



マルチメディアデータの自己組織化と検索に関する研究

波多野, 賢治

(Degree)

博士 (工学)

(Date of Degree)

1999-03-31

(Date of Publication)

2008-03-11

(Resource Type)

doctoral thesis

(Report Number)

甲1925

(JaLCD0I)

<https://doi.org/10.11501/3156326>

(URL)

<https://hdl.handle.net/20.500.14094/D1001925>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



氏名・(本籍)	はたの けん じ 波多野 賢 治	(兵庫県)
博士の専攻分野の名称	博士(工学)	
学位記番号	博い第147号	
学位授与の要件	学位規則第4条第1項該当	
学位授与の日付	平成11年3月31日	
学位論文題目	マルチメディアデータの自己組織化と検索に関する研究	

審査委員	主査 教授 田 中 克 己		
	教授 上 原 邦 昭	教授 瀧	和 男

論 文 内 容 の 要 旨

近年、計算機ネットワークや蓄積メディアの進歩により、World Wide Web (WWW) やオンライン情報サービスなどに注目が集まっている。また、これらのサービスで扱われている情報も文字や数値などの文書を初めとして、静止画像、映像、グラフィックスおよび音声など多岐にわたる。これらマルチメディアデータは情報化社会において日を追うごとに増加を続けており、WWW上に存在するWeb文書のページ数だけでも1997年12月の時点で3億2千万ページにもものぼると推定されている。このように増え続けるマルチメディアデータを設計・管理するための枠組みとしてマルチメディアデータベースが注目されているが、汎用的なマルチメディアデータベースモデルやメディア間の制約を表現できるデータモデル、および一貫性を保持するための構成管理手法は確立されていない。こうした背景から、マルチメディアデータの分類方法や自己組織化の技術、それらデータを効果的かつ効率的にブラウズする方法について様々な研究がなされている。

まず、文書の分類検索機構においては、文書の構造を解析しその集約的な構造の把握を行ったり、WWW上を周期的に巡回して情報を収集しインデックスを生成する検索エンジンとそのインデックスを利用したデータベースを連動したシステムの構築が必要となってくる。しかし、これらの対処を行ってもなお以下のような問題が存在する。

- 対象となるデータが膨大であるため、それらのデータをデータベースに格納する作業に大変な労力と手間を要する。したがって、計算機を利用してそれぞれの情報から適切な特徴量を抽出するシステムが必要となる。

- 分類・検索されるデータ数が多く、1度の問い合わせだけでは的確な情報を得ることが難しい。したがって、言葉による問い合わせ表現に関する枠組みの確立、もしくは検索されるデータを効果的に扱うことができるユーザインタフェースの仕様を確立する必要がある。

一方、映像データの分類・検索に関する研究においては主に2種類の方法が行われている。1つは情報に付属しているインデックスなどの2次情報に基づく分類・検索の方法、もう1つは、映像全体

の色やそれらに映っている物体などの特徴量に基づく分類・検索の方法である。それぞれの研究の中には商用化レベルにまで到達した精度の高いものも存在するが、それでもなお、以下のような問題が存在する。

- それぞれの分野に限って見れば商用化レベルにまで到達しているが、それぞれの分野の技術に関してその性能および検索精度の観点から考慮すれば限界点に到達しているように思われる。よって、さらなる進歩のためにはそれぞれの長所を生かすことが可能な分類・検索基準を新たに確立する必要がある。

- 文書の場合と同様、分類・検索されるデータ数が多く、1度の問い合わせだけでは的確な情報を得ることが難しい。したがって、問い合わせ表現に関する枠組みの確立、もしくは検索されるデータを効果的に扱うことができるユーザインタフェースの仕様を確立する必要がある。

