

氏名(国籍)	安	際	元(中国)
学位の種類	博士(工学)		
学位記番号	博甲第3169号		
学位授与年月日	平成15年3月25日		
学位授与の要件	学位規則第4条第1項該当		
審査研究科	工学研究科		
学位論文題目	Index Method Based on Dimensional Reduction (次元削減による索引法)		
主査	筑波大学教授	理学博士	大保信夫
副査	筑波大学教授	工学博士	西原清一
副査	筑波大学教授	工学博士	海老原義彦
副査	筑波大学教授	理学博士	北川博之
副査	筑波大学助教授	博士(理学)	加藤和彦

論文の内容の要旨

画像、文書、音声などを対象としたマルチメディアデータベースにおける類似検索においては、データを多次元特徴空間上の点として表現し、問い合わせ点からの近傍検索を用いる方法が一般的である。しかしながら、近年、高次元空間における近傍探索においては、いわゆる「次元の呪い」の現象が深刻であり、低次元における索引技法をそのまま適用しても効率が向上しないことが明らかになってきた。本論文では、このような背景のもと、高次元空間上の近傍検索の効率を改善することを目的とした次元削減技法及び次元削減に基づく複数の索引技法を提案している。

提案は4つの分類される。第1の提案は、 L_1 距離を持つ高次元空間をユークリッド空間に埋め込み、ユークリッド空間上の次元削減技法であるFastMapと代表的な空間索引R-木を適用することにより高次元空間の近傍検索を行う手法である。本提案を将棋の局面の類似検索に応用し、その有効性を検証している。第2の提案は、分布に偏りのあるデータに対するデータごとに重要な次元を選び出す時限削減方の提案である。本方法をデータ圧縮に基づく高次元索引VA-fileに適用した新しい索引CVA-file (Compact VA-File) も同時に提案している。提案索引を実データおよび合成データに適用し、既存の各種高次元索引と比較し、その有用性を実証している。第3の提案は、データごとの次元削減法の考え方を、巨大な時系列データセットに適用することにより得られるグリッドに基づく索引技法DDR (Datawise Dimensionality Reduction) である。本提案の特徴は、従来の時系列索引が時間軸のセグメントへの分割をベースとしていたのに対して、データ値の分割を考慮した二次元格子構造への分割に基づいている点である。実データに対して提案方式を適用し、その有用性を実証している。第4の提案は、多次元空間上のデータ分布パラメトリックな可視化を実現するHyper Mapの技法である。本技法は、ユークリッド空間上の次元索引技法FastMapを一般化したものとなっている点に1つの特徴がある。また次元削減の過程をパラメトリックに変化させることが可能である点が大きな特徴となっている。本手法を基にした、会話的なデータクラスタリングシステムのプロトタイプシステムを試作し、その有用性を実証している。

審査の結果の要旨

L_1 距離を持つ高次元データに対する索引技法は独自のものである。高次元データの多くはZipf分布に従うことが多いことを考えると、偏りのあるデータに対するデータごとの次元削減技法は独自であるだけでなく実用面の価値が大きいと思われる。この考え方を時系列データに適用することは独自のものである。Hyper Map の提案は、パラメータの決め方など今後検討の必要な点が多いが、パラメトリックな可視化の基礎的な部分を開拓している点が評価できる。以上により、本研究は情報工学上の貢献が大きいと判断される。

よって、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。