스크린 콘텐츠의 하위 영상 검출을 위한 AdaBoost의 활용

길종인, 김만배
강원대학교 컴퓨터정보통신공학과
e-mail:jigil@kangwon.ac.kr

Utilization of AdaBoost for Sub-image Detection in Screen Content

Jong-In Gil, Manbae Kim
Dept of Computer and Communications Engineering, Kangwon National University

요 약

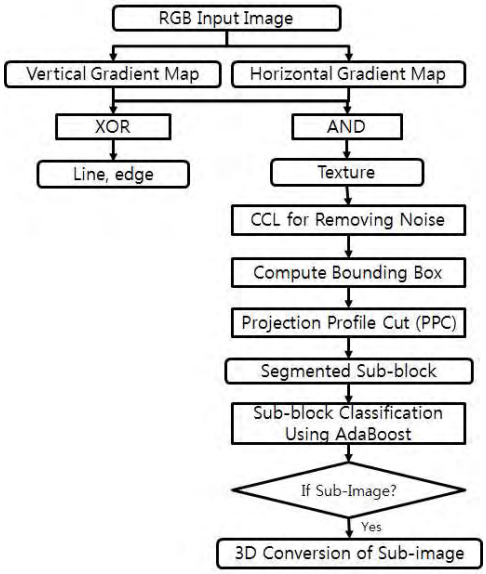
웹페이지와 같은 스크린 콘텐츠는 카메라로부터 획득할 수 있는 자연영상과 달리 텍스트, 로고, 아이콘 및 하위 영상과 같은 여러 가지 요소들을 포함하고 있고, 각 요소들은 서로 다른 형식의 정보를 사용자에게 전달한다. 본 논문에서는 윈도우 영상을 지역적인 특성에 따라 다수의 블록으로 분할한 후, 분할된 각 영역을 배경, 텍스트, 하위영상으로 분류하였고, 기계학습 기반의 알고리즘이 하위 영상 검출에도 좋은 접근법이 될 수 있음을 증명하기 위해 AdaBoost를 이용하였다. 실험결과로부터 93.4%의 검출률, 13%의 거짓 긍정률을 보임으로서, 제안방법이 효과적임을 입증하였다.

1. 서론

스크린 콘텐츠(Screen Content)는 컴퓨터에서 수행되는 응용 프로그램으로부터 생성된 정보를 사용자에게 전달하기 위해 여러 가지 형태로 생성 및 변환하여 출력하였을 때 나타나는 영상이다. 만일 사용자가 이러한 스크린 콘텐츠로부터 다른 정보 및 요소를 포함하고 있는 구역을 정확히 구별할 수 있다면, 각 구역의 목적에 맞도록 적절히 차별화된 영상처리를 적용할 수 있을 것이다. 이러한 윈도우 영상의 영역분할이 적용될 수 있는 분야로써 2D-to-3D 변환이 있다. 그러나 이러한 2D-to-3D 변환기술은 자연영상에 초점이 맞추어져 있기 때문에, 스크린 콘텐츠를 변환하였을 경우, 깊이왜곡의 원인이 된다. 이를 위해 가장 좋은 방법은 입체변환이 불필요한 부분은 변환하지 않는 것이다. 실제로 배경 및 텍스트를 강제로 변환하게 되면 많은 왜곡이 발생하게 되므로, 3D로 변환하지 않는 것이 좋다. 본 논문에서는 스크린 콘텐츠만이 가지는 고유한 특성을 이용하여 하위 영상을 검출하는 방법을 제안한다.

2. 하위 영역 분할 및 영상 분류

그림 1은 본 논문에서 제안하는 방법의 전체 흐름도를 보여준다. 스크린 콘텐츠 영상이 입력되면 수직/수평 그래디언트를 측정한다. 이때 수직/수평 그래디언트 영상에 XOR 연산을 수행하면 하위 영상의 경계가 검출되고, AND 연산을 수행하면 하위 영상 내부의 복잡한 텍스쳐가 검출된다. 본 연구에서는 하위 영상의 경계보다 텍스쳐를 활용함으로써 신뢰도를 높이고자 하였다.