本論文では、以上のような問題点を考慮した手法を提案しその議論を行う。また、実装したプロトタイプシステムの評価についても述べる。

第2章では、本研究を遂行するにあたり利用した基本的な事項および関連研究について述べる。ここでは、近年、情報の分類に利用されているニューラルネットワークの1つである自己組織化マップ (Self-Organizing Map: 以下 SOM と記す) および映像の特徴量抽出の際に利用した Discrete Cosine Transform (DCT) について簡単に説明する。また、システムの実装、評価の際に利用している検索エンジン、評価尺度および本論文に関連する研究について簡単に述べる。

第3章では、データ数が膨大であるマルチメディアデータ、特に文書および映像データ (カット映像) から計算機で扱える特徴量を抽出し、その特徴量 (特徴ベクトル) を基に SOM を利用してデータを分類・検索するシステムの実装方法およびその評価について述べる。ここでは、ベクトル空間法を利用したそれぞれのデータの特徴ベクトルの生成法および overview マップを利用した分類されたデータの可視化の観点から実現した3次元インタフェースの説明を行う。

第4章ではカット映像の分類に焦点をあて、映像の色や映っている物体の形といったコンテンツのみを特徴量として分類する方法、つまり内容検索技術の利用だけでは不十分であるため、映像にその意味内容を記述するための映像記述支援ツールを実装し、記述された意味内容情報とコンテンツ情報を共に利用したハイブリッド型の映像の特徴ベクトルを利用したカット映像3D-SOMの実装およびその評価実験について述べる。ここでは特に、カット映像のハイブリッド型の特徴量の生成法の説明を行う。

第5章では、シーン映像の分類に焦点を当てており、第4章まで行ってきたカット映像に与えられた記述情報を特徴ベクトルとして利用したシーン映像の3D-SOMの実装およびその評価実験について述べる。ここでは、SOMを多段階に利用することにより、初期の段階では見いだすことが出来なかった新たな分類尺度を発見でき、その分類尺度により情報を分類する方法の提案とその有効性について述べる。

第6章では、分類された文書データの分類結果に対してユーザがインタラクティブに操作を行い、一種のユーザビューを生成する分類機構の実装およびその評価実験について述べる。ここでは、ユーザのフィードバック操作をどのように特徴ベクトルに反映させて SOM を学習させるかについて述べ、さらに現在の Web 文書の検索単位をページという概念ではなく関連のあるページ群を検索単位とするためのモデルの提案を行う。

最後に、第7章では、本研究で得られた研究成果をまとめ、さらに、得られた問題点に対する今後の展開について述べる。

論文審査の結果の要旨

近年、計算機ネットワークや蓄積メディアの進歩により World Wide Web (WWW) やオンライン情報サービスなどに注目が集まっている。また、これらのサービスで扱われている情報も文字や数値などの文書を初めとして、静止画像、映像、グラフィックスおよび音声など多岐にわたる。これらマルチメディアデータは情報化社会において日を追うごとに増加を続けており、WWW 上に存在する Web 文書のページ数だけでも、1997年12月の時点で3億2千万ページにもものぼると推定されている。このように増え続けるマルチメディアデータを設計・管理するための枠組みとしてマルチメディア・データベースが注目されているが、汎用的なマルチメディアデータベースモデルやメディア間の制約を表現できるデータモデル、および一貫性を保持するための構成管理手法は確立されていない。こうした背景から、マルチメディアデータの分類方法や自己組織化の技術、それらデータを効果的かつ効率的にブラウズする方法について様々な研究がなされている。文書の分類検索機構においては、文書の構造を解析しその集約的な構造の把握を行ったり、WWW 上を周期的に巡回して情報を収集しインデックスを生成する検索エンジンとそのインデックスを利用したデータベースを連動したシステムの構築が必要となってくる。しかし、これらの対処を行ってもなお以下のような問題が存在する。

対象となるデータが膨大であるため、それらのデータをデータベースに格納する作業に大変な労力と手間を要する。したがって、計算機を利用したそれぞれの情報から適切な特徴量を抽出するシステムが必要となる。

分類・検索されるデータ数が多く、1度の問い合わせだけでは的確な情報を得ることが難しい。したがって、言葉による問い合わせ表現に関する枠組みの確立、もしくは検索されるデータを効果的に扱うことができるユーザインタフェースの仕様を確立する必要がある。

