

ChronoView と MDS の統合による時刻データの可視化手法の開発

安齋 康宏^{1,a)} 三末 和男^{2,b)}

概要: 様々な分野において、時刻付きデータに含まれるイベントの発生時刻の周期性や類似性などを把握することは有用である。発生時刻を分析する際には、ある種のイベント群の時間的特徴、周期的特徴、他のイベント群との時間的特徴の類似性がしばしば着目される。本論文では、大量のイベントを対象として、それらの特徴を効率的に観察するための可視化手法について述べる。著者らが開発した手法は各イベント群の時間的特徴をグリフで表現するとともに、グリフの配置によって周期的特徴と類似性を表現する。グリフの配置を求めるために、周期的特徴を位置で表現する ChronoView と擬似的な多次元尺度構成法 (MDS) を 3次元空間において統合する。本論文ではさらに、実際のデータを用いたイベントの探索に関するユースケースを通して、時刻付きデータの分析における開発手法の有効性を示す。

キーワード: 時刻データ, 視覚的分析, ChronoView, 多次元尺度構成法, グリフ

Integration of ChronoView with Pseudo Multidimensional Scaling for Visualization of Temporal Data

YASUHIRO ANZAI^{1,a)} KAZUO MISUE^{2,b)}

Abstract: In several fields handling event data, exploring periodicity and similarity of the occurrence times of events is an important task. When we analyze event data, we often focus on temporal features of event groups, periodic features of events, and similarity of temporal features among event groups. This paper illustrates a visualization method to efficiently observe such features for multiple events. The method represents the temporal features of each event group with a glyph and represents the periodic features and their similarity with an arrangement of glyphs. To determine an arrangement of glyphs, a ChronoView graph representing periodic features of event groups is integrated with pseudo multidimensional scaling (MDS) in a 3D space. This paper also demonstrates the effectiveness of the visualization method for temporal data analysis with a use case exploring significant events from an actual dataset.

Keywords: temporal data, visual analytics, ChronoView, MDS, glyph

1. はじめに

本論文では時刻付きデータを分析するための可視化手法

について述べる。本論文における時刻付きデータはイベントから構成され、個々のイベントは時刻 (タイムスタンプ) と 1 つ以上の属性を持つ。イベントとは例えば、商品の購入履歴や web ページへのアクセス履歴といったものである。属性とは例えば、購入商品、アクセスされた web ページ、犯罪の種類、ユーザ ID などが想定される。イベントの 1 つの属性に着目することでイベントの集合を抽出することができ、我々はこの抽出された集合をイベント群と呼ぶ。イベント群に含まれるイベントの発生時刻に着目する

¹ 筑波大学大学院システム情報工学研究科
Graduate School of Systems and Information Engineering,
University of Tsukuba

² 筑波大学システム情報系
Faculty of Engineering, Information and Systems, University
of Tsukuba

