

波形の類似性を考慮した 大規模月地震データの 可視化システムの実装

後藤康路¹, 山田竜平², 山本幸生³, 横山昌平⁴, 石川博⁴

¹ 静岡大学大学院 情報学研究科

² 国立天文台

³ ISAS/JAXA

⁴ 静岡大学 情報学部

背景

- * ALSEP(Apollo Lunar Surface Experiment Package)
 - * 地震計を始めとする実験装置群の設置



- * 月地震の観測記録
 - * 得られる知見
 - * 月地震の種類, 活動度
 - * 月の内部構造
- * Webを介したデータ公開
[Yamada et.al. 2011]

背景

- ※ 月地震の分類
 - ※ 発生原因, 震源等
 - ※ 人手による分類が主
 - ※ 未だに未分類とされている月地震が存在

- ※ 分類例
 - ※ 深発月震
 - ※ 同一震源で定期的に発生
 - ※ 震源で分類
 - ※ 浅発月震
 - ※ 隕石衝突

背景: 課題

※ 月地震データの特徴

- ※ 約7年半にわたる膨大な量の地震記録(約545[GiB])

人手による解析が
困難

- ※ 観測環境に起因するノイズが多く含まれている
- ※ 厳密な正解データの入手が困難
- ※ 未知の種類 of 月地震がある可能性

機械的な分類が
困難

目的

月地震の分類

人手による分類結果を教師データとした機械的な分類は困難

人手による分類結果

教師無し機械学習の結果

可視化

- ✳ SOM(Self-Organizing Map)を用い2次元空間に可視化
- ✳ 大規模データを円滑に処理するため、
処理のバックエンドにHadoop+Hiveを利用

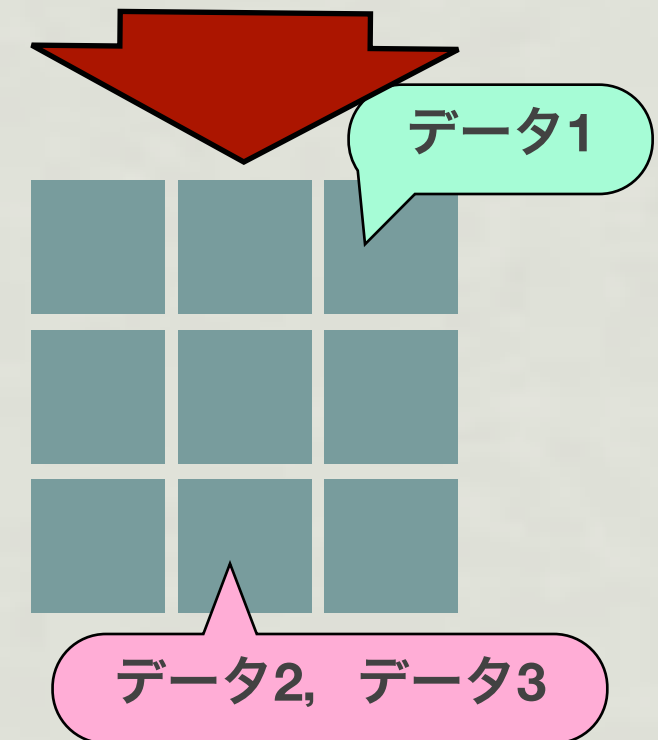
手法

Self-Organizing Map

- * 教師無し機械学習手法のひとつ
- * 可視化手法として利用可能
- * 高次元の入力データを低次元の出力空間マッピング
- * データ同士の類似性は出力空間でも維持

