

# インタラクティブな情報視覚化のための高速レイアウト配置手法

## 重要な情報を注目しながらデータの全体像を一画面で俯瞰(ふかん)できる

近年、情報化社会の発展に伴い、大量のデータが蓄積されるようになってきており、データ全体を俯瞰しながら大量の情報から注目すべき特徴を短時間に把握できる情報視覚化技術が注目されています。

情報視覚化技術の一つに、データの指標を矩形(くけい)の縦と横の長さに対応させ、それらの矩形を平面上に重ならないように配置することにより、大量のデータの特徴を概観する手法があります。しかし、この場合、矩形詰込み問題の解を得るための計算量が非常に大きくなってしまいます。

今回東芝は、情報視覚化技術に適用可能な矩形詰込み問題の近似解を高速に求めるアルゴリズムを考案しました。

### 大量データの視覚化

矩形詰込み問題とは、様々な大きさの矩形を2次元平面上に互いに重ならないように配置できないことで、NP困難<sup>(注1)</sup>として知られています。遺伝的アルゴリズムやシミュレーテッドアニーリング(焼きなまし法)などにより近似解を探すというのが現実的な手法です。

データを矩形に対応させ、複数の矩形を平面上に重ならないように配置することで、大量のデータの特徴を一目で見ることができます。しかし、対話的

(注1) 計算複雑性理論でNP困難とは、クラスNPに属する問題と同程度又はそれ以上に難しい問題のこと。クラスNPとは、解答が与えられさえすれば、それが正答か否かを多項式時間で判断できるような問題のクラスである。

な視覚化システムでは、数千件~数万件のデータを操作しながら変化を即時に画面に表示することが要求されるため、既存のアルゴリズムでは、この要求を満足できません。

今回東芝は、大量のデータに対してインタラクティブ(双方向的)な情報視覚化のために、矩形詰込み問題の近似解を高速で求めることができる、新しいレイアウト配置手法を考案しました。

### レイアウト配置の仕組み

矩形詰込み問題は、最適スライス木(slicing tree)探索の問題に置き換えて近似解を求めるのが一般的ですが、候補スライス木の数は矩形数(n)の指数関数になるため、矩形数の増加により計算量が急増してしまいます。

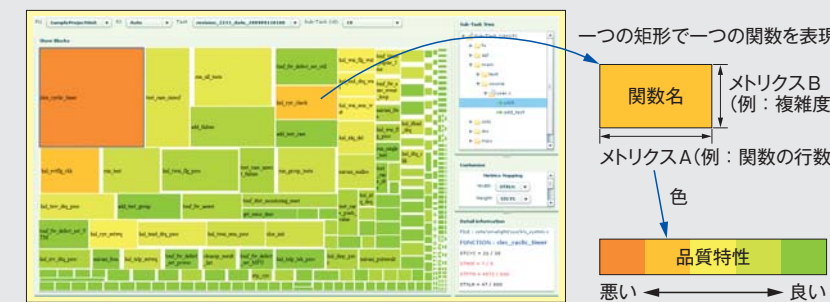
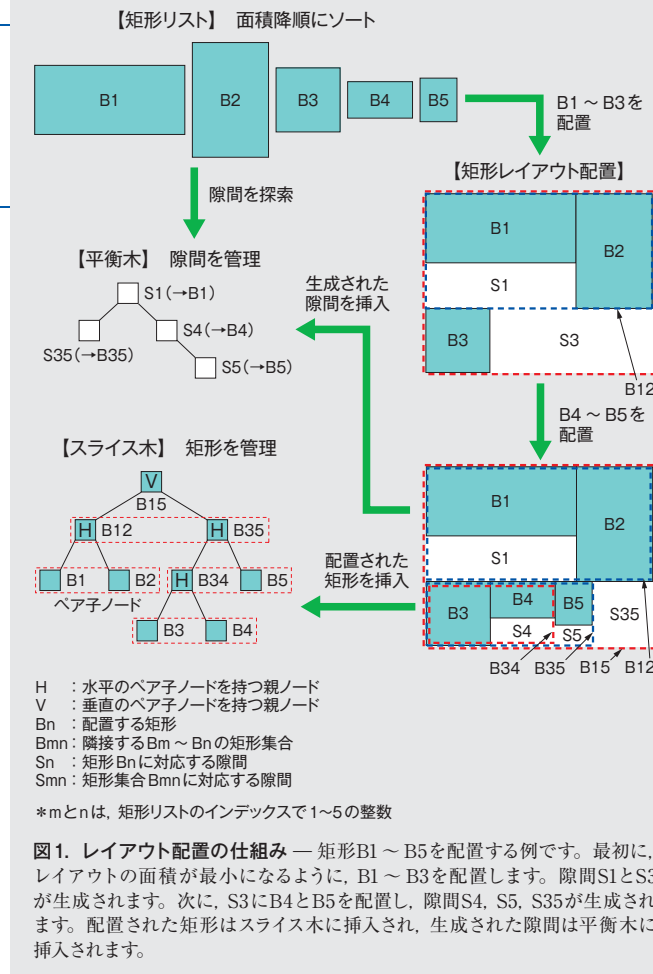


