

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 5 月 20 日現在

機関番号：11301

研究種目：若手研究（B）

研究期間：2011 ～ 2012

課題番号：23700006

研究課題名（和文） 医療データにおける計算幾何を用いた知識発見システム設計

研究課題名（英文） Design of knowledge extraction system for medical data using computational geometry

研究代表者

全真嬉（CHUN JNHEE）

東北大学・大学院情報科学研究科・准教授

研究者番号：80431550

研究成果の概要（和文）：

本研究の目的は医療データにおける現在の精度限界を打破するための可視化知識抽出モデルの提案、理論研究、システムの構築である。

医療画像データ認識の為に、画像切り出し問題を、最大重み領域問題としてアルゴリズムを与えた。基本領域の分割可能な和集合で表される領域族に対する最大重み領域問題について効率的なアルゴリズムを提案した。本結果と交差を許した和集合領域の最適切り出し問題における NP 困難性を比較すると、計算量と幾何学的性質の関連の解明として学術的に興味深い結果を示した。

研究成果の概要（英文）：

In this research gives optimal approximation algorithms for medical image data and numerical data as a theoretical hierarchical segmentation. A given pixel grid and baselines, one can compute in polynomial time a maximum-weight region that can be decomposed into disjoint base-monotone regions. To finding the maximum weight region that is the union of elementary shapes for the corresponding cases. We proved to be NP-hard. This result shows improve to relation of computational complexity and geometric properties.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	3,400,000 円	1,020,000 円	4,420,000 円

研究分野：総合領域

科研費の分科・細目：情報学・情報学基礎

キーワード：アルゴリズム、計算理論、知識検出、情報システム

1. 研究開始当初の背景

近年は社会分野のみならず医療分野でもデータマイニング技術が普及されている。医療分野では、医師の経験だけではなく実験や膨大なデータ解析を通じて客観的な根拠に基づいて診断し、治療を行う医療が重視されるようになった。このような背景により医療分野では統計に基づくデータ分析手法が広く

利用されている。ところが現行のデータマイニングシステムを医療に使うことは困難である。なぜなら現行のデータマイニングシステムは抽出・発見したルールやパターンを提示するが、その知識発見過程を明示的に示し、説明することは困難である。統計手法では一般に、仮説を立ててその仮説を検証するが、仮説や事前知識をあらかじめ固定せずに、データから頻出パターンやルール、知識を抽出

し、データを説明する仮説を構築するのがデータマイニングである。すなわち統計が仮説検証のための手法であるのに対し、データマイニングは仮説構築の手法である。

現行のデータマイニングシステムでは知識発見過程が医師に理解できる形で提示されておらず、医師の診断過程に利用できる客観的な意思決定をサポートしているとはいえない。更に、医療データマイニングで必須である数値データの取り扱いを考えると、属性間の相関関係を求めて可視化する際に、現在のグラフや相関表などの離散データ解析手法をそのまま適用することはできない。また、二値化を行うと数値データの二値化誤差から生じる情報損失のため、正確性と学習汎用性のトレードオフに関する精度限界に直面する。従って、医療データマイニングを考えると医療データに含まれる巨大数値データベースを効率的に処理する最適化アルゴリズムが必要である。

また、医療データには膨大な画像データがある。医療データにおける画像データはCT、MRI、レントゲン、超音波検査などの画像データがある。画像データの解析についても様々な研究が行われている。近年、fMRI (functional magnetic resonance imaging) が開発されており脳機能を計測・画像化する手法が開発されている。ところが、現在の手法では画像データにはノイズが多く、また膨大なデータ容量であるため、解明が困難であり、データマイニング適用に期待が寄せられている。

これらの問題を解決するために、2次記憶上の巨大数値データベースを効率的に処理する最適化アルゴリズム研究と診断過程を明示的に示す可視化システム研究が強く必要とされる。数値・画像データから知識抽出の高速化と精度を上げると共にルール of 可視化を行う事を目標とする。

2. 研究の目的

本研究の目的は医療データマイニングにおける現在の精度限界を打破するための可視化知識抽出モデルの提案である。医療データには膨大な属性の数値データと膨大な画像データがある。ところが、現在の手法では画像データにはノイズが多く、また膨大なデータ容量であるため、解明が困難であり、データマイニング適用に期待が寄せられている。本研究では医療データの数値データと画像データを数値データ集合として幾何学的に扱い、計算幾何学を用いて巨大数値データベースを効率的に処理する新しい最適化アルゴリズムを与え、上記の精度限界を超える精密な手法の提案を行う。

3. 研究の方法

本研究の最終的目的地は、提案する数値デー

タに対する階層化ルール理論のアルゴリズムの高速化と改良、医療画像データに対する知識抽出の際に、画像切り出し問題を、最大重み領域問題として扱い、医療画像の切り出しを行い、新しい最適化アルゴリズムを与え、原稿の医療診断の精度限界を超える精密な手法の提案である。

