Browsing the contents of digital libraries of historical documents by word spotting – Latest Achievements.

Marçal Rusiñol[*][†][1]

[1]Computer Vision Center (Centre de visio per computador) (CVC) – Edifici O - Campus UAB - 08193 Bellaterra (Cerdanyola) - Barcelona, Spain, Espagne

Résumé

The problem of word spotting can be defined as the task of locating the regions within a document image which are likely to contain an instance of a given queried word without explicitly recognizing it. Word spotting techniques are particularly interesting to provide accessibility to document collections in which the performance of OCR engines is still poor. One of the most common scenarios is the use of word spotting techniques in historical documents. Although word spotting is quite a mature research field (the first works on handwritten word spotting were published more than 15 years ago), it is far from being solved. In this talk we will identify the latest achievements in that field and explore the open challenges yet to be solved.

[*]Intervenant
[†]Auteur correspondant: marcal@cvc.uab.es