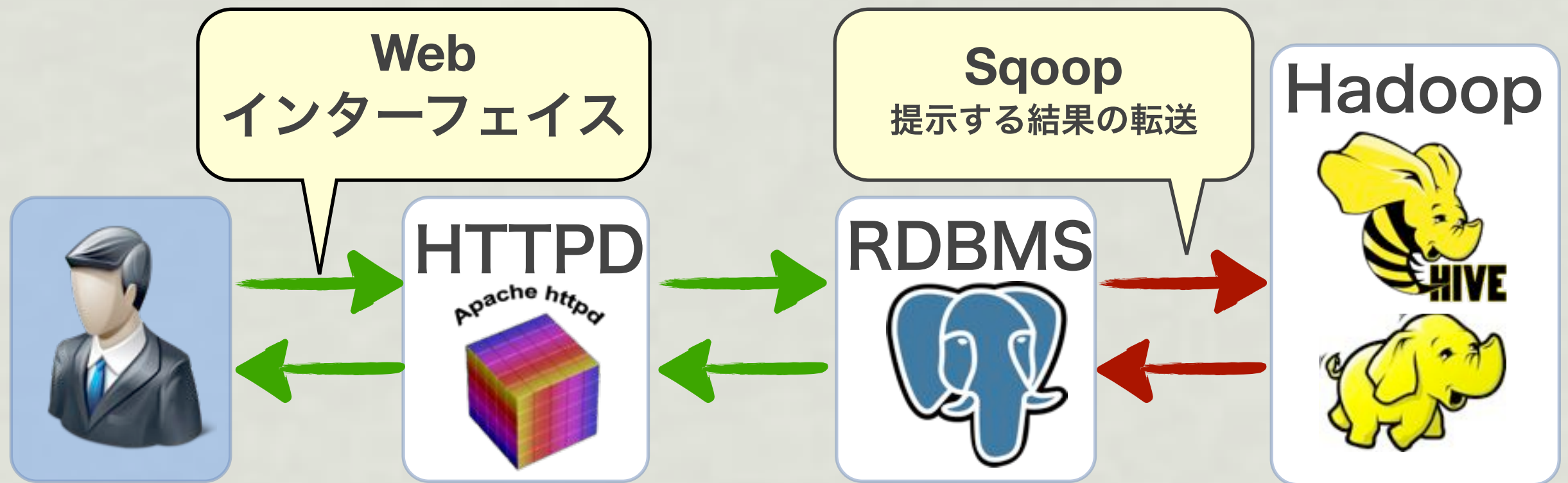
	属性1	属性2	...
データ1	0.1	0.2	...
データ2	0.8	0.5	...
データ3	0.8	0.6	...
:	:	:	

入力データ(高次元データ)



出力空間(低次元空間)

提案システム



- * Hadoop+Hive
 - * 全データセットの管理 (波形データ, イベントデータ, 処理結果)
 - * 波形データ処理
 - * Self-Organizing Mapの処理
- * PostgreSQL
 - * SOMの結果等のWebインターフェイスで用いるデータの管理

提案システム

データ管理

- ＊ 月地震データ

- ＊ DARTS at ISAS/JAXAより取得

- ＊ 容量: 545[GiB] (1,380,000,000レコード)

- ＊ データの要素

- ＊ 観測時刻

- ＊ 長周期地震計測位データ(X, Y, Z軸)

- ＊ 短周期地震計測位データ(Z軸) 等

大規模データ処理

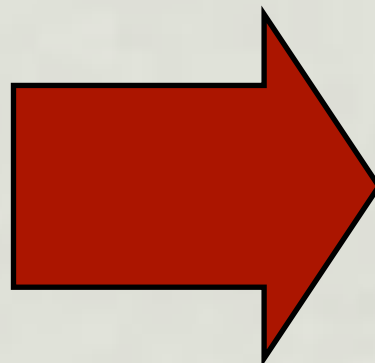
多数の計算機を用いて
負荷を分散させることで
効率よく処理

提案システム

データ管理

- ※ Hadoop
 - ※ 大規模データ処理基盤
 - ※ HDFS: 分散ファイルシステム
 - ※ MapReduce: 分散処理フレームワーク
- ※ Hive
 - ※ Hadoop操作系
 - ※ SQLライクにHadoop上のファイルをテーブルとして処理を行う
 - ※ ユーザ定義関数によって複雑な処理が可能

Hadoop + Hive
によるデータ管理



MapReduce
によるデータ処理

提案システム

データ管理

- ※ Hiveにおける大規模時系列データの管理
- ※ Snappy(高速圧縮アルゴリズム)の適用
 - ※ MapReduce処理のデータ通信オーバーヘッド対策
- ※ RCFile(列指向ファイル形式)の適用
 - ※ 時系列データアクセスの効率化
 - ※ 圧縮アルゴリズムの効果向上
 - ※ テーブル上の列において似た傾向の並びが発生しやすいデータに有利

X軸	Y軸	Z軸
X ₁	Y ₁	Z ₁
X ₂	Y ₂	Z ₂
X ₃	Y ₃	Z ₃

行指向

X軸	Y軸	Z軸
X ₁	Y ₁	Z ₁
X ₂	Y ₂	Z ₂
X ₃	Y ₃	Z ₃

列指向

データサイズを約1/5に圧縮

提案システム

データ処理の流れ

- ＊ 月地震イベントの波形抽出
 - ＊ イベントカタログを参照した抽出
- ＊ 月地震イベント波形の前処理
 - ＊ 単位補正, フィルター処理
- ＊ 特徴量の計算
 - ＊ フレーム毎のPSD(Power Spectral Density)を平均
- ＊ SOM(Self-Organizing Map)
 - ＊ 高次元データを低次元にマッピングする可視化手法



