

GSLetterNeo vol.127

2019年2月

ビジュアルコンポーネントをもちいて計算方法を見比べる

松原 伸人 matubara@sra.co.jp

はじめに

GSLetterNeo Vol.126 「距離が近いデータをグループにして表示する」では、軸上に置いたデータを距離が近いデータをグループにして表示して見る方法のプロトタイプ hakeimodoki を紹介しました。hakeimodoki は、時刻や距離などの数値データを四角形で表し、四角形をプロットした軸上で、指定した距離より近いデータをグループにして表示します。グループになったデータは、縦に積み上げて表示しました。表示して見てみることで、距離が近いデータがわかり、数が数えられるようになり、他の軸と範囲が重なり合う箇所が分かるようになりました。「グループ化」の節では、グループ化の方法を次のように書きました。

"グループ化は、データ間の距離を全て求めたデータ列を作成し、データ列の中からデータ間の距離が短いデータから順に、グループ化の距離より短いデータをペアにします。ペアにしたデータの平均時刻をグループの時刻とします。見つけたペアをデータ列に加え、ペアになったデータをとりのぞいたデータ列を作成し、このデータ列から再びペアを見つける処理を行い、ペアが見つからなくなるまでこれを繰り返します。"

今回は、同じデータに、別のグループ化方法を適用した結果を hakeimodoki で表示して見てみます。

本内容は、科学技術振興機構(JST) 戦略的創造研究推進事業(CREST) 「データ粒子化による高速高精度な次世代マイニング技術の創出」プロジェクト (研究代表者: 宇野毅明 教授 (国立情報学研究所)) で行なっている研究開発の一部です。今回使用している画像は、筆者自身と、共同研究者の山本恭裕 特定教授 (京都大学)、およびKTL所長でもある中小路久美代 特定教授 (京都大学)の協力により得られた研究データの一部です。両者の同意のもと本誌に掲載させていただきました。

計算方法を比べる

Vol.126のグループ化は、距離が近いデータをペアにして徐々に大きなグループを作っていくボトムアップ手法の1つだと思います。これと逆に、データ全体を1つのグループと捉えて、その中から距離が離れている場所で2つのグループに分ける処理を繰り返していくトップダウン手法も考えられます。いずれの場合でも、条件に合うデータから木構造を作っていく処理と考えられます。

今回用いているデータは、Vol.126などで用いた函館公園内を2018年10月9日に歩いて撮影した画像の撮影時刻です。そのうち1名のデータに対して、4つのグループ化方法 **center, left, right, justify** を適用した結果を表示してみました。この4つの方法は、ともにボトムアップに距離を計算してペアを作ってグループにしています。4つの方法それぞれ、ペアにした際の撮影時刻の決め方と、グループ間の距離の計算方法が異なります。

グループの見た目は、グループ化の段階を色で表し、グループのデータの個数を上下の高さ、グループの範囲とグループの時刻から描かれるベジエ曲線で表しています。

center は、Vol.126で紹介したグループ化方法です。ペアにした画像の時刻は、2つのデータの撮影時刻の平均時刻とします。データ同士の距離は撮影時刻から求めます。この方法でグループを作成していくと、四角形が積み上がる位置が真ん中に寄っていきます。[図1]

