존재하는 영역에 컨볼루션을 취함으로써 인해서 에지를 검

출할 수 있는데 1차미분 연산자의 경우 연산자가 예지가 존재하는 영역을 지나면 지나치게 민감하게 반응하는 단점이 있어 이를 보완하기 위해 2차미분을 사용한다. 다음은 본 논문에서 사용하는 LoG(Laplacian of Gaussian)연산자의 함수이다.

$$LoG(x,y) = \frac{1}{\pi\sigma^4} \left[1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (3)$$

이 함수는 맥스칸 모자 필터라고도 불리며, 라플라시안을 적용하기 전에 가우시안 스무딩을 수행함으로써 잡음에 덜 민감하고, 2차미분 연산자의 장점인 추출된 에지 윤곽선이 폐곡선을 이루게 된다는 특징을 가지고 있다. [2]

3.2 허프(hough) 변환

허프변환은 1962년 직선위의 점들을 찾기 위한 방법의 대안으로 Hough가 제안한 방법으로 한 직선상의 점을 파라미터 평면으로 허프변환을 하면, 하나의 점으로 표현할 수 있다는 것이다. 따라서, 같은 직선상의 점들은 허프 변환한 파라미터 평면에서의 어떤 한 점에서 만나게 된다는 것으로 다음과 같은 기울기-절편과 각-거리 수식으로 표현된다. [2]

$$\begin{aligned} y_i &= ax_i + b \\ x\cos\theta + y\sin\theta &= \sigma \end{aligned} \quad (4)$$

IV. 제안된 텍스트 추출 알고리즘

본 논문에서는 먼저 계산량의 감소를 위해 원영상에서 웨이블렛 계수만을 추출하여 사용하였다. 기저함수는 직교 정규화된 특성을 가지며 피크점을 검출하거나 불연속선을 검출하는데 유리한 Coiflet 웨이블렛을 사용하였으며[9], 영상의 데이터양을 줄이면서도 폐곡선 에지 정보를 보존하기에 가장 알맞은 최적의 함수인 coiflets5 함수를 실험에 의해 결정하여 사용하였다.

자연이미지에 포함된 간판, 표지판 등의 텍스트 이미지는 동일한 색상, 채도, 명도를 가지며 특성상 일정수치의 두께가 존재한다. 따라서 색상, 채도, 명도에 해당하는 각 S,V,Cb,Cr값에 2차 미분 연산자인 LoG 연산자를 적용해서 에지를 검출함으로써 폐(Close)영역으로 재구성 한다.

$$LoG(x,y) = \frac{1}{\pi\sigma^4} \left[1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (5)$$

하프변환의 선추출에 의해 연결된 긴 선등의 잡음을 제거하고 모폴로지의 조건적팽창 연산을 적용해 영역을 채우게 된다.[6]

$$X_k = (X_{k-1} \oplus B) \cap A^C \quad k=1,2,3, \dots \quad (6)$$

식6은 제한 조건이 없으면 전 영역을 채우게 된다. 그러나 각 단계에서 A^C 과의 교집합은 결과를 고려 영역의 내부로 국한시키게 된다.

선추출은 다음단계인 모폴로지의 조건적팽창(conditional dilation) 적용시 테두리에 갇혀진 텍스트 영역이 함몰되는 것을 방지해주는 역할도 있다.

다음으로 텍스트만을 추출하기위해 각 컬러영역을 순차적 AND 연산을 수행한 후 OR연산을 수행해 텍스트 영상을 검출한다.

$$SVC_bCr = (S \cap V) \cup (S \cap C_b) \cup (S \cap C_r) \cup (V \cap C_b) \cup (V \cap C_r) \cup (C_b \cap C_r) \quad (7)$$

검출영상에는 크고 작은 잡음이 섞여 있는데 텍스트 영역만 검출하기 위하여 최종적으로 8방향 연결 성분을 구하고 구해진 연결성분들을 가지고 최소거리 식별법에 의해 텍스트 영상만을 검출하게 된다.

$$\begin{aligned} m_j &= \frac{1}{N_j} \sum_{x \in \omega_j} x \quad j=1,2,\dots,W \\ D_j(x) &= \|x - m_j\| \quad j=1,2,\dots,W \end{aligned} \quad (8)$$

m_j 는 평균벡터, N_j 는 클래스 ω_j 로부터의 학습패턴

다음은 제안된 검출 알고리즘의 블록다이어그램이다.