地震計の計測波形

提案システム

月地震イベントの波形抽出

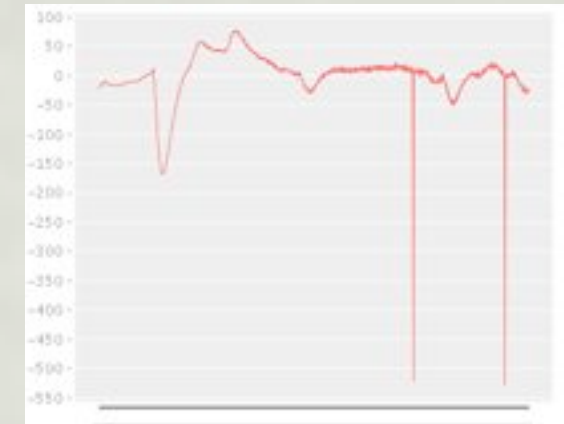
- * 月地震の発生時刻を元にイベントの波形を抽出
 - * 月地震イベントカタログ[Nakamura et.al.1992]等
 - * 発生時刻, 終了時刻
 - * 人手による月地震の分類結果 等
- * 各地震毎に,
地震の傾向が出やすい開始数分間の波形を抽出

提案システム

月地震イベント波形の前処理

1. 長周期トレンドの除去

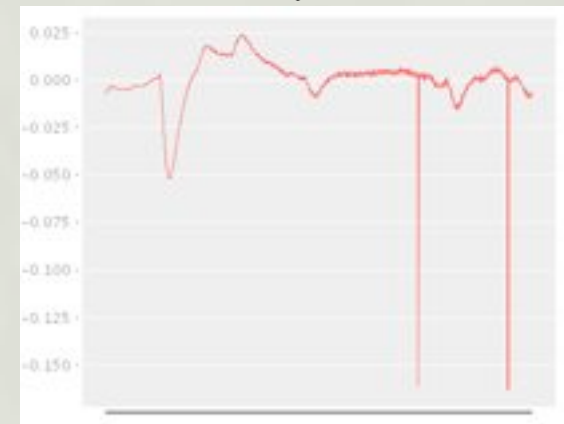
- ＊ 月地震イベントの波形に1次直線をフィットして全体の傾きを除去



1

2. 単位補正

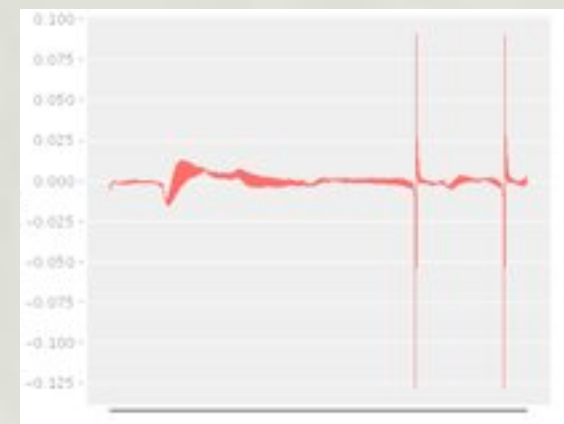
- ＊ 地震計の計測単位[Digital Unit]からメートル法への変換



2

3. フィルタ処理(Butterworth filter)

