

# 時空間変動データからの ホットスポット 自動抽出・要約システムの開発

2017.2.10

宇宙科学情報解析シンポジウム@ISAS

本田 理恵, 林 諒, 松永知也\*

高知大学 理学部 情報科学教室

(\*現日立システムズ)

# 背景

- 大量の時空間データが蓄積
  - 衛星画像、セキュリティカメラの画像、センサーネットワークデータなど
- 時空間ビッグデータから、時間空間の変動パターンを抽出→自然、社会現象の理解、予測に活用
  - 機械学習の手法を適用
    - 宇宙科学情報関連分野にも応用

# 時空間ビッグデータへのアプローチ

- 時空間総合
  - Matsubara et al., 2014 KDDなど
- 時間優先
  - 時系列解析
  - その時空間変化を分析
    - イベントクラスタの抽出と時空間相関ルール(森、本田, JSAI2016)
- 空間優先
  - セグメンテーション、ホットスポット→オブジェクト抽出
  - セグメント、オブジェクトの時間変化として解析
  - (本発表)

# ターゲットとタスク

- ターゲット

- 時系列画像, 時系列の3Dグリッドデータ

- 気象画像、植生などの地球観測画像、フェーズドアレイ気象レーダ画像、天体画像、

- タスク

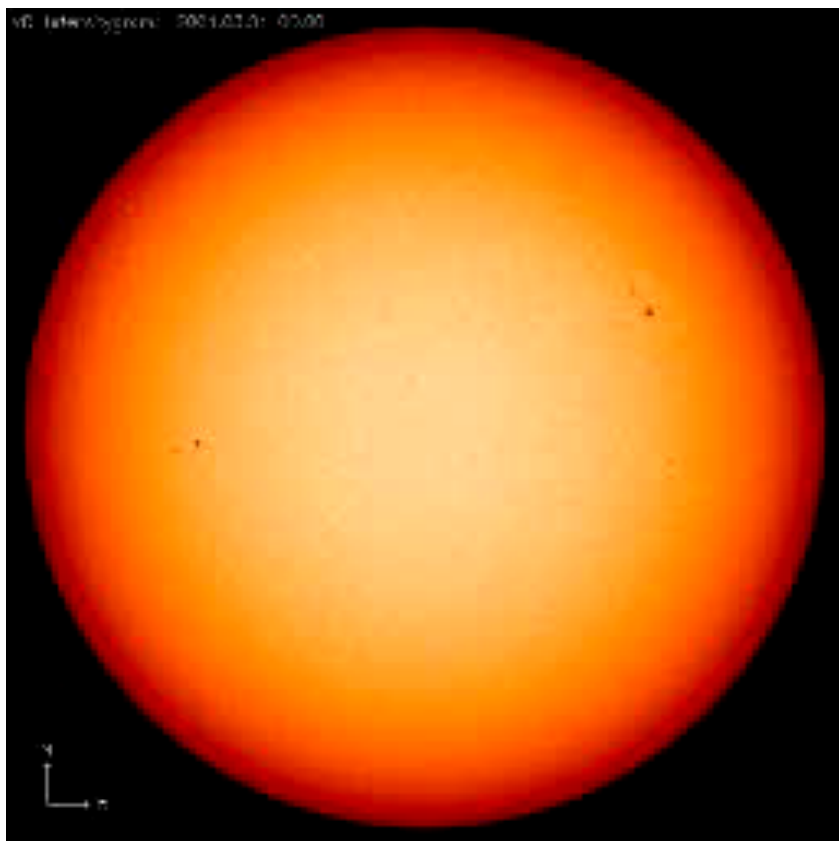
- 時系列データ(2D, 3D)に含まれる”周囲とは異なる領域—ホットスポット—”を抽出し、その時間的変遷を要約

- インタラクティブ知識発見  
高次の知識発見の素材

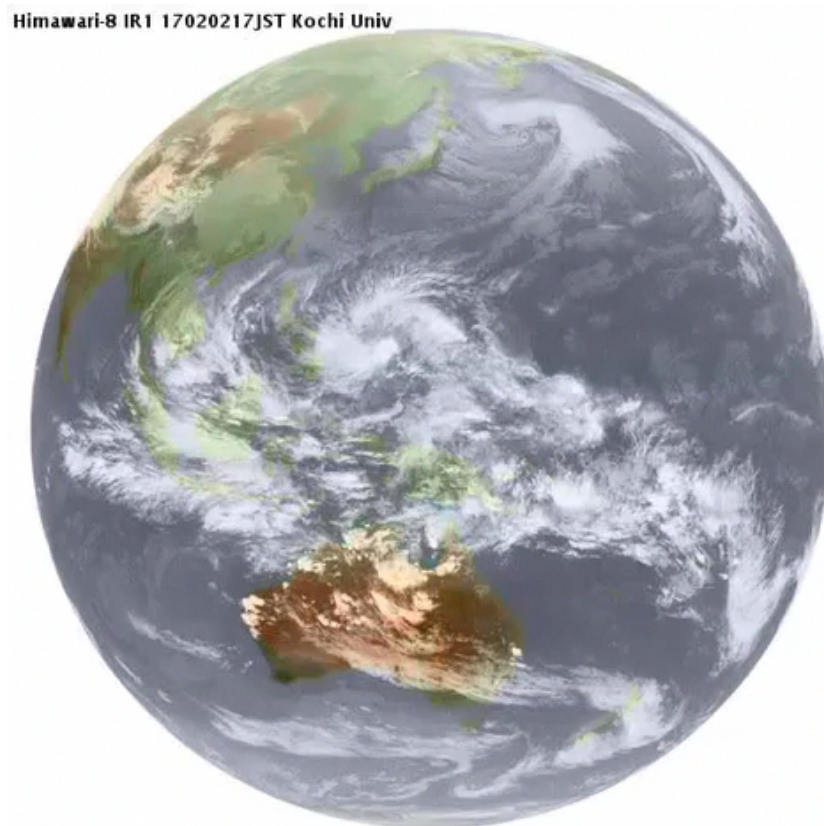
ホットスポット—オブジェクト

- 周囲とは異なる値、分布
- 誕生から消滅まで一定の存続期間
- 生涯に分裂や他のオブジェクトとの相互作用を経験

# 地球観測、天文分野における 時空間データのホットスポット



MDI Sunspot Crossings (2001)



ひまわり8(高知大学気象情報ページより)

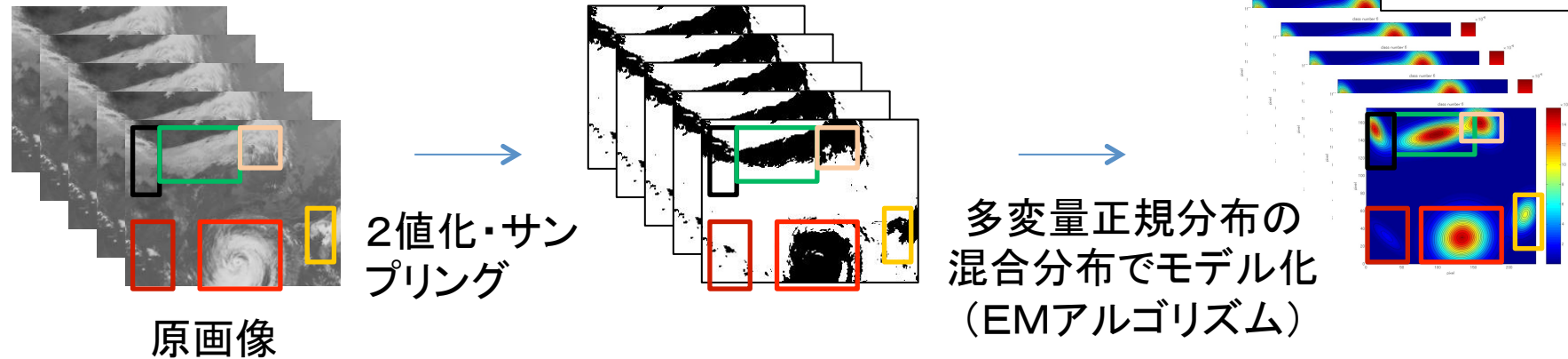
太陽黒点、雲塊 → ホットスポット=オブジェクト

# 研究事例

- 気象画像からのオブジェクト(雲塊)抽出とその変遷要約システムの構築  
– (松永他, DEIM2016)
- 3Dフェーズドアレイ気象レーダからの降水コア抽出(松岡他 JPGU2016、林他 deim2017予定)