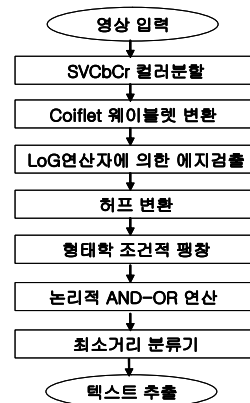


그림1. 제안 알고리즘 블록다이어그램

V. 모의 실험

본 논문에서 제안한 텍스트 추출 알고리즘은 영상의 크기나 해상도 등에 상관없이 자연스럽게 텍스트영역

이 첨가된 20개의 자연이미지를 대상으로 하였으며, Intel Celeron CPU 2.4Ghz, 512MB RAM 환경에서 Matlab을 이용하여 실험하였다. 본 알고리즘은 평균 82%의 검출률(추출된 글자 수/영상 내에서의 전체글자갯수)을 보임으로써 기존 알고리즘에[8] 비해서 검출률 면에서 우수하였으며, 계산량을 현저히 감소시킴으로써 속도면에서도 뛰어난 성능을 보임을 확인하였다. 아래 그림은 각각 원영상과 기존 알고리즘 및 제안된 알고리즘 적용 후의 결과 영상이다.

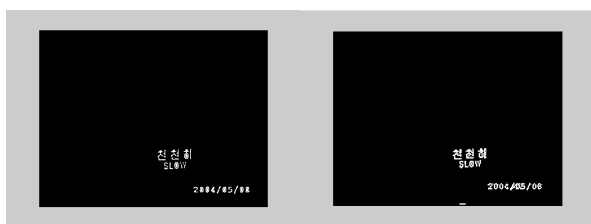
알고리즘	기존알고리즘	제안알고리즘
영상1	7/10	7/10
영상2	8/9	8/9
영상3	4/5	4/5
영상4	4/7	5/7
영상5	7/7	7/7
검출률	79.20%	82.06%

(*검출수/총 문자수)

표1. 기존알고리즘과 제안알고리즘 검출률 비교



그림2. 원 영상



(a)기존 알고리즘 (b) 제안 알고리즘

그림3. 기존알고리즘 및 제안알고리즘 적용 영상

VI. 결 론

본 논문은 자연이미지에 내포된 텍스트의 인식 성능을 좌우하며 질을 결정하는 전처리 과정에 관한 것으로, 기존 연구와 차별화된 방법으로 Coiflet 웨이블릿과 부분통합적인 컬러모델의 적용 및 LoG연산자의 폐곡선 에지 특징을 이용한 텍스트 추출 알고리즘을 제안

하였다. 본 알고리즘은 먼저 직교 정규화 특성을 갖는 Coiflet 웨이블릿 계수를 추출하여 데이터의 양을 감소시켰다. 또한 자연이미지에 포함된 간판, 표지판 등의 텍스트이미지는 동일한 색상, 채도, 명도를 가지며 특성상 일정수치의 두께가 존재한다는데 착안해서 각각의 컬러 성분을 2차 미분 연산자인 LoG 연산자에 의해 폐(Close)영역으로 재구성하고, 하프변환과 각 컬러 모델의 논리적 AND-OR연산 및 최소거리분류법에 의해 최종적으로 텍스트 영역이 추출하였다. 본 텍스트 검출 알고리즘은 기존의 알고리즘들과 달리 영상의 크기나 해상도에 상관없이 텍스트의 검출이 가능하며, 향후 실시간 구현시 텍스트 인식의 전처리 과정으로서 시각장애인을 위한 보행안내시스템, 지능 로봇 등의 자동 주행 등과 같은 다양한 분야에 활용이 사료된다.

참고문헌

- [1] R. Crane, "A simplified approach to Image Processing," Prentice-Hall, 1997.
- [2] R. Gonzalez and R. Woods, "Digital Image Processing," Addison-Wesley Publishing Company, 1992.
- [3] Yu Zhong, Kalle Karu, Anil K. Jain, "Locating Text In Complex Image," Pattern Recognition, Vol.28. No. 10, pp.1523-1535,1995.
- [4] Pyeoung-Kee Kim, "Automatic Text Location in Complex Color Image using Local Color Quantization," TENCON 99. Proceedings of the IEEE Region 10 Technical Conference, Vol.1, pp.629-632, 1999
- [5] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene images," IEEE Transactions Pattern Analysis and Machine Intelligence, PAMI-16(2), pp. 67-82, 1995.
- [6] 송창주, 박태근, "확장성 있는 데이터 재배열 블록을 이용한 효율적인 1차원 이산 웨이블릿 변환 필터 설계," 15회 신호처리합동학술대회논문집 제 15권, 제1호, pp 238, 2002
- [7] R. D. Dony, "Edge detection on color images using RGB vector angles," IEEE Trans. Canadian Conference on Electrical and Computer
- [8] 신성, 노종훈, 백영현, 문성룡, "LoG 연산자의 폐곡선 특징을 이용한 자연이미지에서 텍스트 추출 알고리즘" 대한전자공학회, 2005 하계학술대회 pp. 205-208
- [9] I. Daubechies, Orythonormal Bases of Compactly Supported Wavelets II, Variations on a Theme, SIAM J. of Math. Anal., 24(2), pp.499-519 (1993)