a) anzai@vislab.cs.tsukuba.ac.jp

b) misue@cs.tsukuba.ac.jp

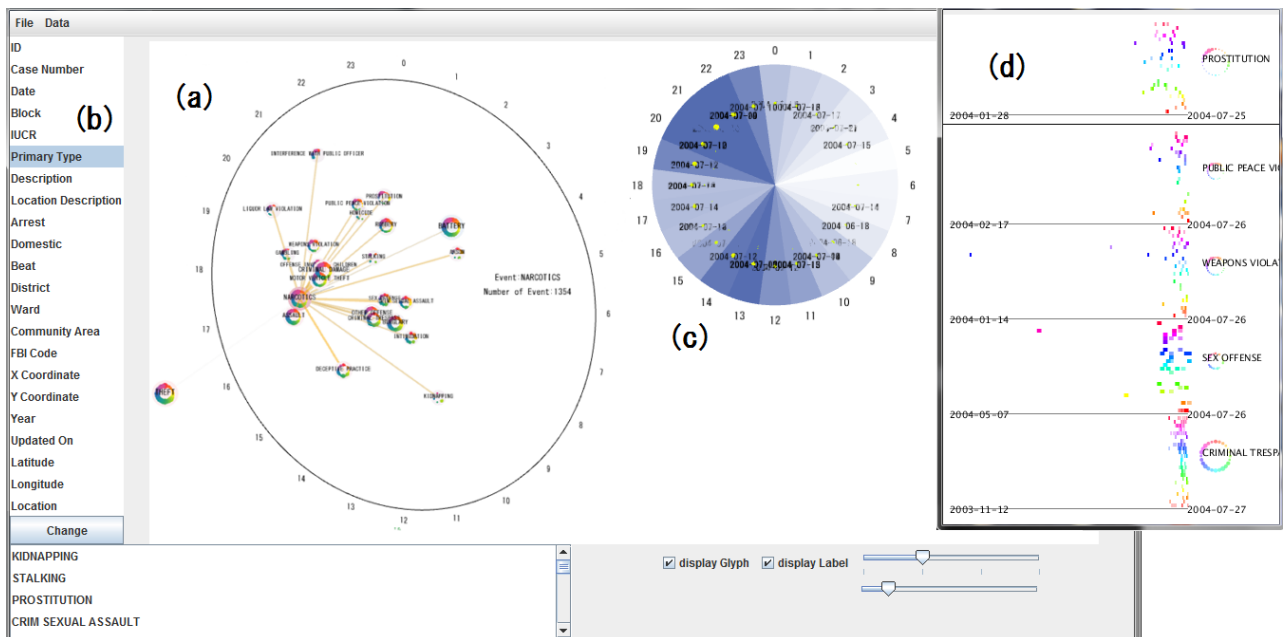


図 1 時刻付きデータ分析のためのインタフェース. メインビュー (a): 分析者がイベントの時間的特徴を観察する. 属性選択リスト (b): 分析者はイベントの属性をこのリストから選択する. 詳細タイムプロット (c): 分析者はイベント群をクリックあるいはイベント群リスト (左下) から選択することで, 詳細な特徴を分析する.

と, 発生時刻の分布や周期性などに特徴が見られることがある. 時刻付きデータを分析して, そのような特徴を把握することで, ビジネスやセキュリティなど様々な領域における知見の獲得と応用が可能となる.

イベント群の時間的特徴を表現する可視化手法には様々なものがある. イベントの発生頻度の時間変化に着目する場合には, 折れ線グラフ, 棒グラフ, 面グラフなどがしばしば利用される. 時間の周期性に着目する場合には, 時間軸を螺旋状に配置するスパイラルチャート [1] が便利である. これらのチャートは, 1つのイベント群の頻度変化をうまく可視化できる. しかしながら, 多数のイベント群を観察する場合には不都合が生じる. 1つのチャートで多くのイベント群を表現しようとする視覚的混雑が発生し, データの読み取りが困難になる. イベント群毎にチャートを描き, それらを画面上に配置する方法では, 各チャートが一定の表示領域を必要とするため視覚的な一貫性が損なわれる. たとえ大画面に配置ができたとしても, イベント群間の類似性や特異性を発見することは容易でない.

本研究の目的は多くのイベント群の時間的特徴を効率的に観察するための概観を提供することである. 我々は時刻データが備える下記の3つの特徴に着目し, それぞれの表現に適した3つの手法を1つの視覚的表現に統合する手法を開発した.

- グリフによる個々のイベント群の時間的特徴の表現
- ChronoView [2] による個々のイベント群の周期的特性の表現
- MDS [3] によるイベント群間の類似性の表現

これにより我々は時刻付きデータを異なる視点から同時に観察できる概観を構築した.

2. 関連研究

2.1 時系列データの表現

時刻付きデータの代表は時系列データである. 時系列データの表現には線グラフや棒グラフがしばしば用いられる [4]. これらは横軸に時間軸を取り, 縦軸に一定時間間隔で集計したイベントの頻度を取ることで, イベント群の可視化にも利用できる. 多くのイベント群を同時に表現したい場合, イベント群の表現を重ねるといった手法がとられる. しかしイベント群の数が増えると視覚的煩雑さがチャート内に生じ, 個々のイベント群の読み取りが困難となる.

多数のイベント群を可視化する手法として, ThemeRiver [5] や EventRiver [6] がある. ThemeRiver は積み上げグラフにおける値の変化を連続的にしたような表現を行う. EventRiver は発生頻度の時間変化をバブル状の図形で表現する. これらは多くのイベント群の時系列的な特徴を同時に空間効率よく表現する. しかしイベント群の数が増えると視覚的に比較する組の数が増え, 類似した時系列パターンの発見が困難になる. Imoto らは x - y 平面で描画される大量の線グラフを z 軸に沿って類似度順に重ねた表現を提案した [7]. この手法は時系列データ全体を概観する x - z 平面 View と, 特定の時系列データを観察するための x - y 平面 View を持つ. この手法は我々の提案手法とイベント群の類似性を同時に観察できる点が共通する.

我々の手法は加えて大量の発生時刻のパターンを観察するための機能を持つ。

2.2 周期的特性の表現

周期的な特性（例：どの時間帯に多くのイベントが発生したか）は重要な時間的特徴の1つである。このような周期性は螺旋や円周状の表現によってよく表現される。そのような表現では、発生時刻はアナログ時計のように角度で表現される。Carlisらはイベント群に含まれるイベントを螺旋状の時間軸上に配置する手法を提案した [1]。同じ時間帯に発生したイベントは直線状に並ぶため、分析者は時刻データの周期性を容易に把握できる。Weberらによる手法 [8] や、Tominskiらによる手法 [9] は色の変化によってデータの値の変化を時系列的かつ周期的に表現している。しかしこれらの手法は1つのイベント群を表現するために大きな描画領域を必要とし、比較のために多くのイベント群を並べるといったことは難しい。ChronoViewはイベントのタイムスタンプの集合を2次元平面上の位置として表現する手法である [2]。多くのイベント群を空間効率良く可視化することを目指して開発された。あらかじめ指定された表現周期に基づいて可視化する手法であるが、複数の表現周期で同時に可視化する2.5D表現への拡張も提案されている [10]。

