

Performance comparison of hash-based methods for similar waveform detection

*Yuki Takayama¹, Shiro Hirano¹, Hironori Kawakata¹

1. Ritsumeikan University

はじめに

これまで、Template matchingを用いて相似地震を検出する研究が多く行われてきた。相似地震とは、類似の波形を持つイベントとして定義される(Igarashi, 2020)。ただし、Template matchingでは、既知の波形の類似波形のみを見つけることしかできない。既知の波形をテンプレート波形として固定しない類似波形の検出を行うことで、Template matchingでは見つけ出せなかった新たな類似波形のペアを見つけ出すことが可能となる。そのような検出手法には、Auto correlation (AC)がある。この手法は連続波形をN個の波形ウィンドウに分割し、それらすべての組み合わせにおいて相関係数を計算するというものである。しかし、ACは解析する期間が長いと計算時間が $O(N^2)$ に比例して、非常に長くなるという欠点がある。そこで、Yoon et al. (2015)は、画像検索や音声検索、Web検索エンジン等で利用されているハッシュ法を用いたFingerprint and Similarity Thresholding (FAST)という手法を開発した。ハッシュ法とは、ハッシュ関数を文字データや画像データなどの任意のデジタルデータに作用させることで、より簡易なデータ(ハッシュ値)に変換し、ハッシュ値をデータの要約として、データ同士の比較を高速化するという手法である。ここでハッシュ関数とは、個々のデジタルデータに何らかのビット列を割り当てる規約であり、類似性検索の場合には高速化のために元データよりも短いビット列を割り当てる。駒形ほか(2020)は、Fingerprintingよりも簡単に計算でき実行時間短縮が見込める、2bit-ahashとkhashの2つを提案した。

一方、ほぼ同じ場所で発生し、ほぼ同一の発震機構を持つ類似の地震を繰り返し地震と呼ぶ(Igarashi, 2020)。しかし、前段落のようにして検出される相似地震のうちどの程度が繰り返し地震であるかについては、検出される波形をもたらすイベントの小ささ故、詳しく調べられてこなかった。Uchide(2020)は、深層学習を利用して日本列島下で発生した小規模地震の初動メカニズム解を推定した。小さいイベントは、メカニズム解が決定できるとは限らないため、まずはメカニズムまで含めて既知のイベントについて類似度指標とメカニズム解の相関を見ることが必要となる。ノイズ(常時微動)とノイズやノイズとイベントは相関の低いものであり、余分な計算となってしまう。駒形ほか(2020)によりkhashならノイズは高確率で除去できることが確かめられているため、イベントとイベントについてのみメカニズム解と類似度の相関を考えることになる。

そこで、本研究では、駒形ほか(2020)が提案した2bit-ahashとkhashに着目し、この2つと相互相関係数を使用して波形の類似度を計算し、波形の類似度だけでなく、発震機構も使用することで、繰り返し地震に対して、どの程度有効となるかを1つの指標として、3つの類似度指標の比較を行うことを目的とする。