- ＊ 地震計感度に合わせて周波数領域を抽出

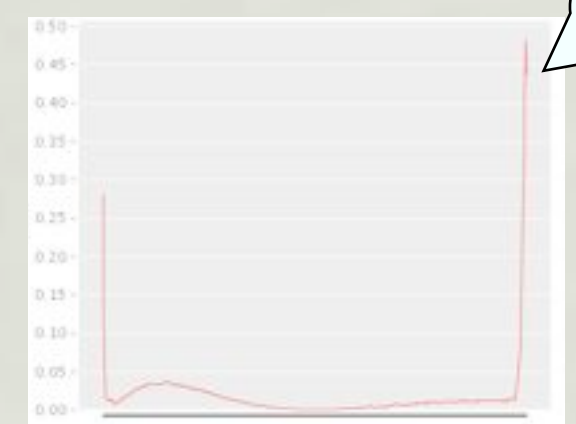
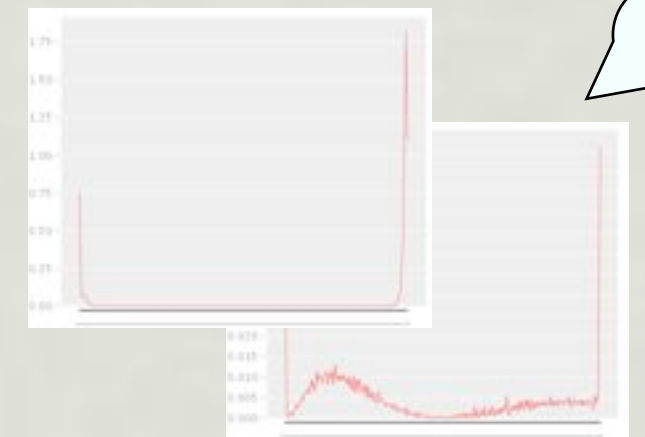
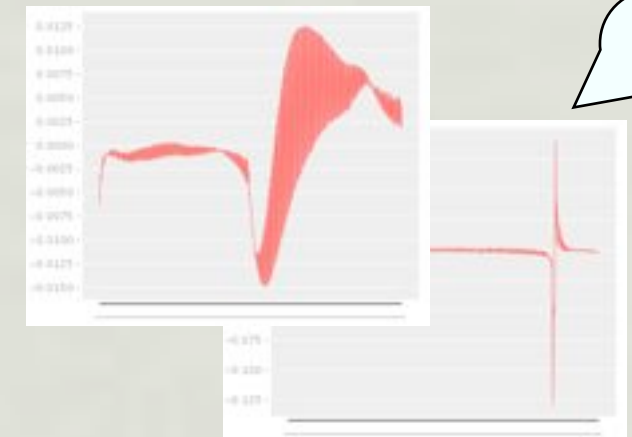
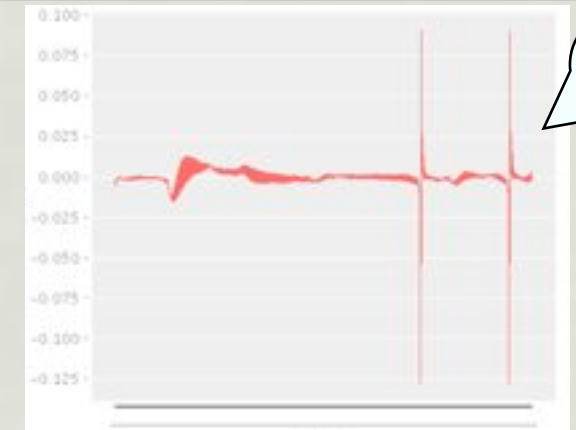