2.3 グリフを用いた時間的特徴の表現

イベント群の時間的特徴の可視化のためにグリフを用いる手法も提案されている。Nguyenらはイベントの発生頻度の時間変化を表現したグリフを地図上に配置することでイベントの発生場所も表現した [11]。この手法はイベント群の時間的特徴と地理的な特徴を提供するが、イベント群間の類似性に関する情報は提供しない。ClockMap [12] はネットワークの階層構造を円状のツリーマップで表現し、ツリーマップに埋め込まれた円形のグリフによってネットワークのトラフィック量の時間変化を表現する。ClockMapでは、ネットワークの論理的な階層構造によってグリフの位置が決められる。

3. イベント群の視覚的表現

我々は時刻付きデータを分析する上で時刻付きデータが備える三つの特徴、(1) 個々のイベント群の時間的特徴、(2) イベント群の周期的特性、(3) イベント群間の時間的特徴の類似性を表現することが重要であると考えた。さらにこれらを1つの視覚的表現で表すことを目指した。

我々の採用したアプローチは以下の通りである。個々のイベント群の時間的特徴を小面積で表現するためにグリフを活用する。イベント群の周期的特性とイベント群間の類似性は、表現空間内に配置されたグリフの位置関係で表現する。イベント群の周期的特性を表現するために

ChronoViewによる配置を利用する。さらに、イベント群間の類似性を多次元尺度構成法 (MDS) のように位置関係で表現する。

3.1 対象データ

イベント全体の集合を G で表す。各イベントはタイムスタンプと属性値を持つとする。イベント $e \in G$ のタイムスタンプと属性値をそれぞれ $t(e)$ と $a(e)$ で表す。ある同じ属性値を持つイベントの集合を「イベント群」と呼ぶ。イベント群 E は G の部分集合である。

イベント群 E に含まれるイベントのうちタイムスタンプがある時間帯 R に含まれるイベントからなる部分集合を E_R で表すことにする。つまりイベント群 E_R は式 (1) のように定義される。

$$E_R = \{e \in E | t(e) \in R\} \quad (1)$$

単位時間ごとのイベントの頻度からなる組を「時刻分布」と呼ぶことにする。たとえば24時間における1時間ごとの時刻分布は式 (2) のように24組 (24次元ベクトル) で表される。ただし、 $b_i = |E_{[i, i+1)}|$ (i 時から $i+1$ 時までの間のタイムスタンプを持つイベントの数) とする。

$$\mathbf{b} = (b_0, b_1, \dots, b_{23}) \quad (2)$$

以下、本論文では24時間周期の時刻を考えることにする。そのため、任意のイベント $e \in G$ において、 $0 \leq t(e) < 24$ とする。

3.2 グリフを用いた時刻付きデータの表現

グリフとは多変量データを比較的小さな図形によって表す視覚的表現である [13]。各変量の値が基本図形の面積、長さ、色などで表現され、それら構成要素を統合したひとまとまりの図形として描かれる。グリフは small multiples の要素や散布図の点として用いられることも多く、小領域で表現できるように座標軸の目盛りやラベルを持たないことも多い。分析者はグリフの概観から多変量の大きな特徴を把握し、詳細な値は別の手段で得る。

グリフを用いることで、分析者は多くのイベント群の時間的な特徴を一度に把握することが可能となる。我々はイベント群を描画するために、イベント群の時刻分布 \mathbf{b} をグリフを用いて表現する。時刻分布をグリフで表現することで、我々はイベント群の時刻分布の特徴を把握できるとともに、おおまかな比較も可能になる。

グリフは、データに対する分析者の知識を活用して、データの直感的な理解を促すことが重要である [13]。加えて、データの概念的なイメージを想起させるグリフが効果的であることがいくつかの実験で示されている [14]。データに周期性があれば、アナログ時計に見られるような円周や螺旋といったイメージが効果的である。以上を踏まえて、

我々は時刻分布を表現するグリフを設計した。

時刻分布 \mathbf{b} の各成分を小さい図形で表現する。1つの成分 $b_i (i = 0, 1, \dots, 23)$ の値を図形 s_i の面積で表し、インデックス i 、つまりイベントの発生時間帯を図形の位置および色相で表すことにした。図形の種類および配置方法には様々なバリエーションが考えられるが、今回の設計では、図形は円周上に配置するものとし、図形としては円と扇型を試みた。図 2 に設計した 2 種類のグリフの例を示す。どちらもグリフの中心からの方位と色相によってイベントの発生時間帯が読み取れる。