# 手法—オブジェクト抽出—

- 成分数毎に以下を取得
- ・重心
  - ・広がり
  - ・重み係数

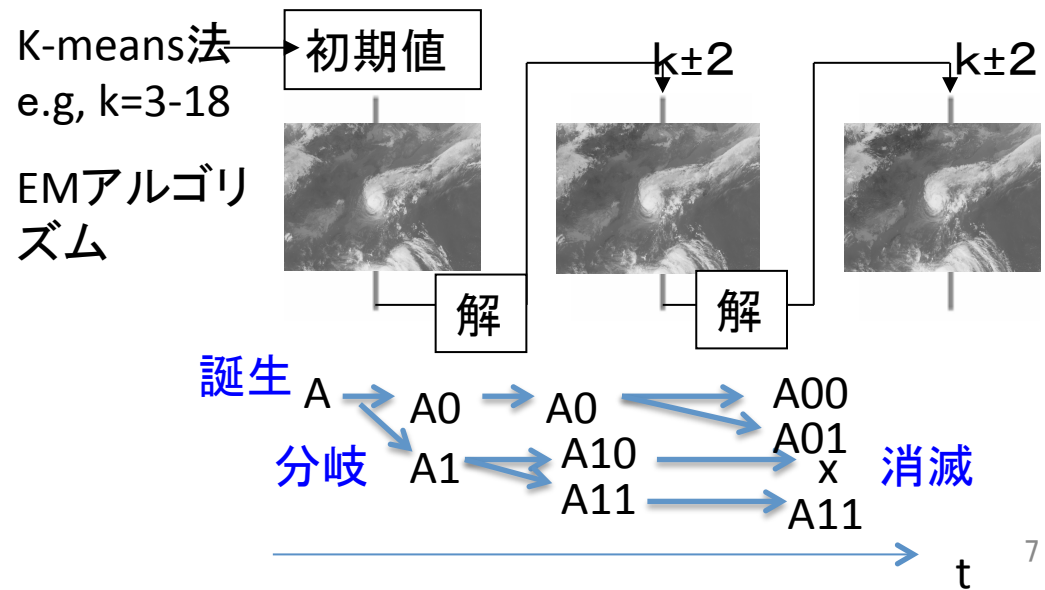


最初の画像の初期値はK-means法  
 与え、ベイズ情報量基準で最適モデル  
 を選択

2枚目以降の初期値は前の時間の  
 解を継承  
 生成消滅を考慮するため  $k \pm 2$  でゆ  
 らぎを与える

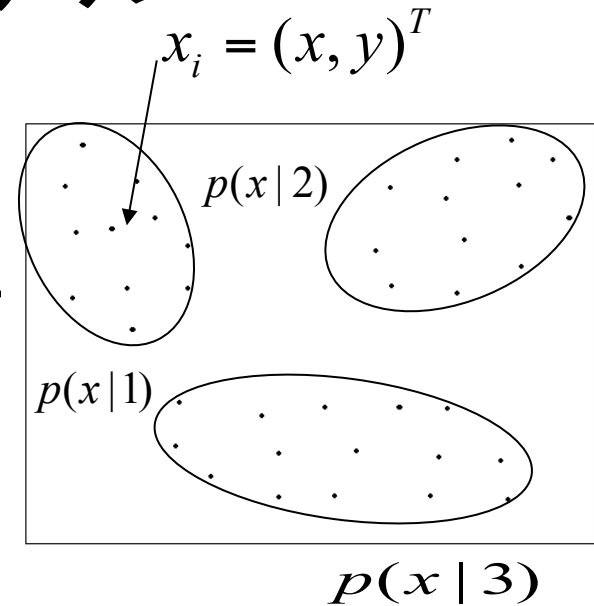
- 処理時間の短縮 (数十分の1)
- 親子関係の記述

求まらなかった場合、一旦リセット



# オブジェクトのモデル

- オブジェクトが存在する点の座標集合  
 $\{(x_i, y_i) \mid i = 1, \dots, N\}$
- 各オブジェクトを多変量正規分布、その混合



$$P(x) = \sum_{j=1}^K \omega_j p(x|j)$$

$$p(x|j) = \frac{1}{(\sqrt{2\pi})^2 \sqrt{|\Sigma_j|}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma_j^{-1} (x - \mu_j) \right\}$$

$$\mu_j = \begin{bmatrix} \mu_{xj} \\ \mu_{yj} \end{bmatrix}, \Sigma_j = \begin{bmatrix} \sigma_{xj} & \sigma_{xyj} \\ \sigma_{yxj} & \sigma_{yj} \end{bmatrix}, \omega_j, K$$