図2. ソースコード品質マトリクスの視覚化 — 関数のマトリクスA, マトリクスB, 品質判定(5段階)をそれぞれ矩形の幅, 高さ, 色にマッピングし、プロジェクト全体を一画面に視覚化します。



図3. フォトコレクションの一覧画面 — 各写真のサイズを重要度に比例させて、すべての写真を一画面に表示します。

かれるスペースを高速に探索でき、レイアウト配置の総計算量は $O(n \times \log n)$ になります。

情報視覚化では、重要な情報(大きい矩形)をどこに配置するかも大切なポイントです。そのために、大きな矩形を注目点の周りに集めるような配置を可能にしました。スライス木の各ノード(節点)のペア子ノードの順番を自由に交換できる特性を利用し、注目点と各ペア子ノードの距離の加重平均が最小になるようにその順番を決定します。スライス木の各ノードに対してBFT(Breadth-First Traversal: 幅優先探索)順にこの処理を適用すると、矩形はおおむね注目点の周囲に大きさ順に配置されます。

### 情報視覚化への適用

情報視覚化への適用例として、二つの例を述べます。

一つ目の例として、ソフトウェア開発プロジェクトの品質管理への適用を述べます。ソフトウェアの品質を評価するためにソースコード静的解析技術が多く用いられます。プロジェクト全体のソースコードの品質特性を直感的に判断するため、静的解析ツールによって得られたソースコードの品質マトリクス<sup>(注2)</sup>を、レイアウト配置手法により視覚化した例を図2に示します。ソースコードの行数、複雑度、品質判定(5段階)をそれぞれ矩形の幅, 高さ, 色にマッピング

(注2) ソフトウェアの品質を定量的に把握するための尺度や測定法。

グし、プロジェクトに含まれる関数に対応する矩形を一画面に表示します。矩形の大きさ, 横縦比, 色と矩形の分布からソースコードの品質を直感的に把握できます。

二つ目の例として、フォトコレクションへの適用例を図3に示します。各写真のサイズは重要度に比例させて、すべての写真を一画面に配置します。画面の中心から画面の周りまで写真のサイズが小さくなり、中心に置かれる重要な写真を認識しやすくなります。

新しいレイアウト配置手法を用いると、CPUのクロック周波数が3.2 GHz, メモリが1 Gバイトのパソコンで、2万件の矩形データの配置計算を約1秒で実行できます。これにより、インタラクティブな情報視覚化を実現しました。

### 今後の展望

この技術を利用すると、大量な情報からその特徴を直感的に把握することができます。

大量のデータを扱う情報視覚化技術として様々な分野での利用を検討しています。また、ほかのインタフェースと組み合わせることによって操作性を更に向上できると考えています。

### 文献

(1) Xinxiao, L., et al. "An aesthetic solution for photo collection page layout". Proc. of the first international conference on knowledge and systems engineering. Hanoi, 2009-10, Vietnam National University, 2009, p.33-38.

李 新肖

ソフトウェア技術センター  
先端ソフトウェア開発担当主務