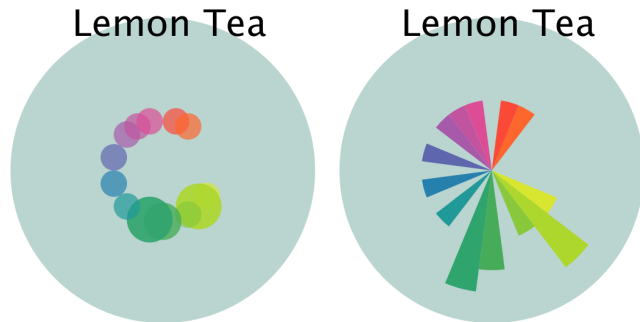


図 2 設計したグリフ。向きと色相は発生時間を意味する。個々の図形の面積は一時間間隔での発生頻度を意味する。

3.3 ChronoView による時刻付きデータの表現

ChronoView は時計の文字盤のような座標系を持ち、時間軸は円周で表現される。1つのタイムスタンプは円周上つまり時間軸上の位置で表現される。イベント群はそのイベント群に含まれるイベントのタイムスタンプの位置の重心に配置される。図 3 は ChronoView の配置規則を示したものである。図 3 では、周期は 24 時間としている。イベント群 A は 1つのイベントしか含まないため、そのイベントのタイムスタンプの位置に配置される。イベント群 B と C は、それぞれ 2つのイベントと 3つのイベントのタイムスタンプ位置の重心に配置されている。イベント群の位置がイベント群に含まれるイベントの発生時刻を大まかに表現する。ある時刻に集中してイベントが発生した場合、イベント群は円周に近い位置に配置される。一方で、発生時刻が様々な時間帯にばらついたイベントからなるイベント群は円盤の中心近くに配置される。

ChronoView はイベント群が持つ周期性を位置で表現する。そのため、ChronoView でイベント群を配置することは、分析者にイベント群の周期的な特徴の発見を促す。加えて、位置での表現は分析者が大量のイベント群を同時に観察することを可能にする。

ChronoView では 1つのイベント群が以下のように配置される。タイムスタンプの集合を T とする。1つのタイムスタンプ $t \in T$ は関数 $f_0 : T \rightarrow R^2$ によって円周上の点へ写される。つまり、1つのイベントは文字盤のような円

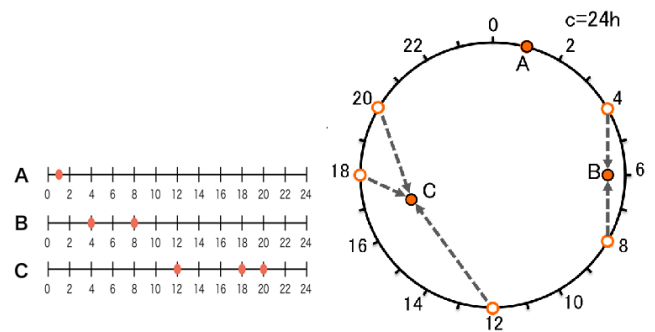


図 3 ChronoView の配置規則。

周上の位置で表される。関数 f_0 は式 (3) のように定義される。

$$f_0(t) = (r \cos \theta_t, r \sin \theta_t) \quad (3)$$

$$\theta_t = \frac{\pi}{2} - 2\pi \frac{t - t_0}{c} \quad (4)$$

ここで、 c は ChronoView が表現する周期である。また、 t_0 は周期の起点となる時刻であり、アナログ時計における 12 時の位置に写される。本論文では $c = 24$ とする。また、 t_0 は午前 0 時とする。イベント群はそれに含まれるすべてのイベントの位置の重心に配置される。イベント群の位置を与える関数 $f : 2^G \rightarrow R^2$ は式 (5) のように定義される。

$$f(E) = \frac{1}{|E|} \sum_{e \in E} f_0(t(e)) \quad (5)$$

3.4 多次元尺度構成法による時刻付きデータの表現

多次元尺度構成法 (MDS) は、多次元データをより低次元で表現する次元削減手法の一種である。MDS はデータ項目間の (非) 類似度が与えられているとき、低次元空間内の項目の配置にその類似度を反映する。二つの項目が類似している場合、それらは近くに配置される。二つの項目が似ていなければそれらは遠く離されて配置される。各項目が n 次元データに関連付けられている場合、二つの項目の (非) 類似度はしばしば n 次元空間上のユークリッド距離やコサイン類似度などで与えられる。

MDS を用いることで、類似の時刻分布を持つイベント群を近くに、そうでないイベント群を遠くに配置することができる。グリフでも個々のイベント群の特徴を読み取ることができるが、類似か否かについては、グリフの特徴を見比べる必要がある。グリフの図形としての類似性を判断するよりも、配置における近さの方が直感的に読み取り易い。

我々は各イベント群の時刻分布を表す 24 次元ベクトル \mathbf{b} を 1つのデータ項目として、それらの類似性を表現空間内の配置によって表現することにした。

4. 3D 空間上における 3つの表現の統合

我々は個々のグリフによって 1つのイベント群の時間的特徴を表現し、グリフの配置によってイベント群の周期的

特徴とイベント群間の類似性を表現する。グリフの配置を決めるために ChronoView と擬似的な MDS を統合した。なお、類似度を位置に反映させるアイデアは MDS に起因しているが、MDS を使用しているわけではないため「擬似 MDS」と呼ぶことにする。

4.1 ChronoView と擬似 MDS の統合

我々は 2 次元平面上で ChronoView による配置を保存しつつイベント群間の類似性を表現することを試みた [15]。しかしながら、ChronoView と MDS の統合には空間を歪める必要があり、視覚的表現としては無理があることが分かった。そのため、3 次元空間内でこれらを統合することにした。