3

提案システム 特徴量の計算

✳ 特徴量: PSD

4. 波形信号をフレーム単位に分割
5. 各フレーム毎にPSDを計算
6. 算出したPSDの平均値を計算



提案システム

特徴量の計算

- ＊ 波形のフレーム分割と前処理

- i. 波形を一定サイズのフレームに分割

- ＊ フレーム長 T に対してシフト幅 $T/2$ で切り出す

- ii. 各フレーム毎にトレンド除去

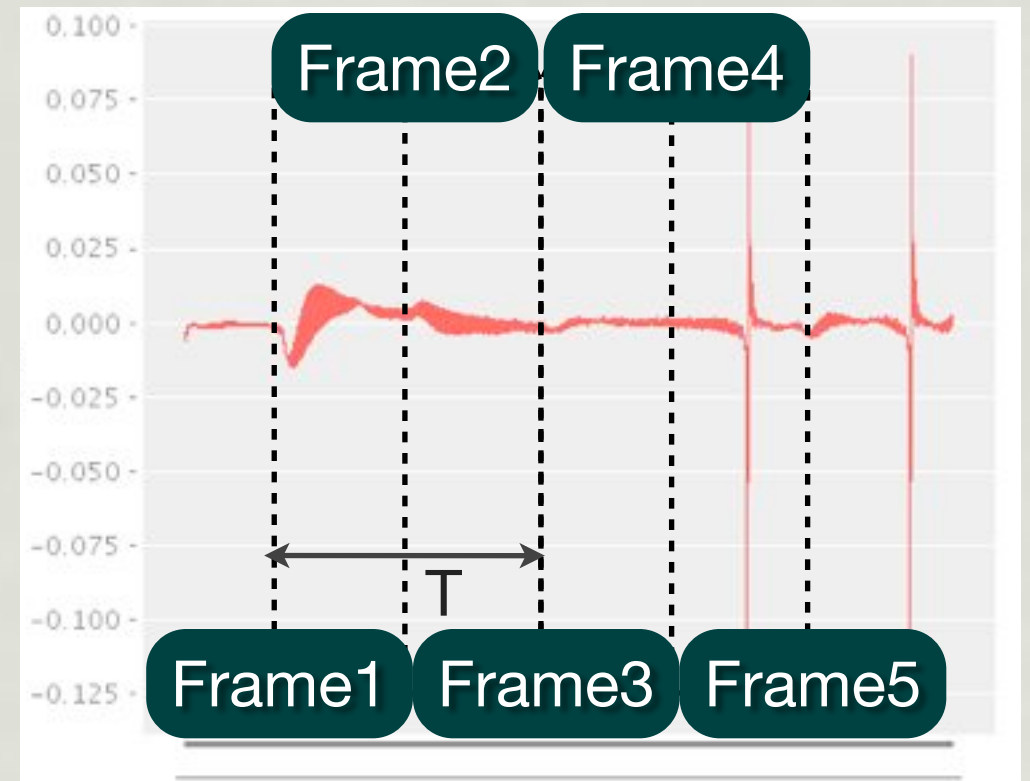
- ＊ 特徴量の計算

- i. 窓関数+FFT(Fast Fourier Transform)

- ii. 周波数応答補正

- ＊ 観測データのスペクトルから地震計の周波数応答を除去

- iii. PSDの計算



提案システム

Self-Organizing Map

- ＊ 波形間の類似関係を2次元空間上にマッピング
 - ＊ 出力空間: $N \times M$ のセルを持つ2次元グリッド
 - ＊ 各セルは特徴量と同じ次元の重みベクトルを持つ
- ＊ 学習
 - ＊ 入力データに最も類似した重みベクトルを持つセルを探索
 - ＊ 最も類似したセルと、その近傍のセルの重みベクトルを更新
- ＊ 可視化
 - ＊ 入力データに最も類似した重みベクトルを持つセルを探索
 - ＊ 最も類似したセルに入力データを割り振る

提案システム

パラメータ

- ＊ 月地震イベントの抽出，前処理，特徴量計算，SOMにおいて，多数のパラメータが存在
- ＊ パラメータの試行錯誤が必要

パラメーター一覧表

月地震波形抽出	前処理	特徴量算出	SOM
抽出データ長	フィルタ処理のパラメータ	フレーム長	入力データのフィルタリング
		窓関数	出力空間
		周波数応答補正の有無	類似度計算手法