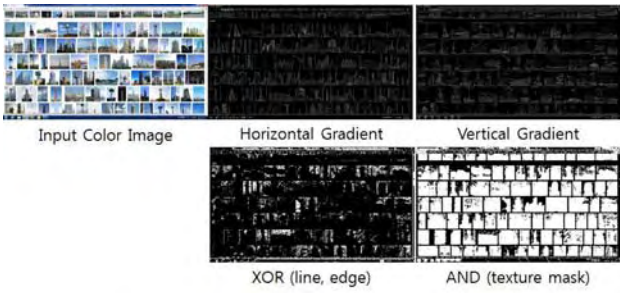
(그림 1) 제안 방법의 전체 흐름도

영상의 그래디언트는 다음의 식 (1)을 이용하여 측정한다.

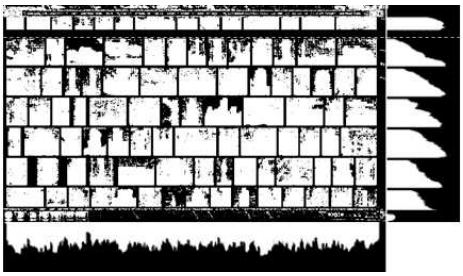
$$f_H(x,y) = I^c(x+1,y) - I^c(x-1,y) \quad (1)$$
$$f_V(x,y) = I^c(x,y+1) - I^c(x,y-1)$$

그림 2와 같이 AND 연산자를 이용하여 획득한 이진 영상에 PPC(Projection Profile Cutting)을 적용하여 영상의 프로파일을 측정한다[1]. 그림 3에서 측정한 그래디언트 프로젝션 프로파일을 보여주고 있다. 이때, 그래디언트

영상은 많은 노이즈를 포함하고 있다. 이를 제거하기 위해 CCL을 이용하여 각 그래디언트 픽셀을 레이블링한 뒤, 적은 수의 픽셀을 가지는 레이블을 제거함으로써 선명한 프로파일을 획득하였다.



(그림 2) 두 그래디언트 영상의 XOR, AND 영상



(그림 3) 그래디언트 프로젝션 프로파일

분할된 각 하위 블록을 배경, 텍스트 및 하위 영상으로 분류해야 한다. 본 논문에서는 AdaBoost + LBP를 이용한 기계학습 기반의 분류를 사용하였다. 기존에 성능이 검증된 여러 가지 특징 기술자가 개발되었으나, 본 연구에서는 특정 객체의 검출이 목적이 아니라 일반 영상과 텍스트 영상의 분류가 목적이다. LBP는 텍스쳐 분석을 위해 개발된 특징으로써, 텍스쳐 정보는 객체 인식에 유용하게 사용될 수 있다. 따라서 LBP는 본 연구와 가장 적합한 특징 기술자라고 할 수 있다. AdaBoost 및 LBP는 P. Viola[2]와 T. Ojala[3]의 논문에서 자세히 설명하고 있다. LBP는 다음 식 (2)를 이용하여 획득할 수 있다.

$$LBP(x,y)=\sum_{p=0}^{P-1} s\left[f(x,y)-f(x_p,y_p)\right]\cdot 2^p \quad (2)$$

$$s(z)=\begin{cases}1, z \geq 0\\ 0, z < 0\end{cases} \quad (3)$$

3. 실험 결과

실험을 위해 AdaBoost의 훈련을 위한 텍스트 및 자연 영상 샘플은 웹페이지 및 포토샵과 같은 툴을 이용하여 수집 또는 제작하였으며, 각 훈련 샘플은 모두 다른 해상도를 갖는다. 다음 그림 4는 AdaBoost+LBP를 이용하여 검출한 하위 영상을 보여준다. 검출된 하위 영상을 붉은색으로 나타내었다.

(그림 4) 실험 결과. 파란색은 하위 영상, 붉은 색은 텍스트 영역, 녹색은 배경

4. 결론

본 논문에서는 윈도우 영상의 3D 변환을 위해 하위 영상을 탐지하는 방법을 제안하였다. 실험 결과로부터 검출율을 측정한 결과 93.3%를 달성하였고, 이로부터 AdaBoost를 이용한 기계학습 방법이 하위 영상 검출을 위한 효과적인 방법임을 입증하였다.

참고문헌

[1] G. L. Nagy, S. Seth, and M. Viswanathan. “A prototype document image analysis system for technical journals,” Computer, 25(7):10-22, July 1992.

[2] P. Viola and M. Jones, “Robust Real-Time Object Detection,” Int’ Conf. on Computer Vision, pp. 12-14, July 2001.

[3] T. Ojala, M. Pietikainen and D. Harwood, “A comparative study of texture measures with classification based on feature distributions,” Pattern Recognition 29, pp. 51-59. 1996.