

要 旨

ファジィクラスタリングを用いた 部分領域に基づく画像検索の性能向上

立花 啓海

インターネットの高速化やカメラ付き携帯電話の普及で、気軽にウェブ上にデジタルデータとして画像をアップロードできるようになり、膨大な量の画像がウェブ上や個人の計算機の記憶領域に蓄積されている。この膨大な画像データを効率的に検索する技術の重要性は高まってきている。

画像検索には様々な手法が提案されているが、その多くの手法が、キーワードやキー画像などのクエリを事前に所持している必要がある。提案手法では、データベース中の画像を基に検索用のキー画像を作成する。ユーザは、それらの位置を指定し、組み合わせて配置することで要求画像に類似するクエリを作成する。これにより、ユーザの要求する画像に対して位置的にも類似する画像を絞り込むことを可能にする。これまでの画像検索ではハードクラスタリングによりインデクシングを行う手法が多く用いられ、ユーザとシステムの間でセマンティックギャップが発生していた。そこで、ファジィクラスタリングを用いることで曖昧さを考慮したインデックスを作成し、セマンティックギャップを減少させている。

実験では、データベースに COREL 画像データベース 2000 枚 (20 カテゴリ × 100 枚) を使用し被験者 16 名により行った。比較対象は、画像管理ソフトである imgSeek の絵を描くことで画像検索を行うドローサーチと比較し、上位に要求画像が含まれる割合と類似画像が含まれる割合を比較した。要求する画像に対する適合率は本手法は 25%、imgSeek は 21% とほぼ同等、類似画像に対する適合率は、本手法 39%、imgSeek が 29% と本手法が 10% 高い結果となりユーザが作成したクエリに類似する画像が検索できることを確認した。

キーワード 画像検索, ファジィ C-means, クラスタリング, CBIR, RBIR

Abstract

Performance Improvement of Region-Based Image Retrieval using Fuzzy Clustering

Hiromi TACHIBANA

A large number of images exists on the web, and many image retrieval researches have been proposed. Particularly content-based image retrieval is an important research area, which does not require any keyword or image for search query.

We propose a new method of image retrieval in which the key images for the retrieval is produced based on the images in the database. A user makes a query by choosing and locating the key images. This enables more accurate image retrieval compared to conventional image retrieval such as visual key image retrieval in terms of the position of sub images.

Conventional systems use clustering algorithms like k-means. However, there is a semantic gap problem that is caused by the interpretation of users and the system when key images are chosen. In this study, we attempt to reduce the semantic gap by a flexible indexing of the location information and the key images, produced by fuzzy clustering.

Experiments are performed using over 2000 (20 categories \times 100 images) images in COREL image database and by sixteen examinees. The ratio of user's demand image in the retrieved images are calculated using both of proposed method and the draw search tool of imgSeek. The ratios of images similar to user's demand in the retrieved images are also calculated. While the relevance ratio to the user's demand image has only minor difference, the relevance ratio to a similar image achieved a result that is 10% higher.

key words Image Retrieval, Fuzzy C-means, Clustering, CBIR, RBIR