4th Global Summit and Expo on

Multimedia & Artificial Intelligence

July 19-21, 2018 | Rome, Italy

Automatic video superimposed text detection based on nonsubsampled contourlet transform

Xiaodong Huang
Capital Normal University, China

Compared with other video semantic clues, such as gestures, motions etc., video text generally provides highly useful and fairly precise semantic information, the analysis of which can to a great extent facilitate video and scene understanding. It can be observed that the video texts show stronger edges. The nonsubsampled contourlet transform (NSCT) is a fully shift-invariant, multi-scale, and multi-direction expansion, which can preserve the edge/silhouette of the text characters well. Therefore, in this paper, a new approach has been proposed to detect video text based on NSCT. First of all, the 8 directional coefficients of NSCT are combined to build the directional edge map (DEM), which can keep the horizontal, vertical and diagonal edge features and suppress other directional edge features. Then various directional pixels of DEM are integrated into a whole binary image (BE). Based on the BE, text frame classification is carried out to determine whether the video frames contain the text lines. Finally, text detection based on the BE is performed on consecutive frames to discriminate the video text from non-text regions. Experimental evaluations based on our collected TV videos data set demonstrate that our method significantly outperforms the other 3 video text detection algorithms in both detection speed and accuracy, especially when there are challenges such as video text with various sizes, languages, colors, fonts, short or long text lines.
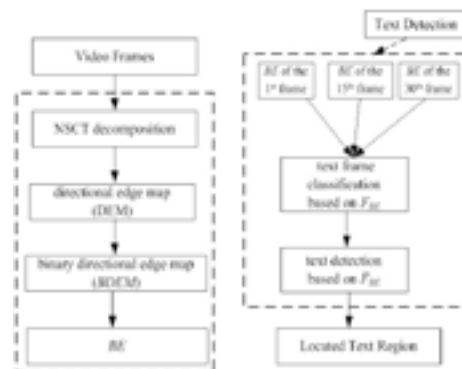

Figure: The framework of proposed method.

Recent Publications:

1. Xiaodong Huang (2014) Automatic license plate detection based on colour gradient map. Computer Modelling & New Technologies 18(7):393-397.

2. Xiaodong Huang, Huadong Ma, Charles X Ling and Guangyu Gao (2014) Detecting both superimposed and scene text with multiple languages and multiple alignments in video. Multimedia Tools and Applications 70(3):1703-1727.

3. Xiaodong Huang (2018) Automatic video superimposed text detection based on nonsubsampled contourlet transform. Multimedia Tools and Applications 77(6):7033–7049.

Biography

Xiaodong Huang is an Associate Professor of Capital Normal University, China. He received his PhD degree in Computer Science from the Beijing University of Posts and Telecommunications in 2010, MS degree in Computer Science from the Beijing University of Posts and Telecommunications in 2006 and BS degree in Computer Science from Wuhan University of Technology in 1995. His research interests include pattern recognition and computer vision.

hxd@cnu.edu.cn

Journal of Computer Engineering & Information Technology
ISSN: 2324-9307

Multimedia 2018
July 19-21, 2018

Volume 7

Page 65