新しい領域族を階層最適化し、より高次元のルールの効率的なアルゴリズムの設計を行い、その結果を結合ルール生成だけでなく、データの視覚化及びデータマイニングへの幾何学的なアプローチにおいても有効に應用を行った。

自動的に抽出し表示された知識形態は、ユーザにより意志決定等の補助として用いられる。重要な条件は、抽出した知識形態がシンプルであり (単純性)、正確にデータの特徴を記述すること、知識としてサンプルに依存しない汎用性を持つ事さらにユーザにとって説得力があり、検証が容易であること (透明性) である。単純性と透明性の観点から、結合ルール及びそれを用いた決定木は有力な手法である。

本研究で提案する確率的な非決定性決定木構造を用いた階層構造は、現行の判定システムにおいて主流になっている決定論的な決定ルールに比較して、強いルールの影響を縮小する方法を適用する。本研究で提案する医療画像データを数値データとして扱い階層化ルールを用いることで拘束力の弱いルールで判定を行う、即ち非決定性を持たせた柔軟な決定システムの構築を行った。

4. 研究成果

本研究の目的は医療データにおける現行の知識発見システムにおける現在の精度限界を打破するための可視化知識抽出モデルの提案、理論研究、システムの構築である。

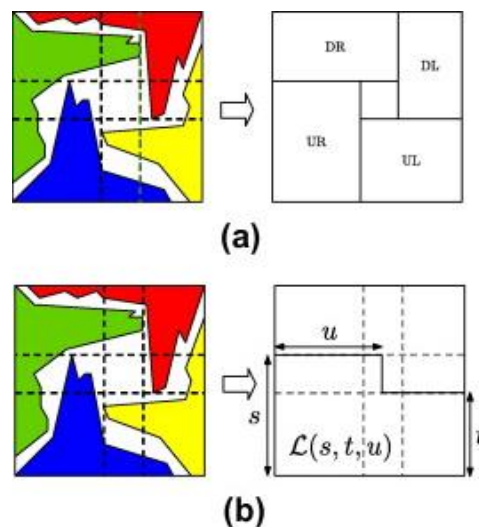


図1. 領域分割の例

- (a) four two-colored rectangles、
- (b) L-shaped regions

本研究で達成するシステムは、視覚的に判りやすく、かつ精度及び効率性の高いものであり、医療分野における意思決定支援という現代社会の必須技術として大きな実用インパクトを持ち現代社会に寄与できる。

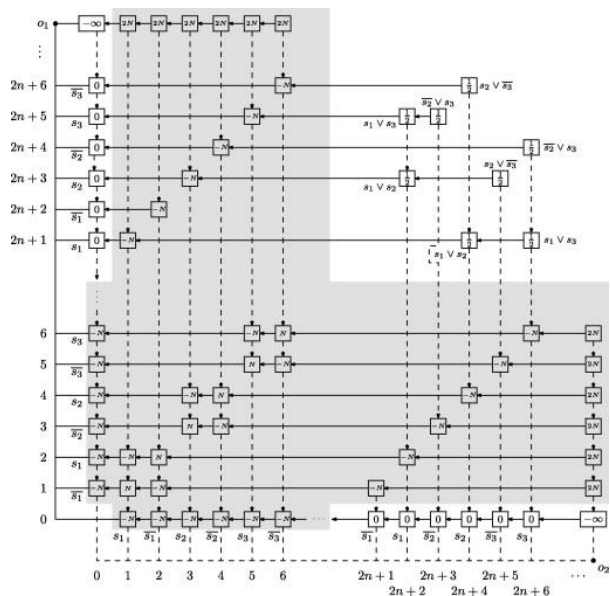


図 2. Union-MWRP instance that encodes the Max-2 Sat instance

本研究で対象とする医療データは、医療検査の数値データと画像データである。本研究では数値データと画像データを、離散アルゴリズムを用いた計算幾何学的な関数の最適近似の独自技法を利用し、データ分布を、ユーザに説明しやすい関数を用いて近似することで数値データベースからルールを抽出し、更にそれを組み立てて知識発見過程が医師にとって明確に理解できるシステムの構築を行うための理論確率を行った。これは、過去のデータマイニング技法に計算幾何学技法と幾何学的可視化を取り入れた画期的なシステムである。提案手法は計算幾何学手法を用いた医療画像データから自動的画像切り出しであり、切り出された画像を知識として提示することにより、医療画像データ解析の際に、知識として提案すると言った、医療画像データ知識抽出理論構築の成果を得た。