一方、映像データの分類・検索に関する研究においては主に2種類の方法が行われている。1つは情報に付属しているインデックスなどの2次情報に基づく分類・検索の方法、もう1つは、映像全体の色やそれらに映っている物体などの特徴量に基づく分類・検索の方法である。それぞれの研究の中には商用化レベルにまで到達した精度の高いものも存在するが、それでもなお、以下のような問題が存在する。

それぞれの分野に限って見れば商用化レベルにまで到達しているが、それぞれの分野の技術に関してその性能および検索精度の観点から考慮すれば限界点に到達しているように思われる。よって、さらなる進歩のためにはそれぞれの長所を生かすことが可能な分類・検索基準を新たな確立する必要がある。

文書の場合と同様、分類・検索されるデータの数が多く、1度の問い合わせだけでは的確な情報を得ることが難しい。したがって、問い合わせ表現に関する枠組みの確立、もしくは検索されるデータを効果的に扱うことができるユーザインタフェースの仕様を確立する必要がある。

本論文では、以上のような問題点を考慮した手法を提案しその有用性の検証を行っている。また、実装したプロトタイプシステムの評価についても述べている。

第2章では、本研究を遂行するにあたり利用した基本的な事項および関連研究について述べている。ここでは、近年、情報の分類に利用されているニューラルネットワークの1つである自己組織化マップ (Self-Organizing Map: 以下 SOM と記す) および映像の特徴量抽出の際に利用した Discrete Cosine Transform (DCT) について説明している。また、システムの実装、評価の際に利用している

検索エンジン、評価尺度および本論文に関連する研究について述べている。

第3章では、データ数が膨大であるマルチメディアデータ、特に文書および映像データ（カット映像）から計算機で扱える特徴量を抽出し、その特徴量（特徴ベクトル）を基にSOMを利用してデータを分類・検索するシステムの実装方法およびその評価実験について述べている。ここでは、ベクトル空間法を利用したそれぞれのデータの特徴ベクトルの生成法およびoverviewマップを利用した分類されたデータの可視化の観点から実現した3次元インタフェースの説明を行っている。

第4章ではカット映像の分類に焦点をあて、映像の色や映っている物体の形といったコンテンツのみを特徴量として分類する方法、つまり内容検索技術の利用だけでは不十分であるため、映像にその意味内容を記述するための映像記述支援ツールを実装し、記述された意味内容情報とコンテンツ情報を共に利用したハイブリッド型の映像の特徴ベクトルを利用したカット映像3D-SOMの実装およびその評価実験について述べている。ここでは特に、カット映像のハイブリッド型の特徴量の生成法について述べている。

第5章では、シーン映像の分類に焦点を当てており、第4章までで行ってきた映像に与えられた記述情報を特徴ベクトルとして利用したシーン映像の3D-SOMの実装およびその評価実験について述べている。ここでは、SOMを多段階に利用することにより、初期の段階では見いだすことが出来なかった新たな分類尺度を発見でき、その分類尺度により情報を分類する方法の提案とその有効性について述べている。

第6章では、分類された文書データの分類結果に対してユーザがインタラクティブに操作を行い、一種のユーザビューを生成する分類機構の実装およびその評価実験について述べている。ここでは、ユーザのフィードバック操作をどのように特徴ベクトルに反映させてSOMを学習させるかについて述べ、さらに現在のWed文書の検索単位をページという概念ではなく関連のあるページ群を検索単位とするためのモデルの提案を行っている。

最後に、第7章では、本研究で得られた研究成果をまとめ、さらに、得られた問題に対する今後の展開について述べている。

総じて、本研究は、マルチメディアデータの自己組織化機構と検索機構について研究し、新規性および有効性の高い方式を提案したものであり、マルチメディアデータベースの分野において重要な知見を得たものとして価値ある集積であると認める。よって、学位申請者波多野賢治は、博士（工学）の学位を得る資格があると認める。