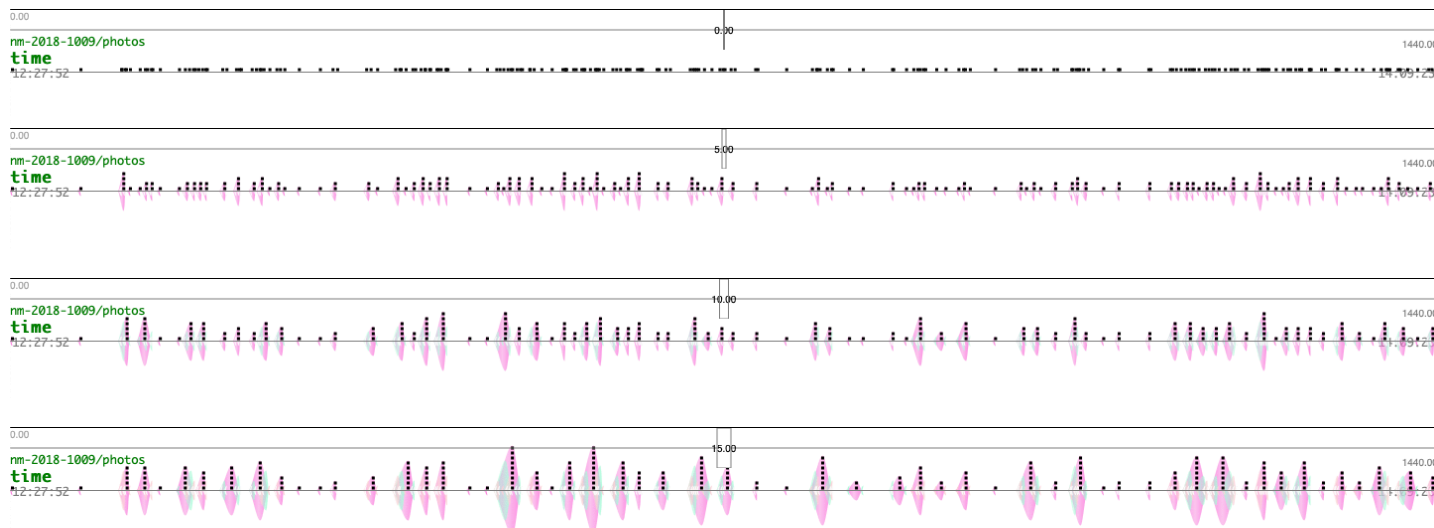


図1 グループ化 center 距離0、距離5、距離10、距離15

left は、ペアにした画像の時刻を2つデータの撮影時刻のうち小さいほうにします。データ間の距離は撮影時刻の差の絶対値です。この方法でグループを作成していくと、四角形が積み上がる位置が左に寄っていきます。[図2]

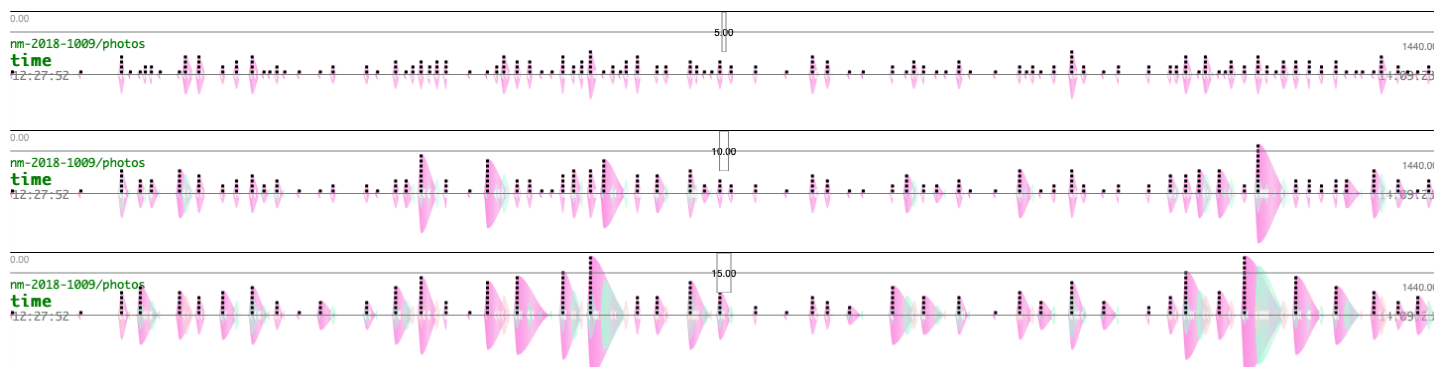


図2 グループ化 left 距離5、距離10、距離15

right は、left と逆に、ペアにした画像の時刻を2つのデータの撮影時刻のうち大きい方にします。データ間の距離は撮影時刻の差の絶対値です。この方法で作成されるグループは、四角形が積み上がる位置が右に寄っていきます。[図3]

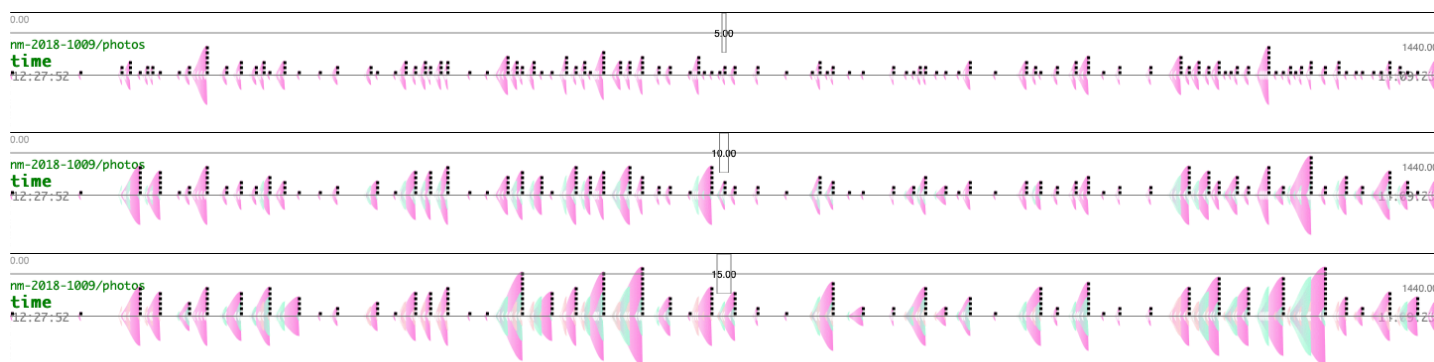


図3 グループ化 right 距離5、距離10、距離15

justify は、グループにした画像の時刻を左端のデータと右端のデータの時刻の平均にします。データ間の距離はグループの両端の画像の時刻の差の絶対値の内の小さいほうの値にします。この方法で作成されるグループは、上記の3つの方法と比べて、グループ間の距離が短くなるため早い段階で大きなグループができていきます。[図4]

