

Word Based Multiple Dictionary Scheme for Text Compression with Application to 2D Bar Code

K. S. NG, L. M. CHENG

Department of Electronic Engineering, City University of Hong Kong
83, Tat Chee Avenue, Hong Kong
96420200@plink.cityu.edu.hk

Abstract

In past decades, researches on the text compression are mainly on documentation applications but have seldomly considered other applications. Significant efforts have also been made to increase both the data capacity and the information density of bar code symbologies. The results of these efforts created the formats of 2D bar codes [1-3]. We will take PDF417 [1] developed by Symbol Technologies as a example. PDF417 is the most popular 2D bar code symbologies. However the storage capacity in PDF417, say around 1K bytes, has limited its widen application. Here, we proposed a text compression technique with back searching algorithm and new storage protocols. Studies on how a word-based multiple-dictionary text compression technique can be used to increase the storage capacity in 2D bar code are described. In order to speed up the search of the text, a hashing function implemented will also be described. For application in data base retrieval such as name, address, the proposed technique is particular useful as the hit rate is high so that copy code can always be sent. For data stored in 2D bar codes which are in the form of limited forms such as part numbers, location, name and reference, the compression ratio can be as high as 2 because the hit ratio can be 100%. The scheme proposed can immediately increase the data storage capacity by a factor of 2 without scarifying the physical constraints and coding mechanisms. For the decoder design, the complexity need not be complex as the decoder just require to know the 'light' and 'dark'. To let the dictionaries becomes more 'intelligent', a sub-dictionary is proposed which let the encoded text more independent.

References

1. T. Pavlidis, J. Swartz, and Y.P. Wang, "Information Encoding with Two-Dimensional Bar Codes," Computer, June 1992, pp.18-28
2. Datastrip International Inc., "<http://www.datastrip.com/>"
3. International DataMatrix Inc., "<http://www.dmatrix.com/>"