図 3. 入力が花の画像の場合のイメージセグメンテーションの例

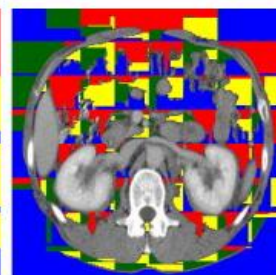
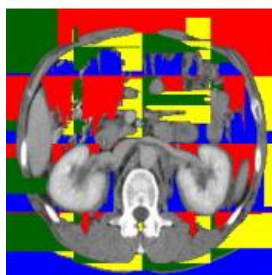
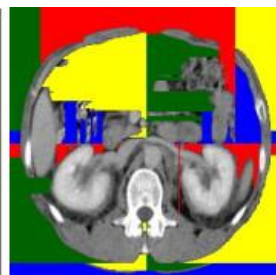


図 4. 医療画像の例 :
C T 画像の Room-edge four-colorings
(for $k = 1, 3,$ and 7)

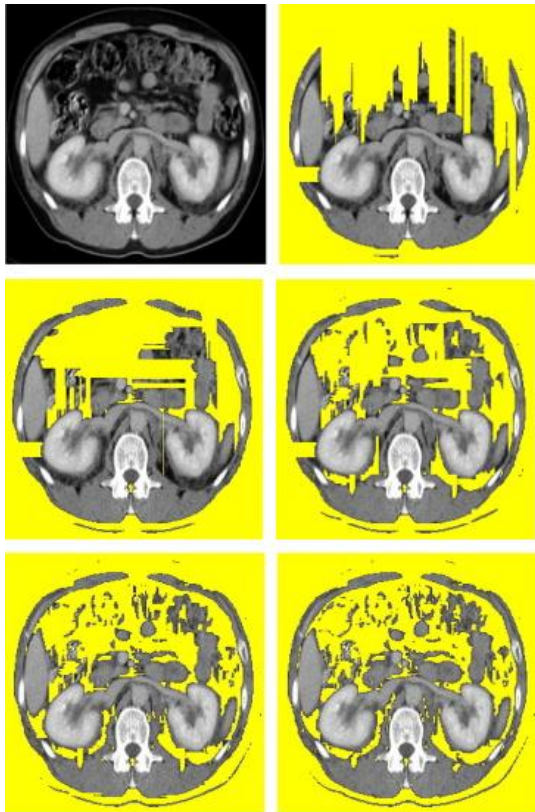


図5. 医療画像の例：
CT画像のイメージセグメンテーション
($k = 0, 1, 3, 7, 15$)

本研究の主な結果として、医療画像データ認識の為に、画像切り出し問題を、最大重み領域問題としてアルゴリズムを与えた(図3, 4, 5)。基本領域の分割可能な和集合で表される領域族に対する最大重み領域問題について効率的なアルゴリズムを提案した(図1, 2)。本論文における結果と交差を許した和集合領域の最適切り出し問題におけるNP困難性を比較すると、計算量と幾何学的性質の関連の解明として興味深い成果と考えられ、学術的に興味深い結果[雑誌論文1, 2, 3, 4, 5]を示した。

特に雑誌論文[3]は世界的に著名なComputer Vision and Image Understandingに掲載された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 5 件)

[1] Jinhee Chun, Takashi Horiyama, Takehiro Ito, Natsuda Kaothanthong, Hiroataka Ono, Yota Otachi, Takeshi Tokuyama, Ryuhei Uehara, Takeaki Uno, “Base location problems for

base-monotone regions,” , Algorithms and Computation Lecture Notes in Computer Science, LNCS7748, pp.53-64, 2013, 査読有
DOI: 10.1007/978-3-642-36065-7_7

[2] Jinhee Chun, Akiyoshi Shioura, Truong Minh Tien and Takeshi Tokuyam, “A Unified View to Greedy Geometric Routing Algorithms in Ad Hoc Networks”, Algorithms for Sensor Systems Lecture Notes in Computer Science , LNCS7718, pp.54-65, 2012, 査読有
DOI: 10.1007/978-3-642-36092-3_7

[3] Jinhee Chun, Natsuda Kaothanthong, Ryosei Kasai, Matias Korman, Martin Nollenburg, Takeshi Tokuyama, “Algorithms for computing the maximum weight region decomposable into elementary shapes”, Computer Vision and Image Understanding, Vol.116, no.7, pp.83-814, 2012, 査読有
DOI: 10.1016/j.cviu.2012.03.003

[4] Natsuda Kaothanthong, Yoshiaki Ota, Jinhee Chun, Takeshi Tokuyama, “Image Retrieval System Using Distance-Based Shape Recognition”, Proceedings of The 14th Korea-Japan Joint Workshop on Algorithms and Computation, Vol.14, pp.137-144, 2011, 査読有

[5] Jinhee Chun, Natsuda Kaothanthong, Yoshiaki Ota, and Takeshi Tokuyama, “Distance-Based Shape Invariants for Image Retrieval”, Proceedings of The 4th Annual Meeting of the Asian Association for Algorithms and Computation, Vol 4, pp.40, 2011, 査読有

