

거리 정보를 활용한 문자 분할

장석우[○], 박영재^{*}, 김계영^{*}, 최현준^{**}

[○]안양대학교 디지털미디어학과

^{*}승실대학교 컴퓨터학부 ^{**}목포해양대학교 전자공학과

e-mail:swjang@anyang.ac.kr[○], {p1y1j1,gykim11}@ssu.ac.kr^{*}, hjchoi@mmu.ac.kr^{**}

Character Segmentation Using Depth Information

Seok-Woo Jang[○], Young-Jae Park^{*}, Gye-Young Kim^{*}, Hyun-Jun Choi^{**}

[○]Dept. of Digital Media, Anyang University

^{*}School of Computing, Soongsil University

^{**}Dept. of Electronic Engineering, Mokpo National Maritime University

● 요약 ●

본 논문에서는 입체영상을 분석하여 3차원의 영상 내에 나타나는 문자 영역을 효과적으로 분리하는 알고리즘을 제안한다. 제안된 알고리즘은 먼저 입력된 영상에서 질감 특징을 이용해 문자영역이 존재하는 후보 영역을 분할하고, 후보 문자영역 중에서 문자열만을 형성하는 영역을 추출한다. 그런 다음, 지역화된 문자영역을 문자와 배경으로 분리하며, 거리 특징을 활용하여 추출된 문자영역이 비 문자영역을 포함하지 않고 문자영역만을 포함하고 있는지를 최종적으로 검증한다. 실험에서는 제안된 방법을 여러 가지 영상에 적용하여 테스트 해 보았으며, 제안된 방법이 기존의 방법에 비해 보다 정확하게 문자영역을 추출함을 확인하였다.

키워드: 문자 추출(character extraction), 질감(texture), 거리(depth), 신경망(neural network)

I. 서론

문자 영역을 강건하게 분할하기 위한 다양한 연구 방법들이 관련 문헌에 계속해서 소개되고 있다. [1]에서는 블록 단위로 영상을 분할한 후, 문자 영역에서 가로 및 세로 방향의 고주파 요소가 매우 크다는 특징을 활용하여 문자를 추출하였다. [2]에서는 문자 영역의 에지를 강하게 남기고 나머지 에지들은 삭제함으로써 문자 영역을 획득하는 방법을 제안하였다. [3,4]에서는 채도 값을 이용하고 배경과 문자 사이의 색상 전이지도(transition map)를 생성하여 문자영역을 검출하였다.

이와 같이 문자를 분할하는 많은 기법들이 연구되어 오고 있으나 여전히 해결해야 할 문제들이 남아 있다. 특히, 기존의 많은 방법들은 주로 2차원 영상을 대상으로 문자 영역 추출을 시도하고 있다. 최근 들어, 3차원 스테레오 입체영상이 많이 보급되고 있으나 3차원적인 특징을 활용하여 3차원 영상에 존재하는 문자를 추출하는 연구는 현재 거의 찾아보기가 어려운 실정이다.

따라서 본 논문에서는 3차원의 스테레오 입체영상을 받아들인 후 질감(texture)과 깊이(depth) 특징을 기반으로 영상에 포함된 문자영역을 강건하게 추출하는 기법을 소개한다.

II. 거리 특징 기반의 문자분할

제안된 방법에서는 3차원으로 촬영된 입체영상을 받아들여 영상으로부터 문자가 존재할 것이라고 예상되는 후보 문자영역을 질감 특징, 즉 식 (1)과 같은 DCT 계수를 이용하여 분할한다. 그리고 후보 문자영역의 문자열을 형성하는 부분에서 에지가 군집된다는 특성을 이용해 문자영역 지역화를 수행한다.

$$c_{uv} = \frac{1}{N} K_u K_v \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I_{xy} \cos \frac{\pi u(2x+1)}{2N} \cos \frac{\pi v(2y+1)}{2N} \quad (1)$$

그런 다음, 지역화된 문자열에서 배경을 분리하고 문자만을 추출한다. 마지막으로, 거리 특징과 신경망을 활용하여 추출된 문자영역이 비 문자영역을 포함하지 않고 문자영역만을 포함하고 있는지를 최종적으로 검증한다.

본 논문에서는 그래프 컷(graph cut) 기반의 방법을 이용하여 3차원의 깊이정보를 추출하였다. 그리고 인공신경망에 사용되는 특징은 후보 문자영역의 전경과 배경 영역 각각에 대해서 에지의 방향에 따른 크기를 16개 방향으로 양자화해서 사용하였다.

III. 실험결과

그림 1은 본 논문에서 제안한 방법을 사용하여 최종적으로 추출한 문자영역을 보여준다. 그림 1 (a)는 좌측과 우측 영상으로부터 추출한 깊이 영상을 보여주는데, 좌측 하단에 표시된 사각형 영역이 문자라고 판단된 영역이다. 그리고 그림 1 (b)는 좌측과 우측 원본 영상에서 추출된 문자영역에 해당하는 부분을 사각형으로 표시하여 보여주고 있다.



(a) Depth image



(b) Segmented left and right image

그림 1. 문자 분할

Fig. 1. Character segmentation

그리고 기존의 인공신경망만을 이용한 방법과 제안된 3차원의 깊이정보를 활용한 문자영역 분할 방법의 정확도를 측정한 결과 기존의 방법은 91.0%, 제안된 방법은 95.4%의 정확도를 보여주었다. 결과적으로, 제안된 방법은 깊이정보의 도움으로 오류가 발생하는 경우를 상당수 해결함을 확인하였다.

IV. 결론

본 논문에서는 입체영상에 존재하는 문자영역을 질감과 깊이 특징을 이용하여 추출하는 알고리즘을 제안하였다. 이를 위해, 영상으로부터 문자가 위치한다고 예상되는 후보 문자영역을 질감 정보를 이용하여 추출하였으며, 후보 문자영역 중에서 배경 영역을 제외하고 문자영역이 존재하는 영역만을 추출하는 지역화를 수행하였다. 그런 다음, 깊이 정보를 이용해 후보 문자영역을 검증함으로써 비 문자영역을 제외하고 문자영역만을 분할하였다. 향후에는 좀 더 많은 영상을 입력받아서 제안한 문자영역 분할 알고리즘의 성능을 최적화할 예정이다. 그리고 왜곡이 심한 문자 영역에 대한 분할도 시도할 계획이다.

Acknowledgements

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2011-0021984)

참고문헌

- [1] X. Qian, G. Liu, H. Wang, and R. Su, "Text Detection, Localization, and Tracking in Compressed Video," *Signal Processing: Image Communication*, Vol. 22, No. 9. pp. 752-768. Oct. 2007.
- [2] M. R. Lyu, J. Song, and M. Cai. "A Comprehensive Method for Multilingual Video Text Detection, Localization, and Extraction," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 2, pp. 243-255, Feb. 2005.
- [3] W. Kim and C. Kim, "A New Approach for Overlay Text Detection from Complex Video Scene," *Journal of the Korean Society of Broadcast Engineers*, Vol. 13, No. 4, pp. 544-553, Jul. 2008.
- [4] W. Kim and C. Kim, "A New Approach for Overlay Text Detection and Extraction from Complex Video Scene," *IEEE Transactions on Image Processing*, Vol. 18, No. 2, pp. 401-411, Feb. 2009.