まず、イベント群の xy 座標を ChronoView で求め、擬似 MDS により 3 次元空間内でのユークリッド距離にそれらの時刻分布の類似性を反映させる。つまり、類似性の表現には z 座標だけを利用する。イベント群の z 座標を求めるために、グラフ描画に利用されるバネモデル [16] を用いる。イベント群がノードであり、どの二つのノードもイベント群の時刻分布の類似度に比例した自然長 d_0 をもつバネで隣接しているとす。バネにより各ノードは式 (6) で与えられる力を受けるとする。安定状態を求めるにあたり、ノードは z 軸方向だけにしか動かないものとする。

$$f_s = s \log \frac{d}{d_0} \quad (6)$$

4.2 3D 空間上でのグリフの表現

グリフの図形要素は、常にイベント群の時間的特徴を観察できるよう、分析者がどの視点から 3D 空間を見ても視認できる必要がある。しかし 2D 表現のグリフを 3D 空間内で描画すると、グリフの形状が傾いて見え、適切に時間的特徴が視認されない恐れがある。そのため、我々はカメラの角度とは別に、個々のグリフの角度を固定して描画することにした。図 4 では座標系が傾いて見えているが、グリフはすべて画面に対して正面を向いて描かれている。なお、グリフ背景の円の面積はイベント群に含まれるイベントの数に対応づけた。すなわち、イベント群 E に関して、円の面積は $|E|$ に比例する。

イベント群間の類似関係をグリフでも把握できるよう、グリフ背景の円に色を割り当てた。グリフの背景の円の色は z 座標の値に対応しており、色の範囲は赤色から青色とした。

3D 空間の表現において、見かけの大きさと位置が問題となる。投影線が 1 点に集中する透視投影では画面手前に描画された図形は大きく見え、画面中央から外へ広がった位置に描画される。これは ChronoView が位置によって周期的特性を表現する上で弊害となる。そのため我々は投影

線を並行に求める並行投影で画面を描画することにした。

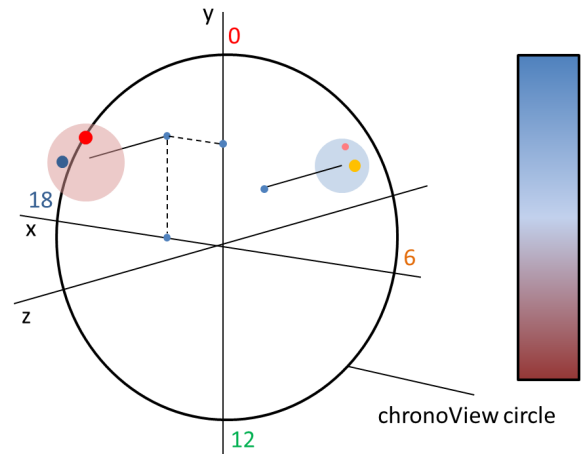


図 4 3D 空間上のグリフ

5. 分析ツールの開発

我々は提案手法を備えた時刻データの分析ツールを開発した。本節ではこのツールの詳細について述べる。

5.1 開発ツールの概観

図 1 は開発したツールの概観を示している。ツールは大きく分けて以下の項目から構成される。

- メインビュー：開発した概観を表示、操作するためのビュー。
- 属性選択リスト：イベントの属性 (例：商品名、ユーザ ID) が複数ある場合、分析者が切り替える事が可能。
- 詳細タイムチャート：分析者は興味あるイベント群を選択し、より詳細なイベントの発生パターンを分析できる。
- 類似イベント群リスト：分析者が選択したイベント群と類似するイベント群の、詳細なイベントの発生パターンを表示するリスト。

このツールは大量のイベント群を観察するための機能と、特定のイベント群の詳細な分析を行うための機能を両方備えている。メインビューではグリフが 3D 空間上に配置されている。分析者はメインビューを用いて 3D 表現空間を観察する。我々は Carlis らの手法 [1] を参考にして詳細タイムチャート (図 1(3)) を開発した。このチャートではイベント群に含まれる個々のイベントが螺旋状の時間軸上に配置される。このチャートはより詳細な周期的特性 (例：どの時期に多く発生したか) を示す。イベントは 1 時間ごとに集約され、円の面積はその頻度に比例する。背景の扇形の図形の色は各時間帯で発生したイベントの数を示している。

5.2 グリフの集約表現

表示するイベント群の数が多くなるにつれ、グリフの重なりが煩雑さをもたらす。煩雑さを回避する1つの方法は複数のグリフを1つにまとめてしまうことである。そこで我々は、イベント群をクラスタ化して時刻分布が類似している複数のイベント群を1つのグリフに統合する機能を用意した。図5はクラスタ化前後の様子を示している。ChronoViewと擬似MDSを1つの視覚的表現に統合したことにより、イベント群間の距離に類似性が反映されているため、互いに距離が近いグリフの集まりが1つのグリフに統合されたように見える。

