

知識情報工学専攻	学籍番号	041019
申請者氏名	小塚 崇洋	

指導教員氏名	村越 一支
--------	-------

論文要旨 (修士)

論文題目	改良成長型階層自己組織化マップによる画像の自動分類
------	---------------------------

大量の画像データを自動的に分類することでより効率的に画像データを利用することができるようになると考えられる。画像の自動分類に関する研究には、教師あり学習によるものと、教師なし学習によるものがある。あらゆる画像の分類を自動的に行なうためには、教師データを必要としない教師なし学習が望ましいと考えられる。そのため、本研究では教師なし学習における画像の自動分類精度の向上を目指す。

画像の自動分類に関する研究として、教師あり学習による画像の自動分類を行なっている研究は多くあり、高精度な分類が実現されている。しかし、教師なし学習による自動分類に関する研究はあまり行なわれていない。そこで、本研究では高精度な分類を実現している教師あり学習による分類を行なう際に用いている画像特徴量を教師なし学習に適用し、教師なし学習による画像の自動分類を行なう。分類に用いる特徴量は、教師あり学習により高精度な分類を実現している Varma and Ray(2007) の手法で用いられている特徴量のうち、特徴量の次元数が固定であり、あらゆる画像に対して汎用性の高い Shape₁₈₀ と AppGray(Bosch et al., 2007) という二つの特徴量を用いる。

画像の自動分類を行なうための教師なし学習アルゴリズムには Category Utility を用いた成長型階層自己組織化マップ (Category Utility with Unification - Growing Hierarchical Self-Organizing Map: CUU-GHSOM)(藤川, 2007) を元に、画像分類を対象として改良を行なったアルゴリズムを用いる。CUU-GHSOM はマップサイズや階層の成長を自律的に行なっているため、分類すべきカテゴリをあらかじめ定めることなく、階層的な分類が実現できると考えられる。しかし、CUU-GHSOM には階層成長処理のためのクラスタ統合処理が十分に行なわれていないことや、画像特徴量間の比較において適切な距離計算方法を用いていないといった問題がある。これらの問題を解決することで、CUU-GHSOM による画像の自動分類精度の向上を目指す。従来のクラスタ統合処理が十分ではないという問題においては、単一のクラスタへの統合しか行なっていない従来手法に対し、複数のクラスタへの統合を許容することでクラスタ統合処理の改善を行なった。距離計算方法においては CUU-GHSOM で用いられているユークリッド距離の代わりに、Bosch et al.(2007) の様々な距離関数による画像分類精度の比較において、高い分類精度を示していた χ^2 距離を用いることで、画像分類により適した距離計算方法への変更を行なった。

提案手法の分類精度を検証するため、画像の分類研究に広く用いられている画像データベースである Caltech-101(Li et al., 2004) を用いた実験を行なった。実験は Caltech-101 の 102 カテゴリ全てを用いた分類と、58 カテゴリを用いた分類を行なった。この 58 カテゴリは、教師あり学習において分類率が低いカテゴリを、教師なし学習による分類も不可能であると仮定し除外したものである。実験の結果、従来の CUU-GHSOM と提案手法の分類率を比較すると、102 カテゴリでは 21.5% から 26.4% に、58 カテゴリでは 32.4% から 41.0% にそれぞれ分類精度の改善が見られた。これらの結果、提案手法により教師なし学習における画像の自動分類精度の向上が確認できた。