

論 文 要 旨

Time-series Visualization adopting Overview and Summarization

(時系列データの俯瞰と縮約のための可視化)

人間文化創成科学研究科 理学専攻 情報科学領域 林 亜紀

本論文では、時系列データの俯瞰と縮約のための可視化方法を提案する。近年、様々な時系列データが蓄積されている。Web のアクセスログ・決済情報などのシステムログや、ソーシャルネットワークサービス上で蓄積される位置情報ログなどのライフログに加え、音楽メディアをテキスト化した MIDI も時系列データの一つである。それらが持つ時間的周期性に着目した分析により、将来予測や状況・内容理解の精度が向上し、リコメンデーションやマーケティングなどの経済効果が期待される。形状や空間構造をもたない抽象的なデータに対し、形状を与えて提示することにより、人間による視覚的な理解を支援する技術として、情報可視化への注目が高まっている。特に、既存の可視化手法を効率的に活用して自然現象や社会現象を効率的に分析する Visual Analytics の有効性が示されてきている。一般に、Visual Analytics では、まずデータ全体の傾向を俯瞰した上で、情報を縮約して興味深い情報を詳細観察する、といった俯瞰・縮約の分析手順が提唱されている。しかしながら時系列データは大容量かつ多属性なため、効果的な俯瞰・縮約を実現する可視化手法は少ない。本論文の目的は、大規模時系列データに対し、データ特性や観察目的に応じた効果的な俯瞰・縮約を実現することである。特に、観察支援のための自動抽出指標の導入により、可視化の可読性を向上させるとともに、俯瞰表示の中から縮約による詳細観察にふさわしい興味深い部分を抽出する作業を支援することに注力する。本論文では、一般的な多次元時系列データを扱う一般的な手法に加え、前述した位置情報データや楽譜データのような、独特なデータに対しても対応した俯瞰・縮約のための可視化手法の構築を目指して 3 つの手法を提案する。

1 つ目は、決済情報やアクセスログなどのシステムログに代表される、一般的な多次元時系列ログの俯瞰・縮約のための手法である。システムログの分析目的として、サイト運営やマーケティングなどの経営支援が挙げられる。大規模かつ多属性で、俯瞰が難しいシステムログを、高い可読性を保ちながら可視化するために、時系列を横軸に、他の属性のうち 1 つを縦軸に割り当てたヒートマップを切り替えながら観察する俯瞰方法を採用する。この際、例えば急激な値変化がある項目が多い属性など、興味深い属性を自動的に推薦する指標を導入する。また、縮約すべき項目の発見を支援するために、時間刻みの調節による集計粒度制御と、ソート・クラスタリングによる重要項目抽出を実現する。決済情報とアクセスログの実データを用いた分析により手法の有効性を確認し、一般ユーザとデータ分析に長けた専門家の両方から実用面での評価を得た。

2 つ目は、位置情報ログに代表されるライフログの俯瞰・縮約手法である。ライフログの分析目的として、情報推薦の満足度向上が挙げられる。しかしながら、ユーザニーズに合った推薦内容とタイミングの決定は難しい。ユーザの状況は時々刻々と変化し、その状況を詳細に抽出するのが難しいためである。本論文では、ライフログの様々な時間軸に基づく周期性に着目し、各行動がどの程度本人の習慣に沿って

いるのかを自動で定量的に評価する習慣度という指標を導入する。例えば旅行中などの非習慣時には、他の旅行者が現地で訪れやすい観光スポットを推薦するなどの活用が考えられる。提案手法では、多くのユーザに共通する、各場所の習慣性傾向を地図上に可視化する俯瞰表示により推薦内容決定を支援する。また、複数ユーザの習慣度変遷の羅列から選択した、興味深いユーザの詳細な習慣度変遷観察により、推薦タイミングを向上する縮約を実現する。位置情報ログの実データを用いて提案手法の妥当性を定量的・定性的に評価した。

3 つ目は、より特徴的なデータへの対応を目指した、楽譜情報における音楽構造の俯瞰・縮約手法 **Colorscore** である。楽譜情報の中でも特にオーケストラ楽曲のスコアは段数が多く、楽曲理解は難しい。**Colorscore** は、各フレーズが担う主旋律、伴奏などの役割を半自動的に抽出することにより、作曲者や編曲者、演奏者による、全体像把握と他編成へのアレンジを支援する。本論文では各パートの役割変遷や類似メロディの繰り返し構造を一画面上に描画した俯瞰表示を提案する。また、楽曲の骨格部分を自動的に判断し、他編成へのアレンジを支援するために、重要な小節のみを拡大表示する横方向の縮約と、重要なフレーズを残して段数を段階的に削減する縦方向の縮約を実現する。複数のオーケストラ楽曲への適用を通して手法の有効性を検証し、初学者と専門家両方の目線から手法を評価した。

以上の 3 手法の構築を通し、時系列データの俯瞰・縮約による可視化の実現において、可視化結果の可読性向上と、縮約箇所選択支援の両方の場面で、提案手法が導入した自動抽出指標が有用であることが分かった。時系列データの可視化分析システムを構築する際には、観察目的とデータ特徴・時間的周期性を考慮した自動抽出指標を導入した上で、時間軸、項目の両方の観点から俯瞰・縮約方法を決定することが有効であると考えられる。