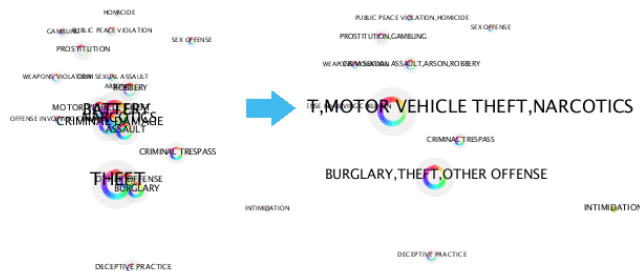


図5 The example of aggregation

5.3 類似度リンク

擬似MDSにより、イベント群間の類似度は、3次元空間内における対応するグリフ間のユークリッド距離に反映されているが、MDSによる次元削減と同様に距離で類似度が厳密に表現されているわけではない。分析者がイベント群間の類似度をより正確に把握できるように、イベント群間の類似度を示すリンクを描画する機能を用意した。リンクの太さと色で類似度を表現しており、類似度が高いリンクほど太くそして明るい黄色となる。逆に類似度が低いほどリンクは細く、暗い色となる。分析者があるイベント群を選択すると、選択されたイベント群と他のイベント群との間にリンクが描かれる。

6. ユースケース

本節では、時刻データの分析における提案手法の有効性を説明する。分析対象のデータとして購買データを用いる。この購買データは購入商品（属性）と購入時刻（タイムスタンプ）を持つ。分析者の目的は購買データを分析し、大量のイベント群の中から購入時刻に特徴がある商品を発見することである。

6.1 発生時間の観察

最初に分析者は、どの時間帯に商品が購入されているのかを調べることにした。分析者はイベント群の周期的特性が表現されたChronoViewのレイアウト(図7(a))を観察した。分析者はそれからx-y平面上のグリフの配置を観察

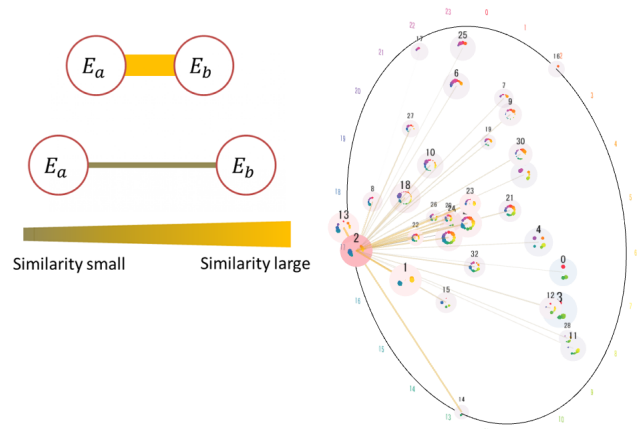


図6 「2」イベント群と他イベント群の間に類似度リンクが描画されている。より正確な類似関係の概観が得られる。

した。図7(a)では多くのグリフが、ChronoViewの円盤の中心から見て左下や左上に配置されている。円盤の中心よりも左下にグリフが配置されるというのは、イベント群に含まれるイベントの多くが昼の時間帯に発生したことを意味する。同様に、円盤の中心よりも左上に配置されるというのは多くのイベントが夜に発生したことを意味する。以上の特性を踏まえ、分析者はグリフの配置から、12時～16時の時間帯と18時から夜中にかけて購入された商品が多いという情報を得た。

購入商品の傾向として、Lemon Tea, Milk Tea, Green Teaが円盤の中央から下に位置しており、飲み物が昼に購入されていると見る事が出来る。これはグリフに昼の時間帯を示す緑色の円が大きく描画されていることから分かる。Pieは円周(22時)近くに位置しており、22時あたりに集中して購入されていると見る事が出来る。

円盤の中心付近では、分析者はグリフの位置からイベントの時刻分布を読み取ることが難しくなる。だが分析者はグリフを観察することで時刻分布の読み取りを補うことができる。例えば「Vitamin Drink」のグリフでは、中心から見たどの方位にも大きく円が描かれており、これを見ることで時間帯でも満遍なく商品が購入されていることが分かる。

6.2 類似性の観察への遷移

図7(a)では多くのグリフ背景の円の色が赤あるいは紫である。だが逆に青色の円を持つグリフは少ない。イベント群間の類似性は背景の円の色で類似性でも表されている分析者は画面を俯瞰して背景の円の色が青い「Water」(図8(左))に興味を持ち、「Water」が他と比べ変わった時刻に購入されていると考えた。分析者は発生時刻の類似性を調べるため、周期的特性が見える画面からイベント群間の類似性が見える画面へと移動した。ChronoViewとMDSを1つの表現へと統合しているため、分析者は見たい特徴に