提案システム

動作例

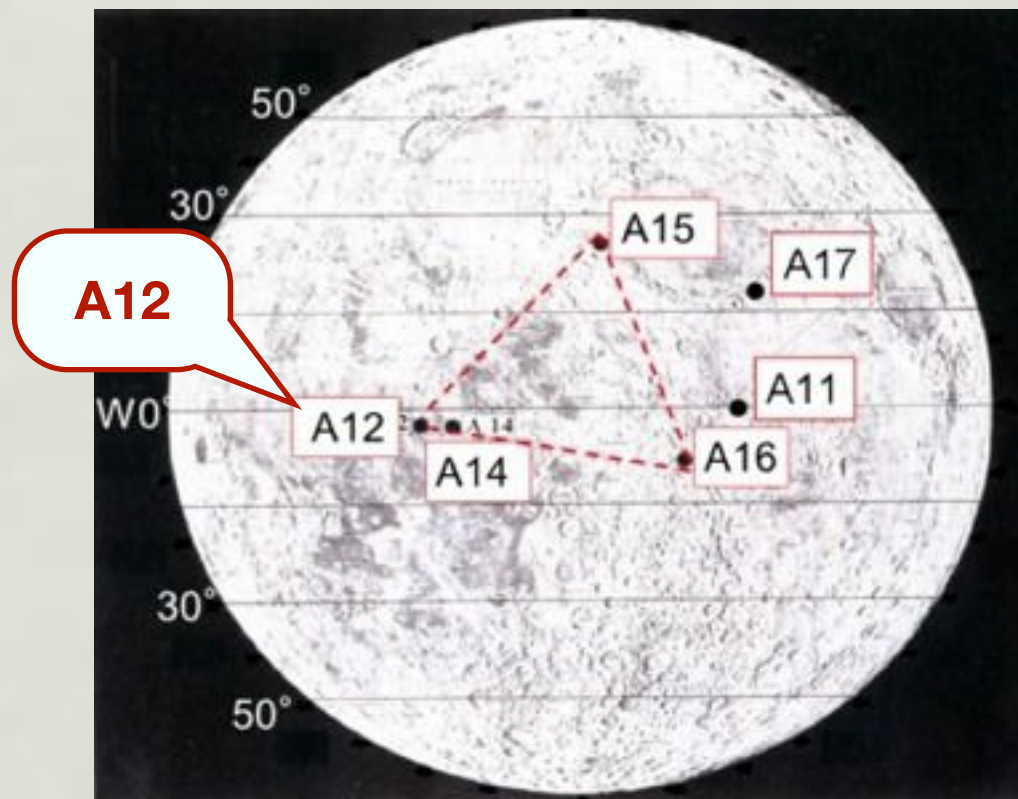
- ＊ データ

- ＊ A12の地震計が観測した長周期地震計波形

- ＊ XYZ成分各軸毎にSOM

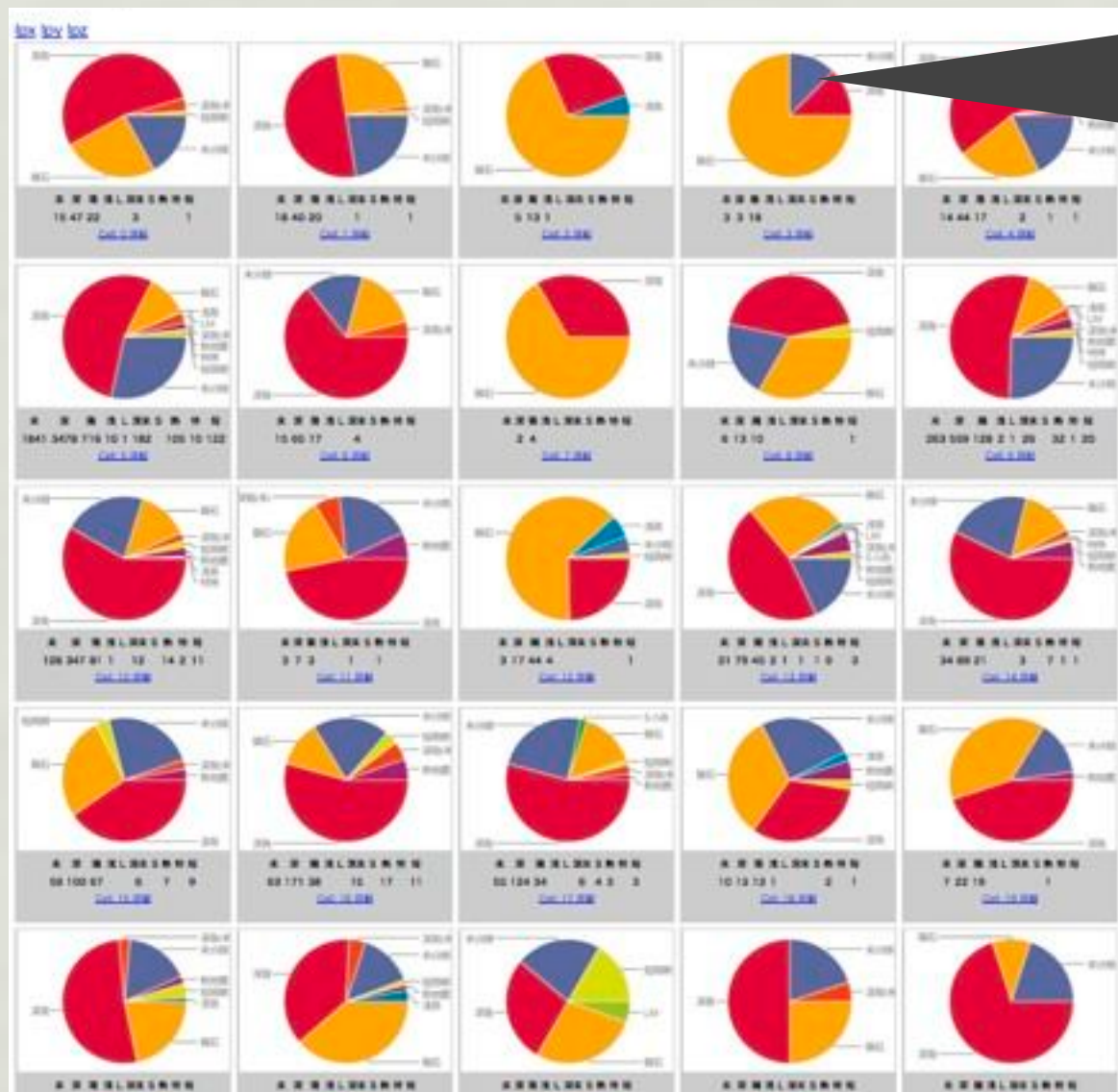
実験パラメータ

抽出データ長	5分間
フィルタ処理の パラメータ	0.3[Hz]-1.5[Hz] 次数: 1
フレーム長	512
窓関数	hanning
周波数応答関数	有
SOM出力	5x5(2次元空間)
類似度計算	ユークリッド距離