中心座標      形状      重み係数, 成分数

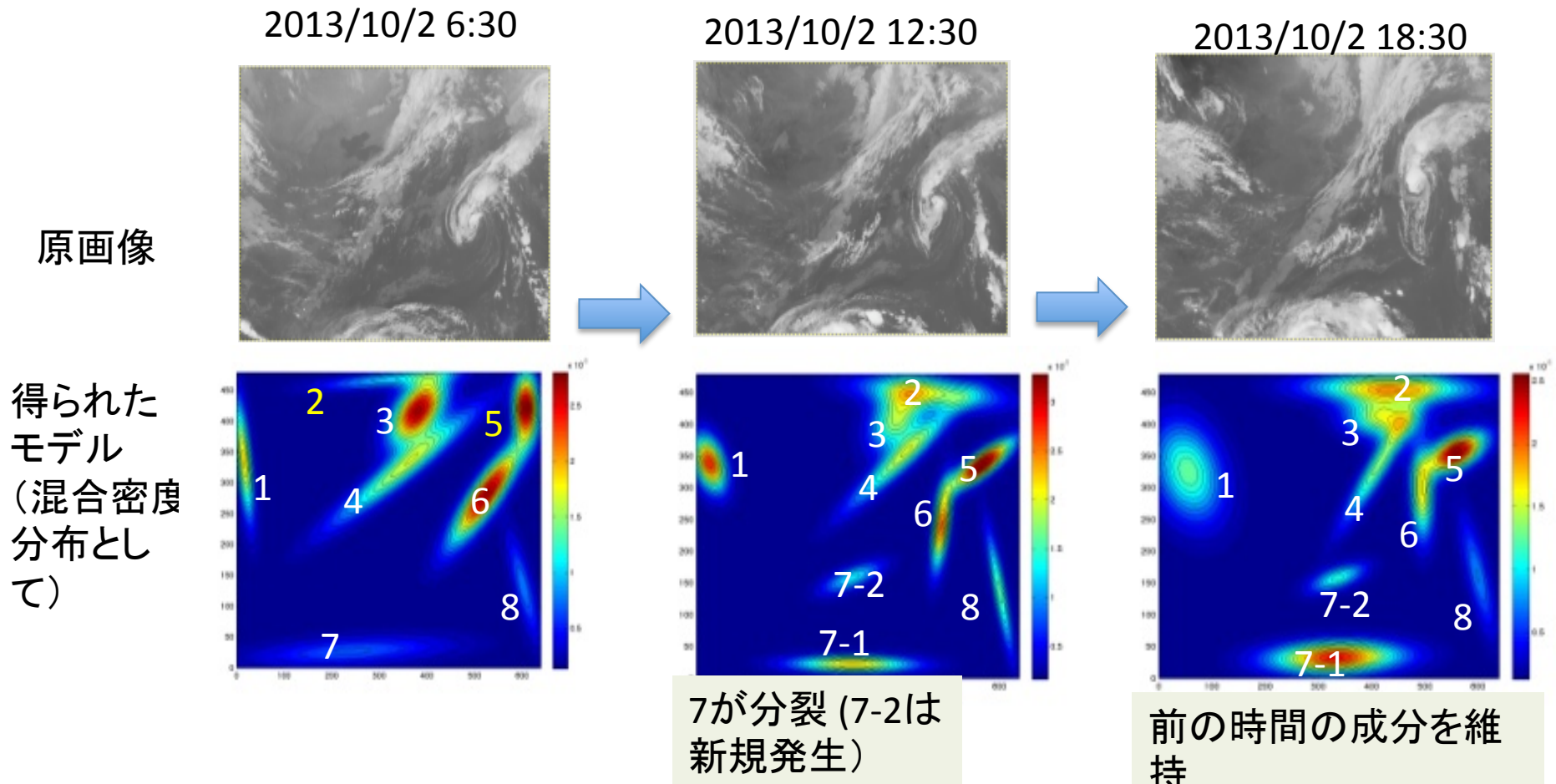
この値も未知

EMアルゴリズムで、  
モデルパラメータを決定

あり得る範囲で総当たりで  
解を求め、BIC(ベイズの情報  
量基準)で最適解を選択す  
る

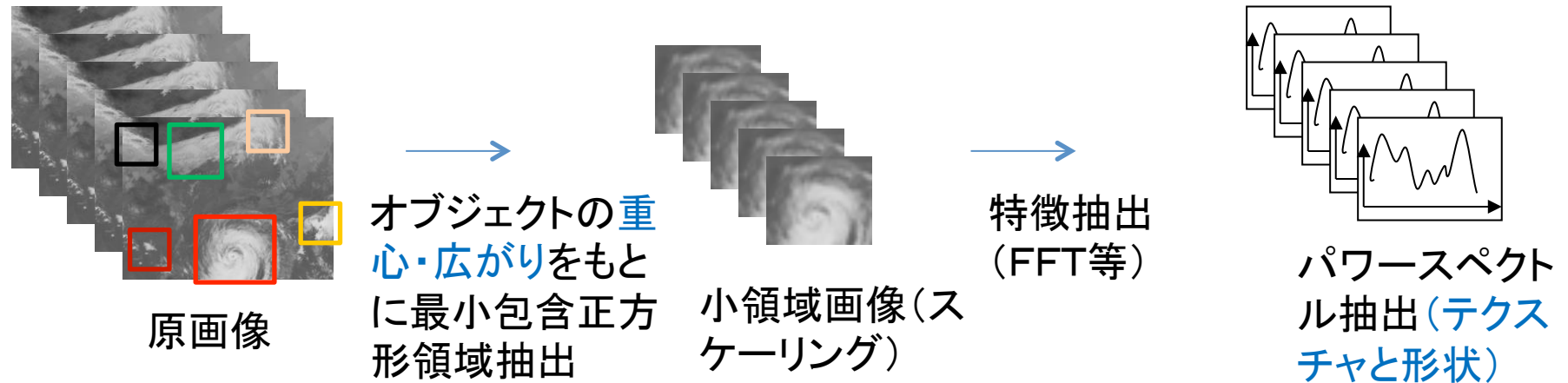


# 気象画像での実験結果例 —オブジェクト抽出—

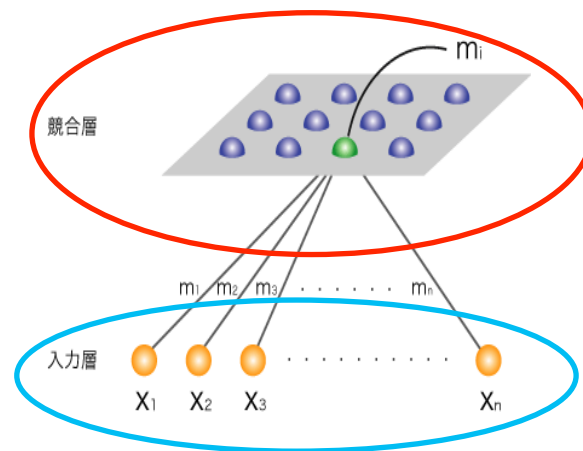


- 成分数の自動決定→概ね成功、消滅や、分裂、新規発生(分裂として)にも対応
- 各成分の中心位置、広がりが自動記載できるようになる
- 同時に親子関係も記載できる

# 手法—意味的ラベリング—



自己組織化マップ(SOM)による分類



似た特徴をもつ画像が各クラスターに分類される

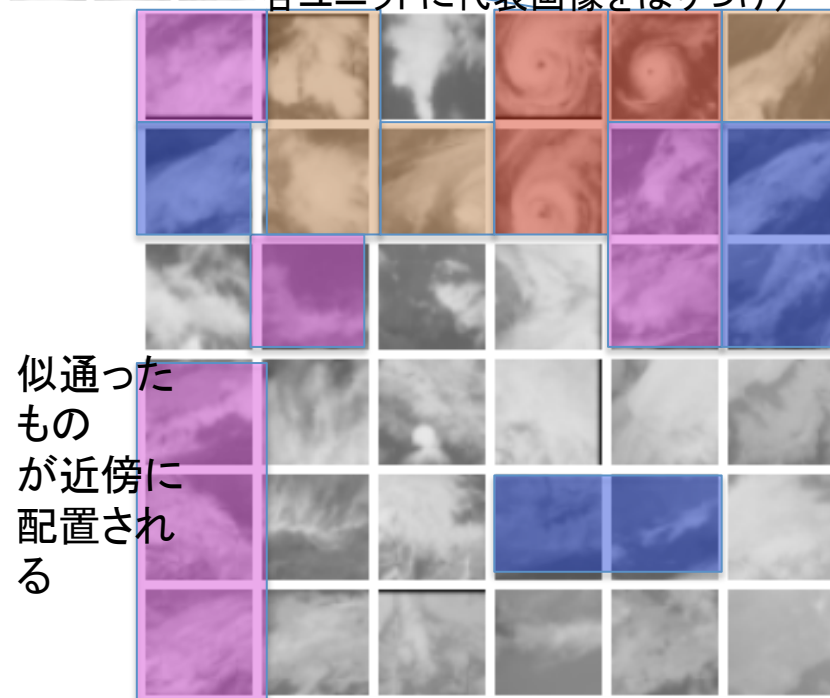
結果をもとにクラスターに意味的ラベル付け(台風、前線など)

パワースペクトル

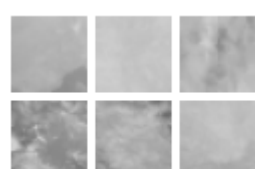
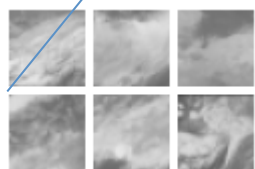
(松永、本田, deim2015)

# SOMによるクラスタリング結果 — 意味的ラベル付けと精度評価 —

各クラスタの代表画像  
(自己組織化マップの競合層の  
各ユニットに代表画像をはりつけ)



似通った  
ものが近傍に  
配置される



(松永ほか 2015)

精度: 同一クラスタ内の適合画像の割合

梅雨前線、 90%	台風の一部、 60%	粗い雲、 75%	台風、 67%	台風、 100%	台風の一部、 100%
寒冷前線、 60%	台風の一部、 20%	台風の一部、 46%	台風、 40%	梅雨前線、 53%	寒冷前線、 39%
粗い雲、 31%	梅雨前線、 70%	粗い雲、 32%	広い雲、 36%	梅雨前線、 53%	寒冷前線、 62%
梅雨前線、 57%	粗い雲、 63%	広い雲、 50%	真っ平らな雲、 100%	真っ平らな雲、 33%	薄い雲、 20%
梅雨前線、 50%	冬特有の雲、 9%	粗い雲、 50%	寒冷前線、 29%	寒冷前線、 33%	広い厚い雲、 83%
梅雨前線、 55%	粗い広い雲、 40%	冬特有の雲、 46%	薄い広い雲、 68%	薄い広い雲、 86%	薄い広い雲、 88%