[学会発表] (計 14 件)

[1] Jinhee Chun, Ricardo Gonzaro and Takeshi Tokuyama, “Space-efficient polygon reconstruction from visibility angle information”, 6th Annual Meeting of the Asian Association for Algorithms and Computation (AAAC2013), 2013年04月21日, Matsushima, Japan

[2] Jinhee Chun, Natsuda Kaothanthong and Takeshi Tokuyama, “Shape Description using Classified Distances”, 6th Annual Meeting of the Asian Association for

Algorithms and Computation (AAAC2013), 2013年04月20日, Matsushima, Japan

[3] Jinhee Chun, Takashi Horiyama, Takehiro Ito, Natsuda Kaothanthong, Hirota Ono, Yota Otachi, Takeshi Tokuyama, Ryuhei Uehara, Takeaki Uno, “Base location problems for base-monotone regions”, 7th International Workshop on Algorithms and Computation, 2013年02月15日, Kharagpur, India

[4] Jinhee Chun, Takashi Horiyama, Takehiro Ito, Natsuda Kaothanthong, Hirota Ono, Yota Otachi, Takeshi Tokuyama, Ryuhei Uehara and Takeaki Uno, “Algorithms for Computing Optimal Image Segmentation using Quadtree Decomposition”, Thailand-Japan Joint Conference on Computational Geometry and Graphs(TJJCCGG2012), 2012年12月06日, Bangkok, Thailand

[5] Jinhee Chun, Akiyoshi Shioura, Truong Minh Tien and Takeshi Tokuyama, “A Unified View to Greedy Geometric Routing Algorithms in Ad Hoc Networks”, 8th International Symposium on Algorithms for Sensor Systems, 2012年09月13日, Ljubljana, Slovenia

[6] Jinhee Chun, Akiyoshi Shioura, Truong Minh Tien and Takeshi Tokuyama, “A Unified View to Greedy Routing Algorithms in Ad Hoc Networks”, 15th Korea-Japan Joint Workshop on Algorithms and Computation(WAAC2012), 2012年07月10日, Tokyo, Japan

[7] Jinhee Chun, Natsuda Kaothanthong, Hiromi Takahashi, Takeshi Tokuyama, “Optimal Grid Decomposition for Maximum Weight Region Computation with Application to Image Segmentation”, Computational Geometry: Young Researchers Forum(CG:YRF), 2012年06月19日, Chapel Hill, USA

[8] Jinhee Chun, Natsuda Kaothanthong, Hiromi Takahashi, Takeshi Tokuyama, “How to cut a complicated figure by using scissors? Computing the maximum weight region consisting of base monotone regions”, The 5th Annual Meeting of the Asian Association for Algorithms and

Computation (AAAC2012), 2012年04月21日, Shanghai, China

[9] Jinhee Chun, Natsuda Kaothanthong, Yota Otachi, Hiromi Takahashi, and Takeshi Tokuyama, “An Algorithm for Optimally Locating Baselines using Quad Decomposition”, 2012年電子情報通信学会総合大会, 2012年3月21日, 岡山

[10] Jinhee Chun, Natsuda Kaothanthong, Hiromi Takahashi, and Takeshi Tokuyama, “An Image Segmentation using Maximum Weight Region with Shape Constraint”, The First ETH-JAPAN Workshop on Science and Computing, 2012年3月13日, Engelberg, Switzerland

[11] Jinhee Chun, Natsuda Kaothanthong, Takeshi Tokuyama, “Image Segmentation Using Maximum Weight Region”, 電子情報通信学会コンピュータシミュレーション研究会, 2011年12月16日, 名古屋

[12] Natsuda Kaothanthong, Yoshiki Ota, Jinhee Chun, Takeshi Tokuyama, “Image Retrieval Using Shape Recognition”, 第10回情報科学技術フォーラム(FIT2011), 2011年9月8日, 函館

[13] Natsuda Kaothanthong, Yoshiki Ota, Jinhee Chun, Takeshi Tokuyama, “Image Retrieval System Using Distance-Based Shape Recognition”, The 14th Korea-Japan Joint Workshop on Algorithms and Computation (WAAC2011), 2011年7月9日, Pusan, Korea

[14] Jinhee Chun, Natsuda Kaothanthong, Yoshiki Ota, and Takeshi Tokuyama, “Distance-Based Shape Invariants for Image Retrieval”, The 4th Annual Meeting of the Asian Association for Algorithms and Computation(AAAC2011), 2011年4月16日, HsinChu, Taiwan

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

名称:
発明者:
権利者:
種類:

番号：
出願年月日：
国内外の別：

○取得状況（計0件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

全真嬉 (CHUN JNHEE)
東北大学・大学院情報科学研究科・准教授
研究者番号：80431550

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：