

發明名稱： 中文書類の自動分類方法

Method for automatic classification of Chinese documents

發明名稱：	中文書類の自動分類方法
專利證書番号：	135410
出願番号：	087121388
出願日：	1998/12/21
公開番号：	439042

説明

The present invention relates to a method for automatic classification of Chinese documents which breaks expressions of every documents to be classified by the method of breaking expressions and puts all vocabulary processed by breaking expressions into the vocabulary table. Every vocabulary in the vocabulary table is processed by feature identification to get the feature value of every vocabulary of the documents to be classified in the vocabulary table. A feature combination process is performed to the feature value of every vocabulary of one document to be classified in the vocabulary table to get the feature vector of the document to be classified. Then, a document similarity identification step is performed to the feature vector of every document to be classified to get the similarity coefficients of every and other document to be classified. At last, every document to be classified is classified by sequence of the similarity coefficients to combine the similarity coefficients of the preceding sequence into the same kind, and thus the classification of documents is completed.