- 台風、前線、薄く広がる雲等の意味的インデクス付け  
→ 平均精度(適合率) 61%

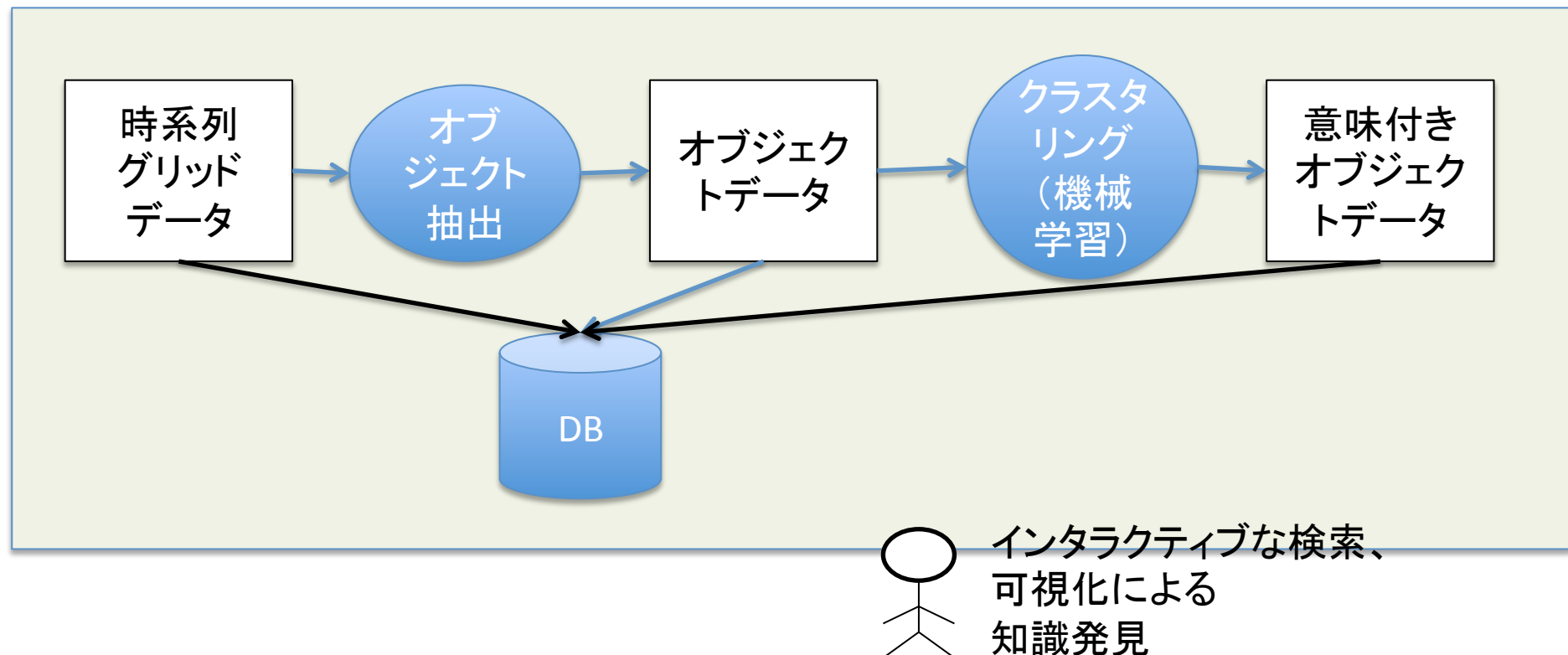
# システム概要

下記の機能を満たすシステムを構築

オブジェクトを自動的に時系列2次元、3次元グリッドデータから抽出

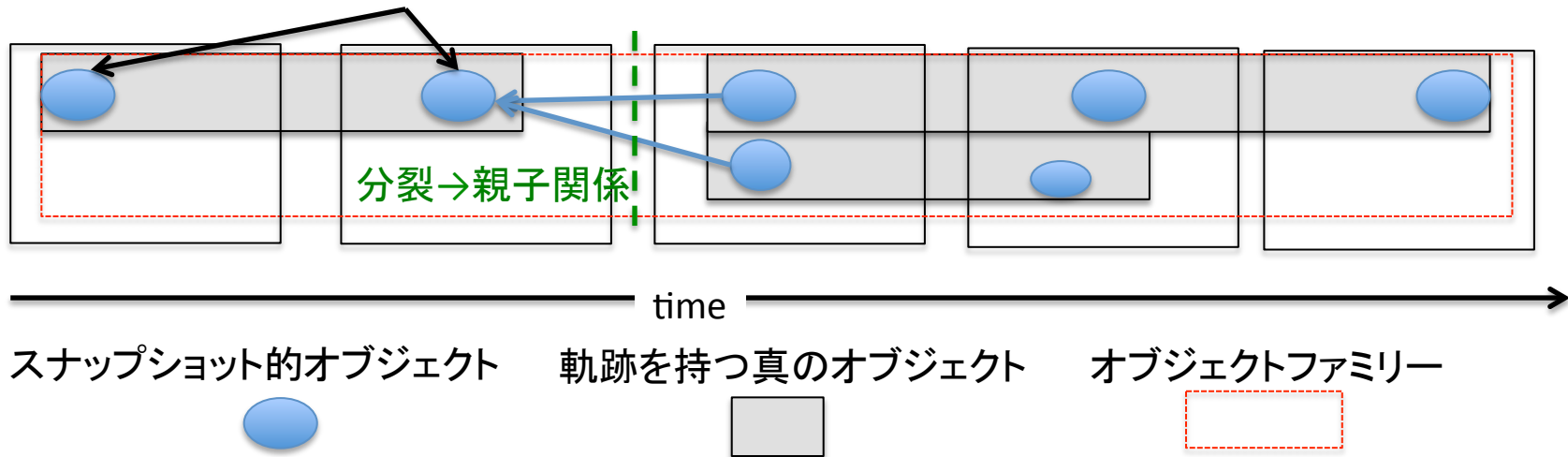
オブジェクトの特徴に応じて意味的ラベル付け

オブジェクトの相互作用や変遷を時間、空間の両面で要約、可視化



# オブジェクトの階層構造

同一のラベル名が付与されている→追跡が可能

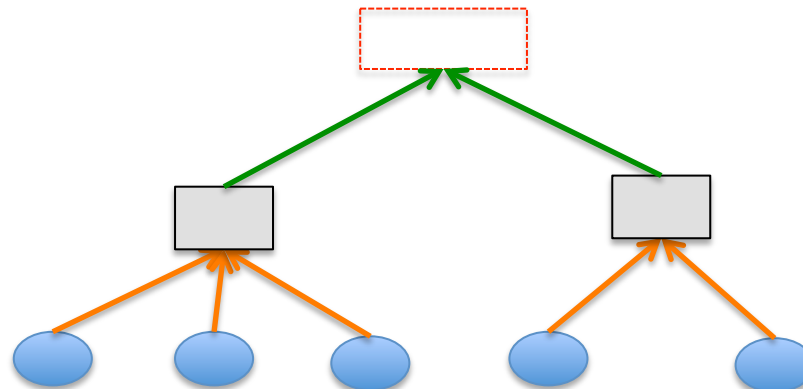


## オブジェクト情報を階層的な構造で表現

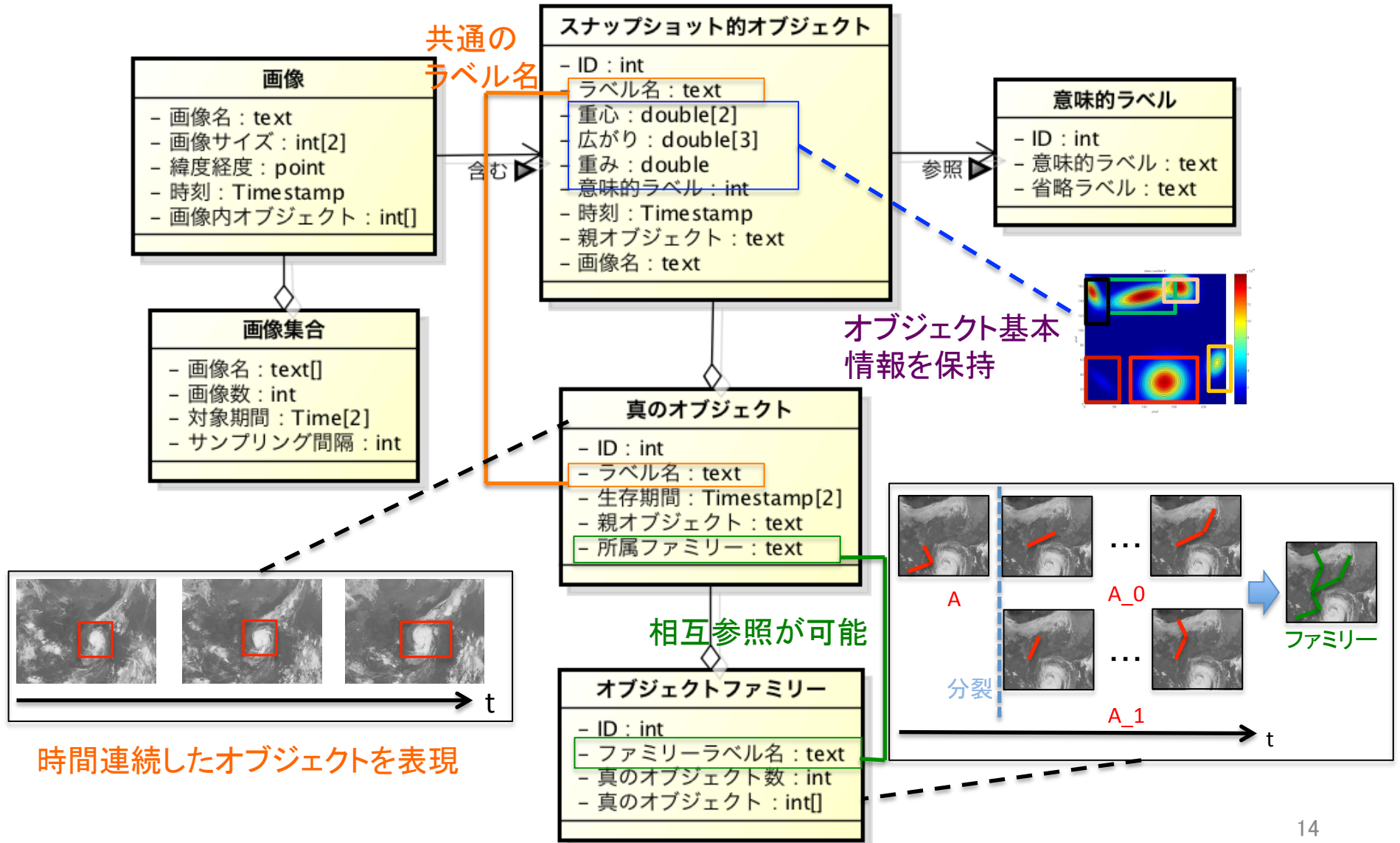
オブジェクトファミリー

軌跡を持つ真のオブジェクト

スナップショット的オブジェクト