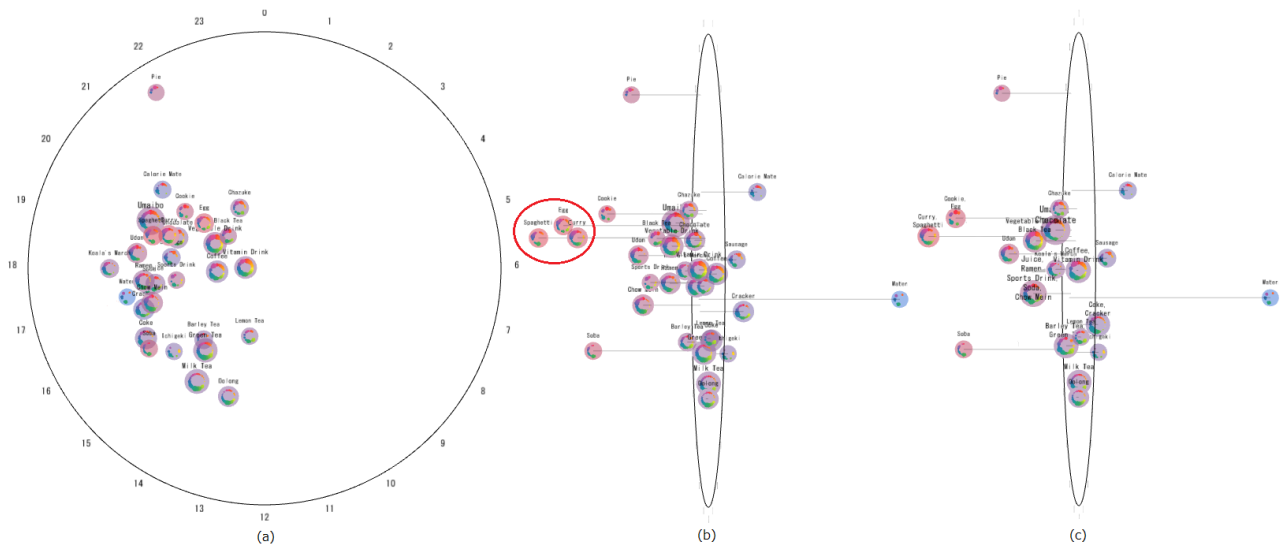


図 7 3D 空間上で商品の購入時刻の特徴が表現されている。(a) イベント群の発生時刻の特徴が表現されたレイアウト。グリフの位置がイベント群の大きな周期的特性を示す。(b) 類似性を観察するために y 軸に対して回転を行ったレイアウト。グリフの背景の色とグリフの位置が特徴的なグループの存在を示す。(c) 密集していたイベント群をクラスタ化した結果。

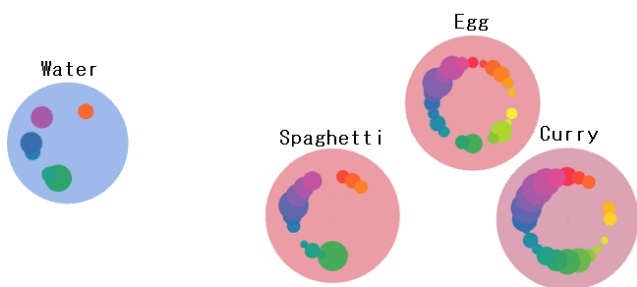


図 8 グリフ背景の円の色が青い「Water」(左)。類似した発生時刻の特徴を持つ 3つのイベント群 (右)

応じ滑らかに画面を遷移させることができる。

6.3 類似性の観察

図 7(b) は 3 次元空間の z-y 平面を示している。ChronoView による配置を維持しながら、イベント群間の類似性を 3 次元空間上の距離で表すために、我々は z 軸の 1 軸方向の自由度だけを利用している。そのため、イベント群間の類似性は、z 座標が見える z-x 平面あるいは z-y 平面の方がより把握しやすい。

分析者は概観を観察して確かに「Water」が他イベント群から離れていることを発見した。イベント群が他の多くのイベント群から離れていることは、そのイベント群の発生時刻が特徴的であることを意味する。つまり「Water」は他の商品と比べ購入される時刻に特徴があることがわかる。更に、「Egg」、「Spaghetti」「Curry」(図 7(b) 赤線)らが隣接しており、それらが他イベント群から離れていることを発見した。これは「Egg」、「Spaghetti」「Curry」が似た時刻に購入されており、かつ発生時刻に特徴がある事を

意味する

図 8 を見ると、これら 3つのイベント群のグリフ上には 12 時、19 時、2 時あたりに共通して円が描画されている。これはこれらの商品に共通する購入時刻の特徴を示す。以上より、類似性が反映された画面に滑らかに変化したことで、分析者は二つの特徴的な時刻分布を持つ商品グループを見つける事が出来た。

6.4 インタラクションの活用

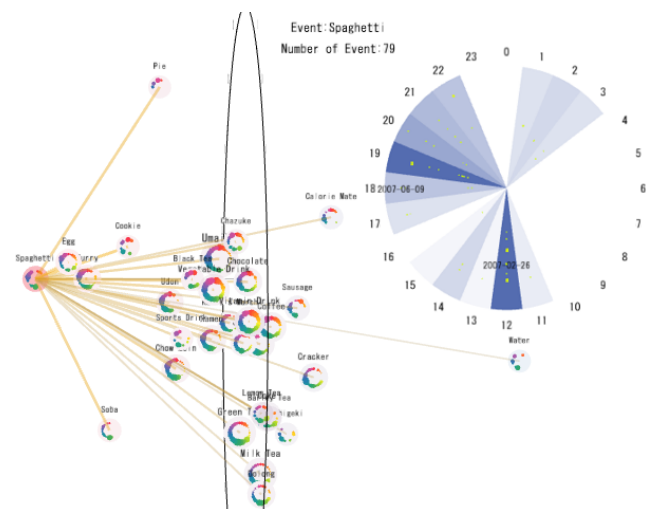


図 9 「Spaghetti」と他イベント群にリンクを描画し、「Spaghetti」の詳細タイムプロットを表示した画面。

本節では類似度リンクとグリフ集約の活用例を示す。分析者は図 7(b)において、発生時刻に特徴があると思われた「Spaghetti」に着目した。このとき「Spaghetti」とほかの

