Text Line Extraction in Natural Scene Images

汪, 留安

（別紙様式 2）

氏　　名　：汪 留安

論 文 名　：Text Line Extraction in Natural Scene Images
(情景画像中の文字列検出)

区　　分　：甲

論 文 内 容 の 要 旨

In this thesis, we study a text line extraction method with user-intention and complete text line extraction methods. The text line extraction is to localize the exact text bounding box by grouping individual characters into text lines. As a crucial prerequisite, text line extraction plays an important role in Optical Character Recognition (OCR) applications. Text line extraction with user-intention focuses on user-interested text line extraction to improve the user experience. We perform the user interaction by tap or swipe gesture, and make full using of user assigned location. The text character size, skew angle of text line, and reduction ratio are estimated to generate one sub-region of image. Then, the candidate text characters are decomposed to connected components in the generated sub-region of image, and the heuristic and texture features of which are fed into cascade of Gentle Adaboost classifier to remove non-text candidates. Finally, the user-intention text line path is accumulated according to the local and global permutation similarity of text characters. We achieve precision of 0.81 and 0.84 by tap gesture and swipe gesture, respectively, which outperform state-of-the-art methods. The complete text line extraction is to extract all of text lines in the image automatically without any user interaction. We formulate the text line extraction problem as a global paths optimization solution by taking advantage of the particular structure of constructed directed graph upon the text characters. The text characters are extracted by Maximum Stable Extremal Region (MSER) method. Then, we construct the directed graph upon the MSER vertices with directed edges assigned an appropriate cost value. Unary cost function and pairwise cost function are employed to calculate the cost of each directed edge from one character vertex to other character vertex. The directed graph imitates the human reading sense and text spatial arrangement from left to right. The text line paths are included in the directed graph, which constructs the relationship and eliminate the disorder of candidate text characters. The text line path global optimization produces the text line paths with minimum cost value iteratively by previous text line paths augmentation. The global optimization can identify the text line number automatically to avoid exhaustive searching, and it benefits various appearances of text lines. Furthermore, we improve the text line extraction performance with complementary usage of red, green, and blue channels. All of candidate characters are integrated into one directed graph, and text line paths in different channels are optimized together. Experimental results on ICDAR2011 demonstrate that the recall measure

is improved by 6.4%. We also achieve promising f-measure as high as 0.820, which is competitive with state-of-the-art methods. To improve the efficiency and decrease the complexity, we transform the directed graph of multi-channels into reduced graph by making full using of paths verification in different channels. The key idea is to reduce the duplicate text line or segment with numerous vertices into two vertices and one directed edge. The transformation can reduce the computation complexity successfully for text line paths global optimization with no performance sacrificing.