# データモデル





# 時系列画像からのオブジェクト要約・検索システム

MTSAT7, IR1, 日本周辺の画像

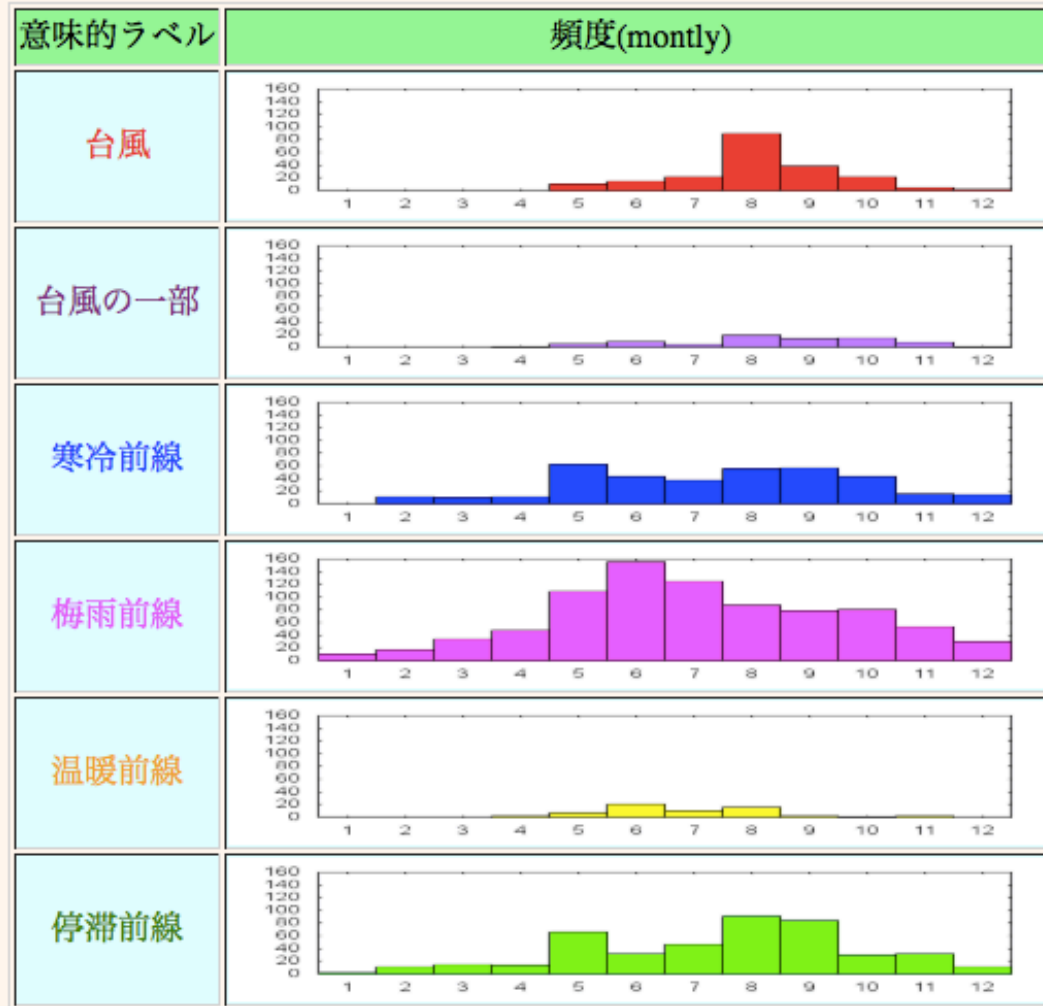
期間: 2006.1.1 ~ 2006.12.31 サンプル間隔3hour

オブジェクト検出結果のサマリ [ [monthly](#) | [weekly](#) ]

## 構築したシステム例

2006年1年間の日本付近のひまわり画像 (MTSAT-7)  
 (高知大気象情報ページ)  
 640x480pix  
 サンプル間隔: 3時間おき

全2920枚  
 抽出された全オブジェクト 14058  
 所要時間 7094sec  
 学習時間4239sec  
 @Intel core i5  
 3.2GHz, iMac



検索対象期間と可視化情報の形式を選択して下さい(現在の検索期間目安は3日間です)