イベント群との間にリンクが描画される。このリンクの太さ、明るさが類似性を意味する。図9では「Spaghetti」と、「Curry」「Egg」「Cookie」との間に太いリンクが見える。逆に、距離が離れている「Water」「Cracker」「Calorie Mate」との間には細いリンクが見える。太いリンクであるほど類似性が高いことを意味するため、分析者は「Spaghetti」と購入時刻の特徴が類似する商品が何であるかをより正確に把握できた。詳細タイムプロット(図9(右))を見ると、12時と19時を示す扇形の図形の上に円がいくつかプロットされている。これは「Spaghetti」が主に12時と19時に周期的に購入されていたことを意味している。

イベント群の数が多い場合、分析者はクラスタ化を行うことで視覚的な煩雑さを排除することができる。図7(c)はクラスタ化を行った結果を示しており、図7(b)と比べてグリフの数が減っている。特に $z=0$ 付近に位置するイベント群には類似するイベント群が多く、それらが1つのグリフにまとめられたことを示している。

7. まとめ

我々は時刻付きデータの分析を支援することを目的として、大量のイベント群の時間的特徴を観察できる概観を開発した。我々はこの概観を3つの手法、グリフ、ChronoView, MDSを1つの視覚的表現へ統合することで構築した。この概観の重要な特性は、大量のイベント群が小面積で表現されること、イベント群の周期的特性およびその類似性が同時に表現されることである。分析者は大量のイベント群の中から特徴的なイベント群を発見し、他の手法を用いたより詳細な分析へ繋げる事ができる。我々は開発した概観のデータ分析における有効性をユースケースを通して説明した。

謝辞

本研究はJSPS 科研費 JP17K00264 の助成を受けたものです。

参考文献

- [1] John V. Carlis and Joseph A. Konstan. Interactive visualization of serial periodic data. *in Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 29–38, 1998.
- [2] S. Shiroy, K. Misue, and J. Tanaka. Chronoview: Visualization technique for many temporal data. *in Proceedings of 16th International Conference on Information Visualisation*, pp. 112–117, 2012.
- [3] J. B. Kruskal and M. Wish. *Multidimensional Scaling*. Sage, London, 1978.
- [4] W. Javed, B. McDonnell, and N. Elmqvist. Graphical perception of multiple time series. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 6, pp. 927–934, 2010.
- [5] S. Havre, B. Hetzler, and L. Nowell. Themeriver: Visualizing theme changes over time. *in Proceedings of IEEE*

- Symposium on Information Visualization*, pp. 115–123, 2000.
- [6] D. Luo, J. Yang, M. Krstajic, W. Ribarsky, and D. Keim. Eventriver: Visually exploring text collections with temporal references. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 1, pp. 93–105, 2012.
- [7] M. Imoto and T. Itoh. A 3d visualization technique for large scale time-varying data. *in Proceedings of 14th International Conference on Information Visualisation*, pp. 17–22, 2010.
- [8] M. Weber, M. Alexa, and W. Müller. Visualizing time-series on spirals. *in Proceedings of the IEEE Symposium on Information Visualization*, pp. 7–14, 2001.
- [9] C. Tominski and H. Schumann. Enhanced interactive spiral display. *in Proceedings of the Annual SIGRAD Conference – Special Theme: Interactivity*, No. 34, pp. 53–56, 2008.
- [10] 石井貴大, 三末和男. データの周期的特性の探索を目的とした chronoview の 2.5d 拡張. 情報処理学会ヒューマンコンピュータインタラクション研究会研究報告, Vol. 2018-HCI-177, No. 9, pp. 1–8, 2018.
- [11] D. Q. Nguyen, C. Tominski, H. Schumann, and T. A. Ta. Visualizing tags with spatiotemporal references. *in Proceedings of the International Conference on Information Visualisation*, pp. 32–39, 2011.
- [12] F. Fischer, J. Fuchs, and F. Mansmann. Clockmap: Enhancing circular treemaps with temporal glyphs for time-series data. *in Proceedings of EuroVis Short Papers, Eurographics*, pp. 97–101, 2012.
- [13] M. O. Ward. Multivariate data glyphs: Principles and practice. *In Handbook of Data Visualization*, pp. 179–198, Berlin Heidelberg, 2008. Springer.
- [14] J. Fuchs, P. Isenberg, A. Bezerianos, and D. Keim. A systematic review of experimental studies on data glyphs. *IEEE Transactions on Visualization and Computer Graphics, Institute of Electrical and Electronics Engineers*, Vol. 23, No. 7, pp. 1863–1879, 2017.
- [15] 安齋康宏, 三末和男. イベントデータ可視化のためのグリフ配置手法の開発. 第22回情報処理学会シンポジウム, インタラクション 2018, pp. 1054–1059, 2018.
- [16] P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, Vol. 42, pp. 149–160, 1984.