

A morphological representation of characters and its application in font identification

Une représentation morphologique des caractères appliquée à l'identification de polices

Wei Ming, Noboru Babaguchi, and Tadahiro Kitahashi*

Font identification is very important in document processing. To date, most methods use a font database and therefore require expensive computation. In this paper, a novel method is presented to distinguish between different fonts in Japanese documents. In our method, a new character representation, called a G font pattern, is generated from a character using mathematical morphology operations. This G font pattern is used as a reference to identify fonts. Experimental results have shown a high degree of accuracy in identifying more than 10 000 characters of multiple fonts and sizes. The method is effective for different types of characters: Kanji, Hiragana, Katakana and alphanumeric characters as well.

L'identification de la police de caractères est très importante dans le traitement des documents. Jusqu'à maintenant, la plupart des méthodes utilisent une base de données de polices et requièrent donc une quantité imposante de calculs. Cet article présente une nouvelle approche d'identification de polices de caractères japonais. Cette approche utilise une méthode de représentation des caractères appelée patron de caractère G. Cette représentation est générée par des opérations de morphologie mathématique appliquées sur le caractère original. Le patron G est utilisé comme référence pour l'identification des polices. Les résultats expérimentaux démontrent que l'exactitude de la méthode est très élevée dans l'identification de plus de 10 000 caractères de polices multiples de diverses dimensions. La méthode est applicable à divers types de caractères: Kanji, Hiragana, Katakana, de même que pour les caractères alphanumériques.

*The authors are with ISIR, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan. E-mail: {ming.babaguti, kitahashi}@am.sanken.osaka-u.ac.jp