START  yyyy-MM-dd hh:mm:ss

END  yyyy-MM-dd hh:mm:ss

画像
  生命線図
  軌跡
  ファミリー

期間を入力して選択

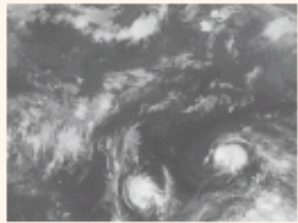
## サマリ

月ごとの抽出されたオブジェクト (ホットスポット) の種類ごとの頻度

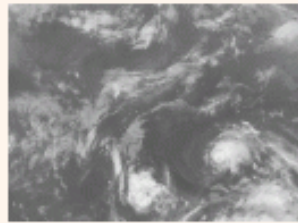
自動的に生成される

## 検索結果一覧

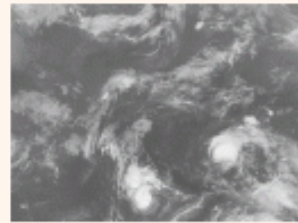
対象期間 2006-08-07 00:00:00 から 2006-08-08 00:00:00 までの気象画像(yymmddtt.pgm)を表示しています。  
画像名をクリックすると詳細情報を表示します。(※実際の画像はpgm形式ですが、ここでの表示はgif形式となっています)



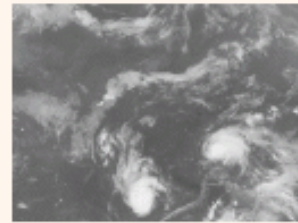
[fe.06080700.pgm](#)



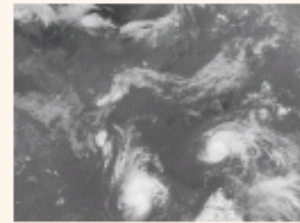
[fe.06080703.pgm](#)



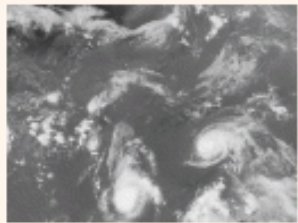
[fe.06080706.pgm](#)



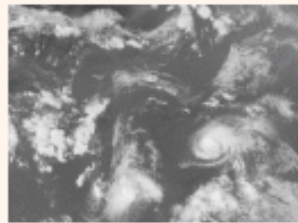
[fe.06080709.pgm](#)



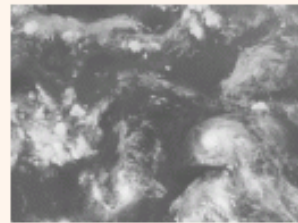
[fe.06080712.pgm](#)



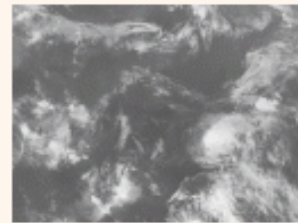
[fe.06080715.pgm](#)



[fe.06080718.pgm](#)



[fe.06080721.pgm](#)



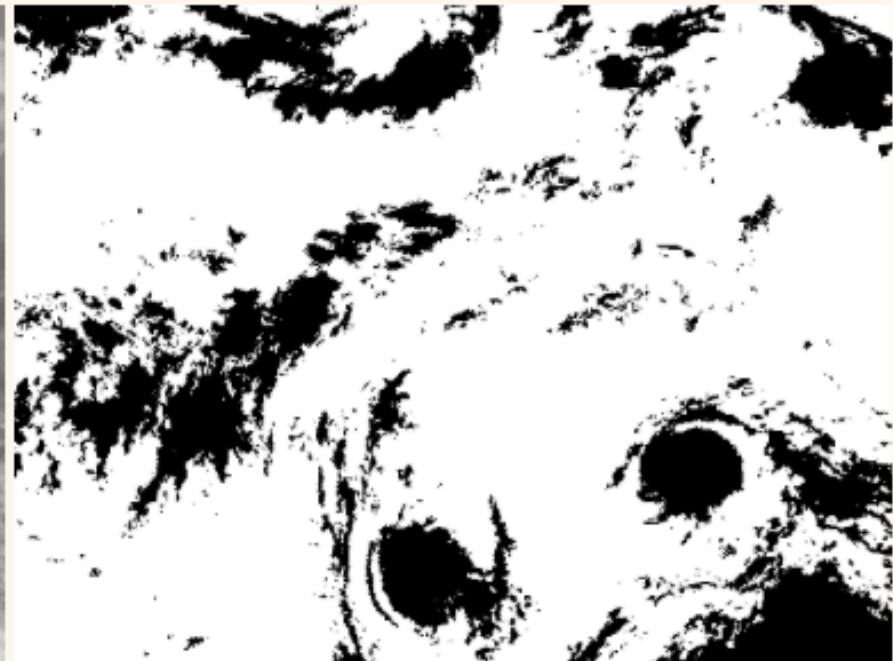
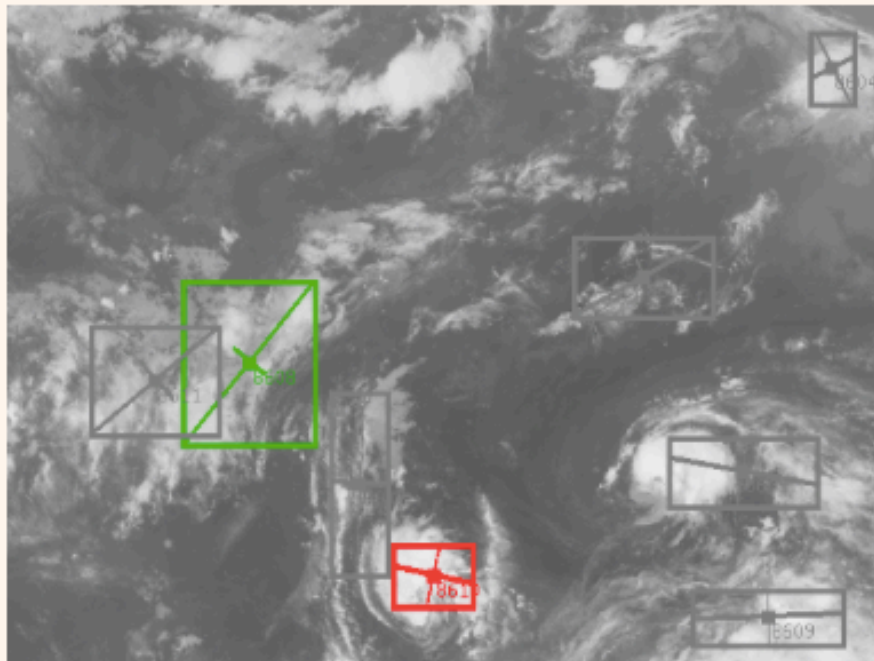
[fe.06080800.pgm](#)

[別条件で検索](#)



画像 fe.06080700.pgm に含まれるオブジェクト (右は2値化画像)