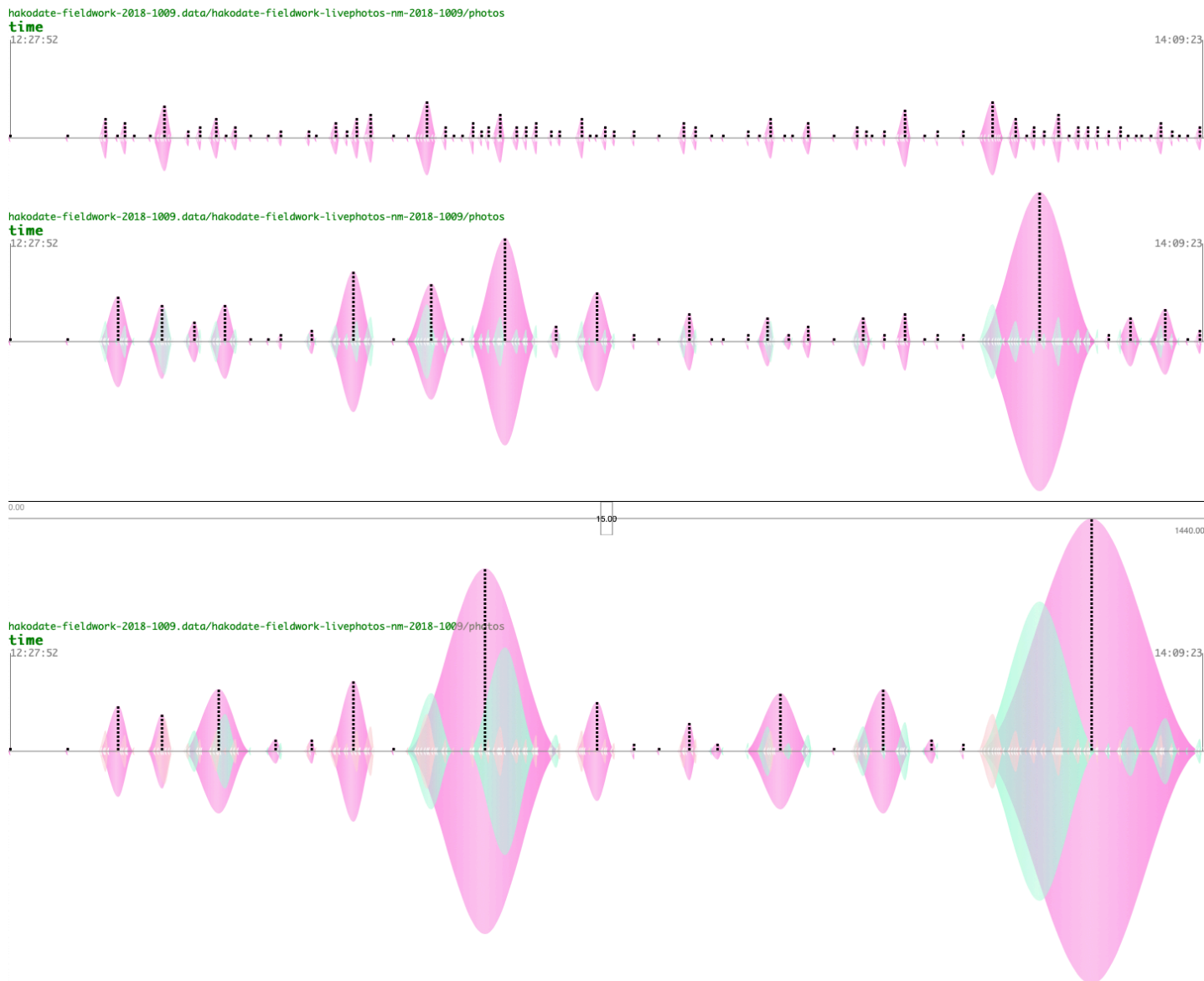


図4 グループ化 justify 距離5、距離10、距離15

グループ化の処理方法を変えて見ると結果が違ってくることがわかりました。計算方法にかかわらずに、適用できるようにビジュアル表現をプログラムしておくことで、計算結果を見て比べられるようになりました。

GSLetterNeo vol.127

発行日 2019年2月20日

発行者 株式会社 S R A 先端技術研究所

編集者 土屋 正人

バックナンバー <https://www.sra.co.jp/gslletter/>

お問い合わせ

gsneo@sra.co.jp

〒171-8513 東京都豊島区南池袋2-32-8