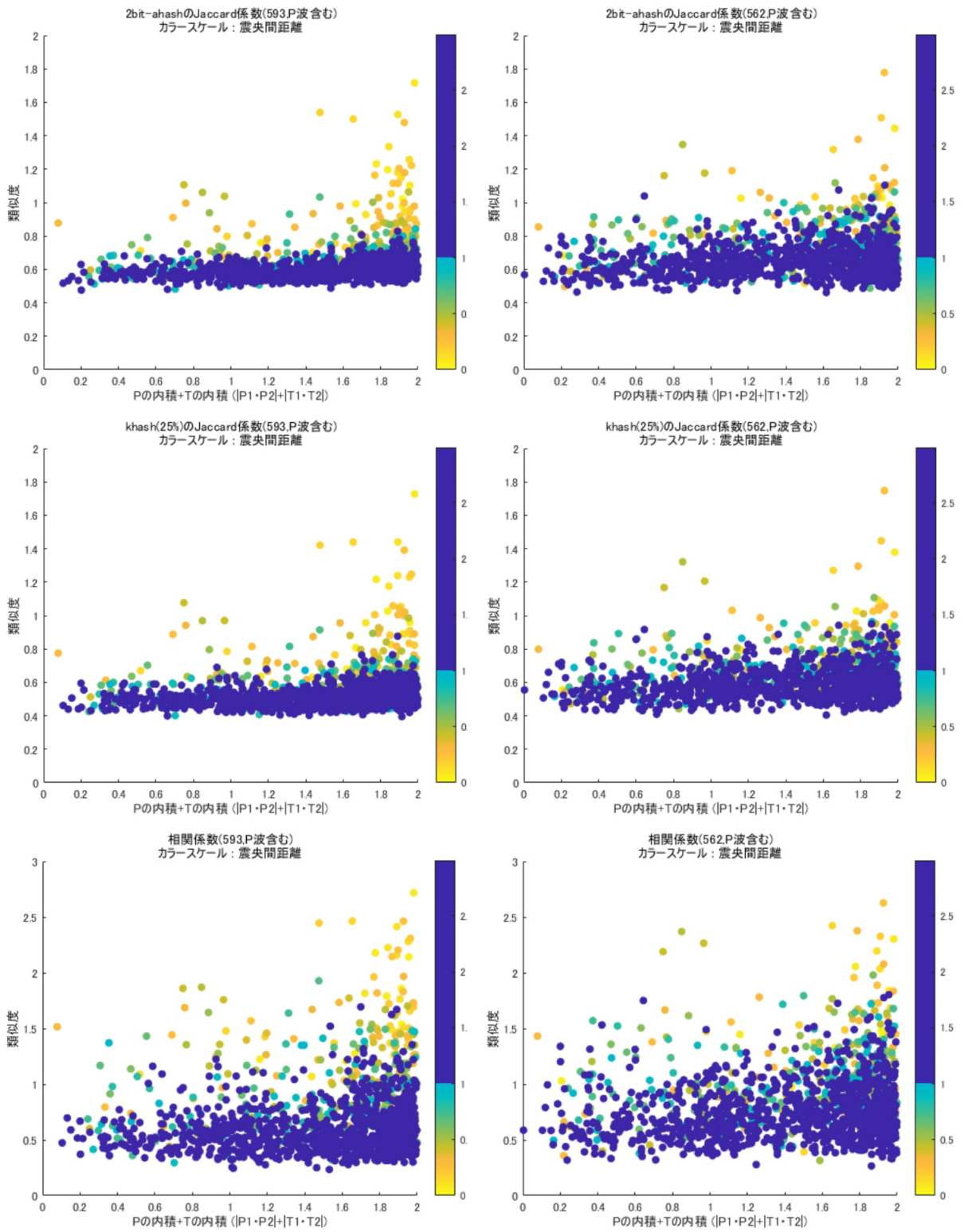
解析手法

解析データとして、2018年6月18日7:58に大阪府北部で発生したMj6.1の地震直後から16時間のHi-net交野、久御山観測点で記録された連続波形記録の3成分を用いた。サンプリング周波数は100Hz、周波数帯域は5~30Hzを使用した。まず、非定常的なイベントをピックし、それらについて3つの類似度指標それぞれで観測点ごと総当たりにイベント同士の類似度を計算した。その後、Uchide(2020)で初動解が決定されているイベントのうち信頼性のたかいもののみについてUchide(2020)の初動解からP軸、T軸、N軸を計算した。イベント同士のP軸の内積とT軸の内積の絶対値の和を横軸に、イベント同士の類似度を縦軸に、イベント同士の震央間距離をカラースケールでとり、図を作成した。ここで、深さについては、決定精度が低い可能性があるため、無視した。

結果

図より、2bit-ahashとkhashは相互相関係数に比べて、イベント同士の震央間距離が遠いペアの類似度は相対的に低くなり、震央間距離が近いペアの類似度が相対的に高くなり、その境界が明確になることが明らかに

なった。このことから、類似度閾値の設定を容易にすることが期待できる。また、相互相関係数よりも2bit-ahashやkhashの方が、震央間距離が遠いイベントについては、メカニズムが似ていたとしても、類似度が高くないことが期待できる。また、解析の結果、3つの類似度指標間の順位相関(特に相互相関係数)が必ずしも高くないことが明らかになった。波形が類似していることと、イベント間の震央間距離と震源メカニズムが似ていることが、どの程度一致するのかを知ることは、3つの類似度指標が繰り返し地震についてどの程度確証を持てるかの指標になる。今回の解析結果では、相互相関係数よりも2bit-ahashやkhashの方が繰り返し地震の判定に有効であることが考えられる。



左：交野観測点 右：久御山観測点
上：2bit-ahash 真ん中：khash 下：相互相関係数