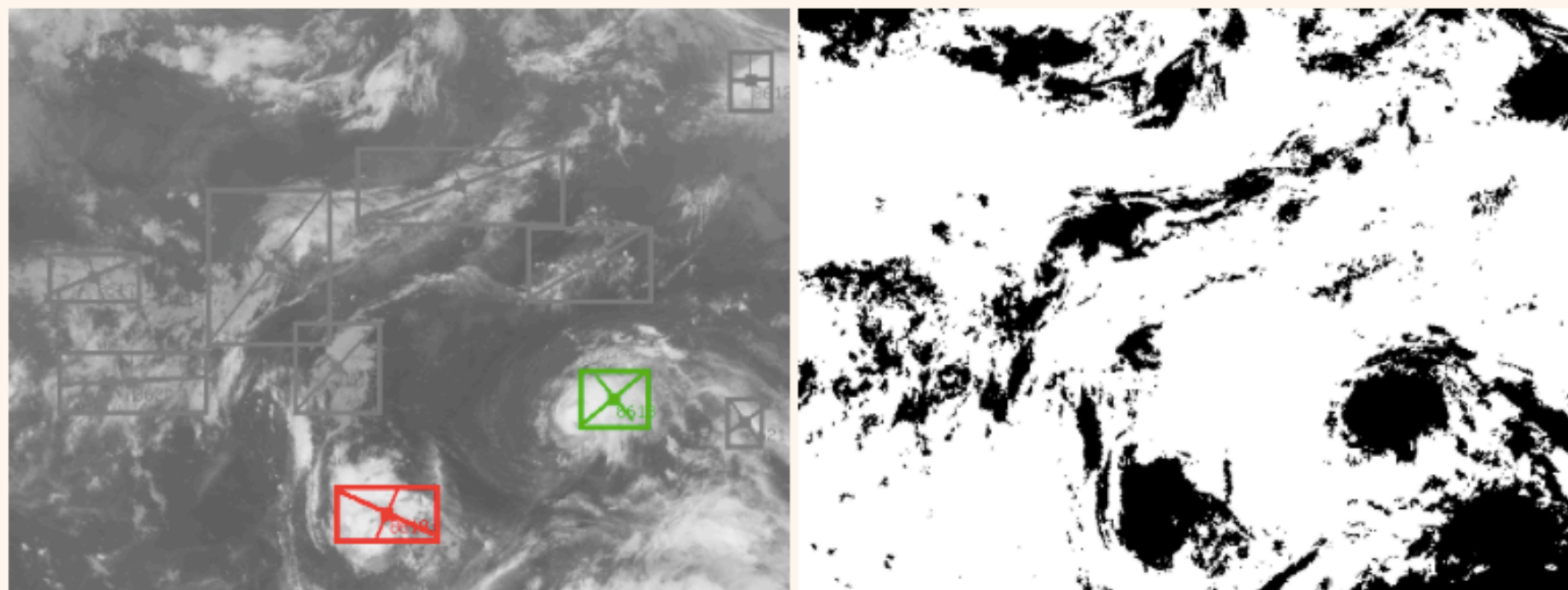
<前時刻> <次時刻>



ID	オブジェクトラベル名	意味的ラベル	発生時刻	消滅時刻
8604	06072109.d_01_0_0_1_1_0_0_0_0_0_0_1_1	7	2006-08-02 09:00:00	2006-08-07 06:00:00
8605	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0	7	2006-08-06 12:00:00	2006-08-07 00:00:00
8606	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_0	7	2006-08-04 18:00:00	2006-08-07 03:00:00
8607	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_0	7	2006-08-06 15:00:00	2006-08-07 03:00:00
8608	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_1	6	2006-08-06 09:00:00	2006-08-08 03:00:00
8609	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_1_0_1_0_1_0	7	2006-08-06 15:00:00	2006-08-08 21:00:00
8610	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_1	1	2006-08-06 21:00:00	2006-08-10 21:00:00
8611	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_1	7	2006-08-06 21:00:00	2006-08-07 09:00:00

画像 fe.06080703.pgm に含まれるオブジェクト (右は2値化画像)

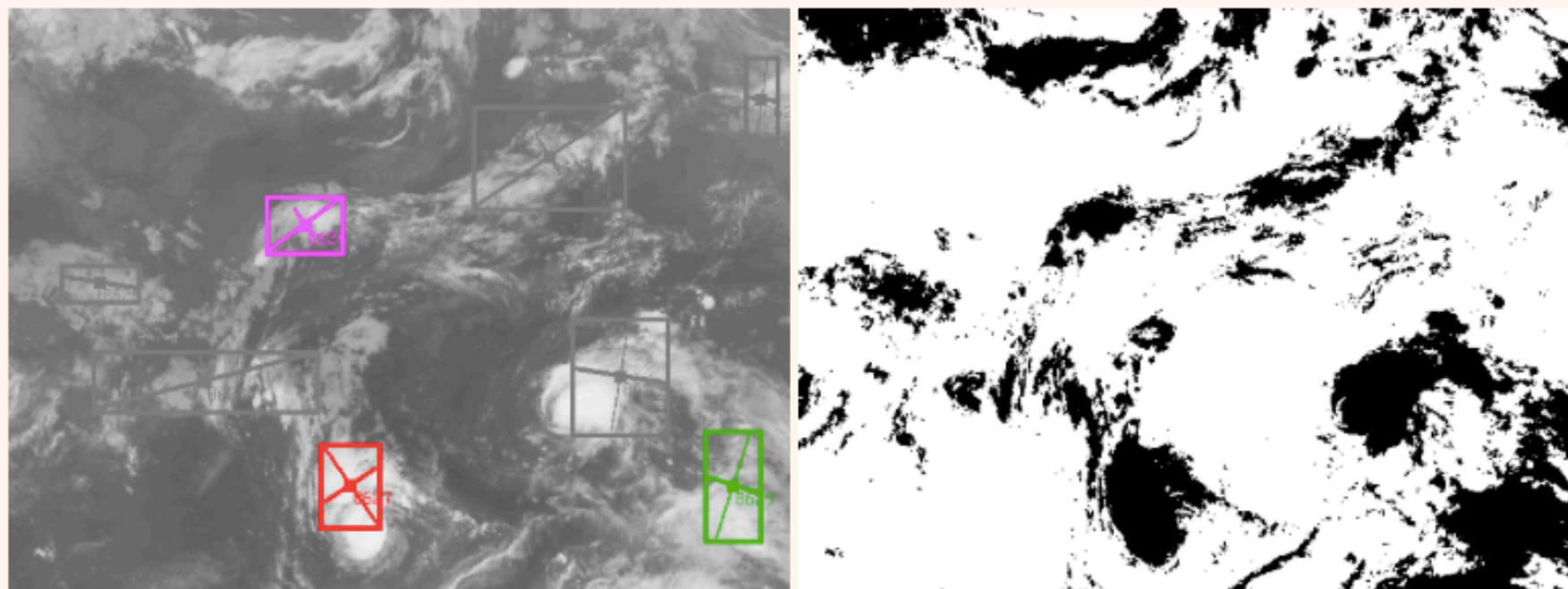
[<前時刻>](#) [<次時刻>](#)



ID	オブジェクトラベル名	意味的ラベル	発生時刻	消滅時刻
8612	06072109.d_01_0_0_1_1_0_0_0_0_0_0_1_1	7	2006-08-02 09:00:00	2006-08-07 06:00:00
8613	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_0	6	2006-08-07 03:00:00	2006-08-08 06:00:00
8614	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_0	7	2006-08-04 18:00:00	2006-08-07 03:00:00
8615	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_0	7	2006-08-06 15:00:00	2006-08-07 03:00:00
8616	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_0	7	2006-08-06 00:00:00	2006-08-08 21:00:00
8617	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_0_0	7	2006-08-06 21:00:00	2006-08-07 21:00:00
8618	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_1	7	2006-08-06 09:00:00	2006-08-08 03:00:00
8619	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_1_1	1	2006-08-06 21:00:00	2006-08-10 21:00:00
8620	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_1	7	2006-08-06 21:00:00	2006-08-07 09:00:00

画像 fe.06080706.pgm に含まれるオブジェクト (右は2値化画像)