提案システム 動作例

※ SOM全体の結果表示画面



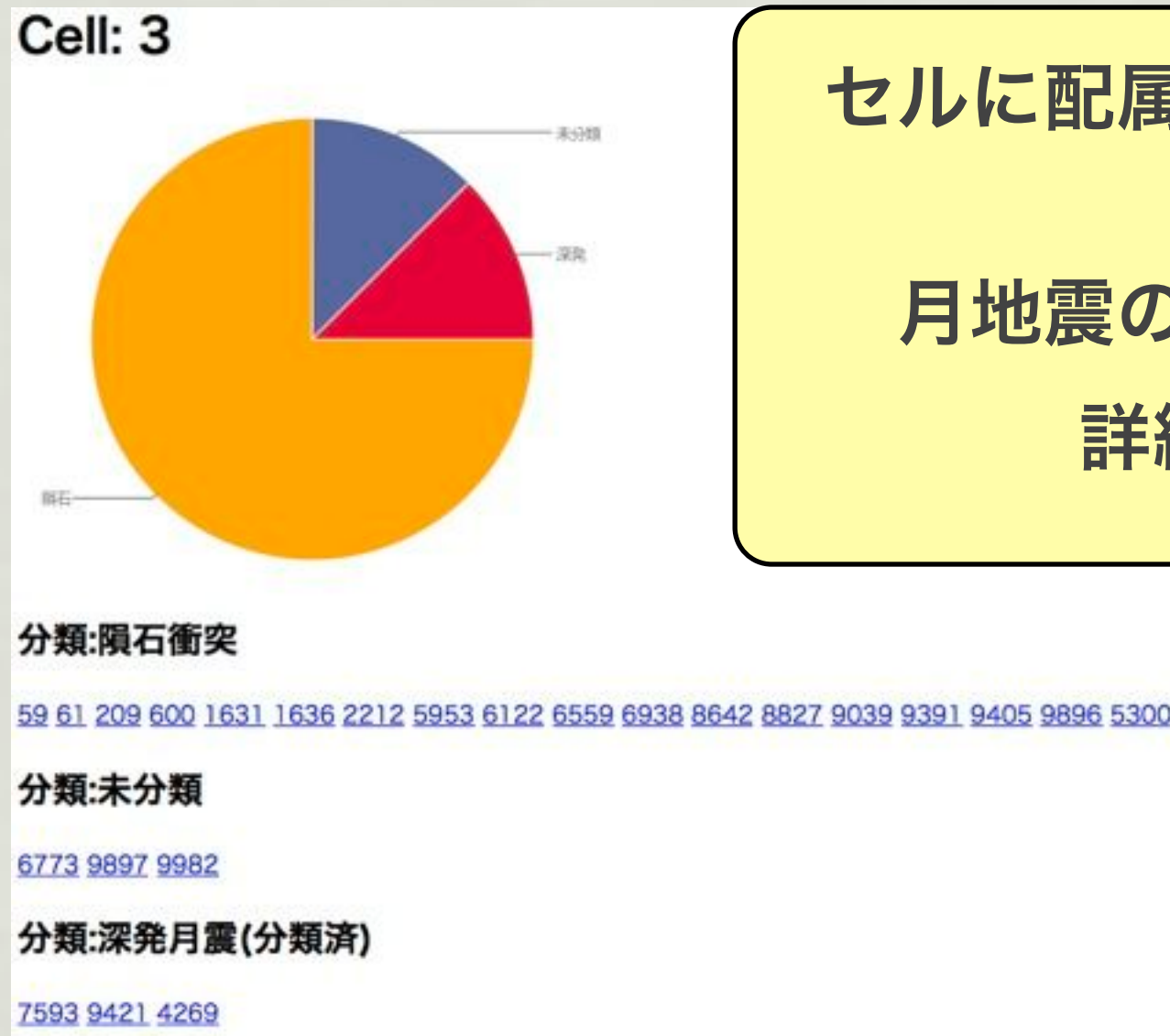
データの内訳

X軸のSOM結果

提案システム

動作例

＊ SOM出力のセルの詳細画面



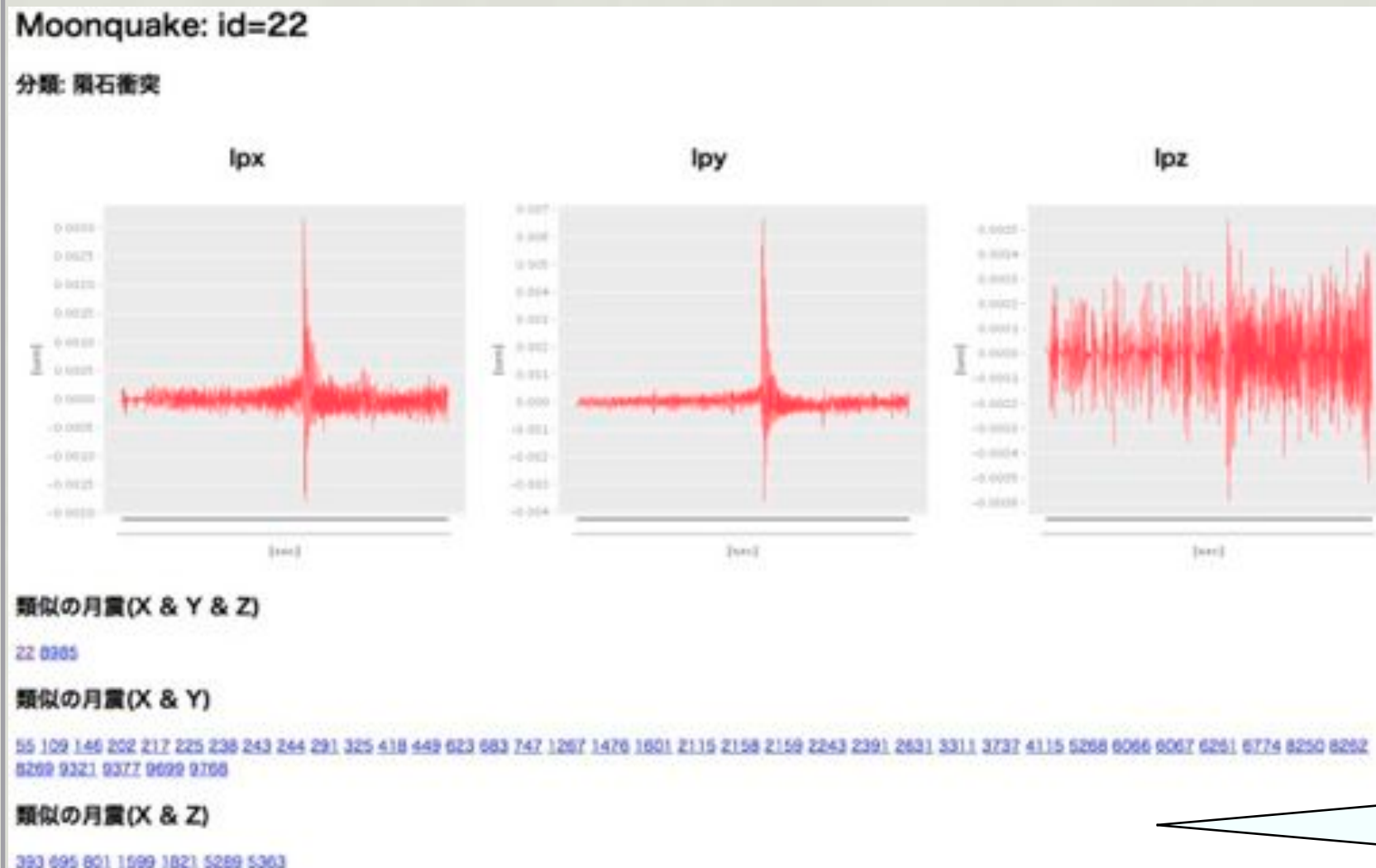
セルに配属されたデータ

月地震のIDによって
詳細表示

X軸のSOM結果のセル番号3の詳細

提案システム 動作例

＊ 各月地震の詳細画面



月地震(ID=22)の詳細画面

3軸それぞれのSOMの
結果に基づいて
月地震間の類似関係を
可視化

X, Y, Z軸それぞれの
SOMにおいて
同じセルに配属されたものを
リストアップ

まとめ・今後の課題

＊ まとめ

- ＊ 月地震研究の促進のために、
大規模月地震データの可視化システムを構築
 - ＊ Hadoop+Hiveを用いた月震データ管理
 - ＊ Self-Organizing Mapを用いた2次元マッピング

＊ 今後の課題

- ＊ 月地震研究者からのフィードバック
- ＊ 複数の観測地点での波形のSOM結果を
突き合わせた可視化