

要 旨

部分領域画像検索における分割粒度の検索精度に与える影響

河野拓司

類似画像検索の中で画像部分領域の類似性に基づくビジュアルキー型画像検索が提案されている。しかし、従来の検索方法ではユーザは、要求する画像に提示されているビジュアルキーが使われているか否かの選択しかできず、類似するビジュアルキーが要求する画像のどの部分にあるかという位置関係は考慮されていなかった。そこで位置情報を含むビジュアルキーによりクエリを構成するシステムが提案されたが、 3×3 に分割数を固定しているため、分割数を変更することによる影響は明らかになっていない。そのため、被験者と画像における視覚的ギャップがビジュアルキー選択の誤りの原因になり、検索精度に影響を及ぼしている可能性があると考えられる。本研究では、画像の分割数を変更した場合の検索精度の変化を検証し、ビジュアルキーの生成方法についての基礎的な知見を得ることを目指す。本研究の実験内容は、画像を様々な分割数で分割してビジュアルキーを生成し、分割数の違いによる検索精度の影響を測定する。性能評価では、被験者 10 名による実験を行い、従来手法のビジュアルキー型画像検索と提案手法による検索の場合との適合率、再現率、F 値を比較する。実験の結果、 5×5 の分割数の検索精度が高いことを示す。従来手法のビジュアルキー型画像検索の場合、適合率 11%、再現率 24%、F 値 0.15 であるのに対し、 5×5 の分割数では適合率 36%、再現率 33%、F 値 0.35 となり、検索精度が向上していることを確認している。

キーワード ビジュアルキー，類似画像検索

Abstract

Effectiveness of Sub-Image Scale to the Precision of Region-Based Image Retrieval

KONO , Takushi

Visual-key image retrieval based on the similarity of partial area of image , called sub-image , has been studied as a region-based image retrieval . Conventional visual-key image retrieval measures a similarity between sub-image and query , which is input by a user regardless the location of the sub-image in images . Location based visual-key image retrieval has been proposed , however the size of scale of sub-images are fixed to 3×3 division . The size of sub-images or the number of division is an important issue to the accuracy of image retrieval . The size of sub-image affects the semantic gap between user's impression and therefore this affect the precision of retrieval . In this thesis , all images in the image database are divided to 2×2 , 3×3 , 4×4 , 5×5 and sub-images are constructed . These sub-images are clustered by Ward clustering algorithm . Twenty visual-keys are produced . Precision , recall , and F-values are derived by image search assessment . Image search assessment are conducted with ten subjects . Precision , recall , and F-values are compared with those of conventional method , 3×3 division . The result shows that the accuracy is the most highest in the case of division of 5×5 . The precision is 36% , the recall is 33% , and F-value is 0.35 for 5×5 division , while the precision is 11% , the recall is 24% , and the F-value is 0.15 for conventional 3×3 division .

key words visual-key , content-based image retrieval