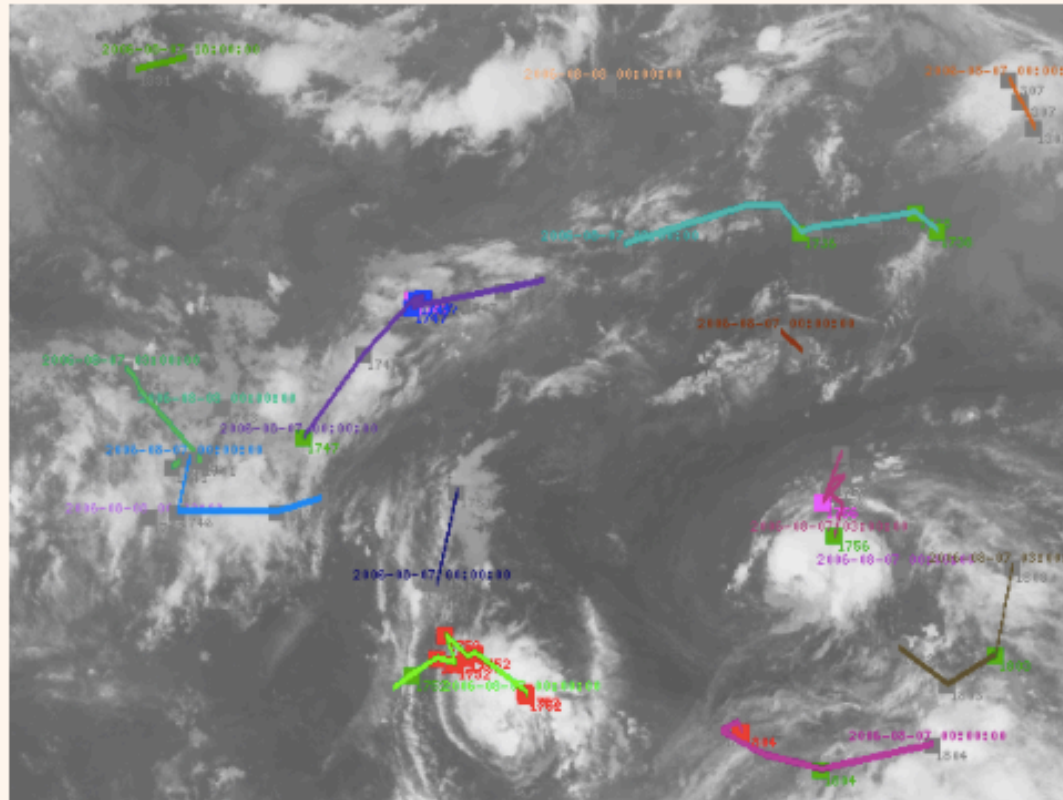
[<前時刻>](#) [<次時刻>](#)



ID	オブジェクトラベル名	意味的ラベル	発生時刻	消滅時刻
8622	06072109.d_01_0_0_1_1_0_0_0_0_0_0_1_1	7	2006-08-02 09:00:00	2006-08-07 06:00:00
8623	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_0	7	2006-08-07 03:00:00	2006-08-08 06:00:00
8624	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_0	7	2006-08-06 00:00:00	2006-08-08 21:00:00
8625	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_0	7	2006-08-06 21:00:00	2006-08-07 21:00:00
8626	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_1	4	2006-08-06 09:00:00	2006-08-08 03:00:00
8627	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_1	1	2006-08-06 21:00:00	2006-08-10 21:00:00
8628	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_1	7	2006-08-06 21:00:00	2006-08-07 09:00:00
8629	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_1_0_0_1	6	2006-08-07 03:00:00	2006-08-11 00:00:00



2006-08-07 00:00:00 から2006-08-08 00:00:00の間に発生していた軌跡



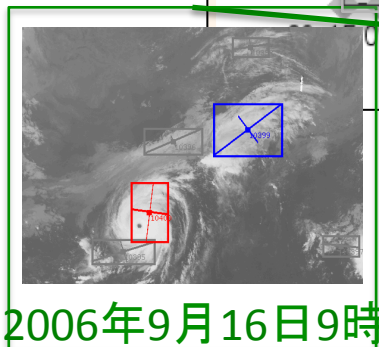
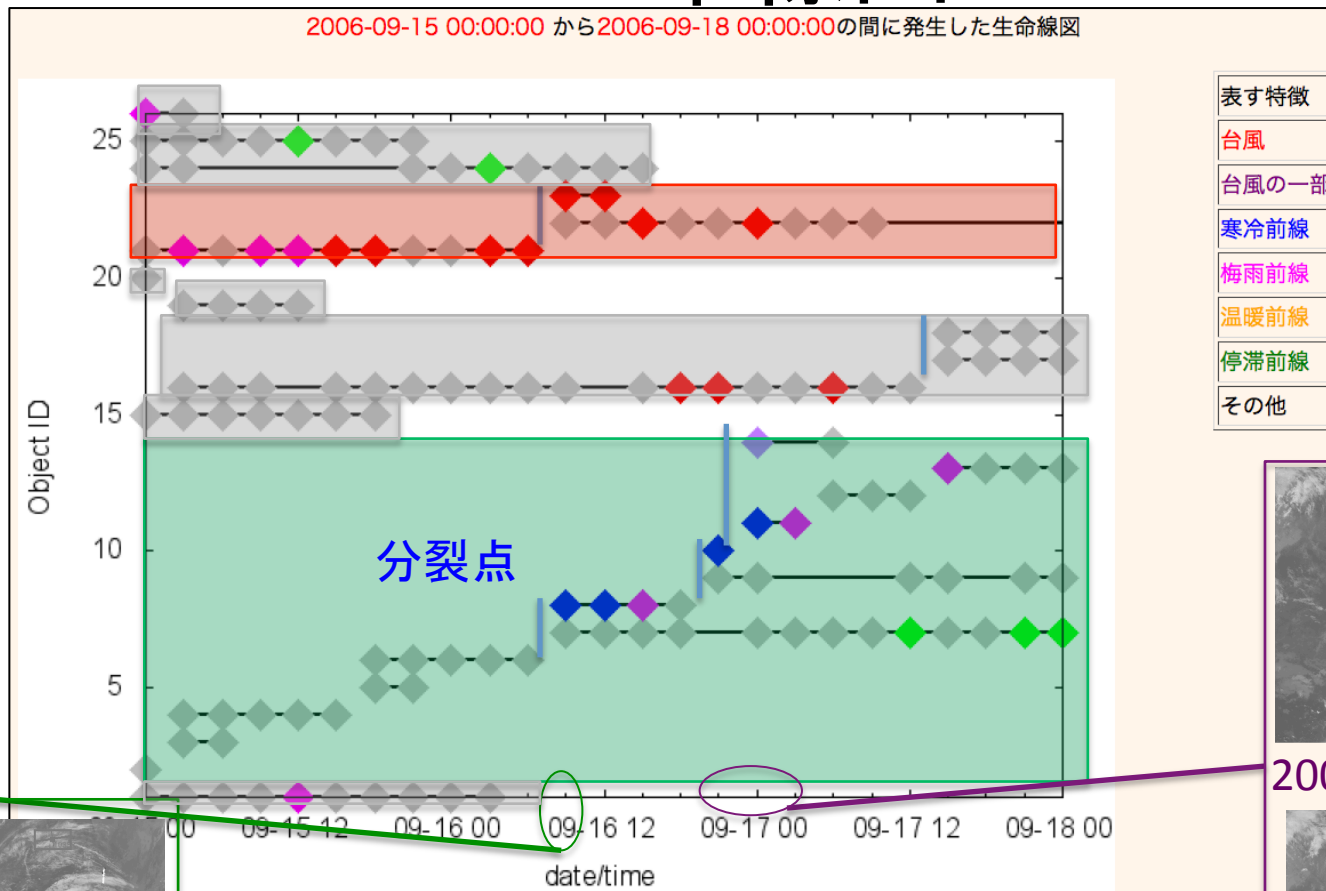
意味的ラベルの対応表は以下の通りです。

意味的ラベル	表す特徴
1	台風
2	台風の一部
3	寒冷前線
4	梅雨前線
5	温暖前線
6	停滞前線
7	その他

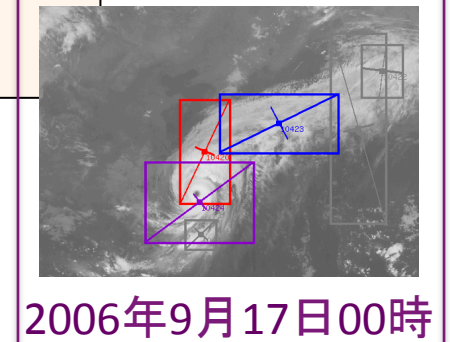
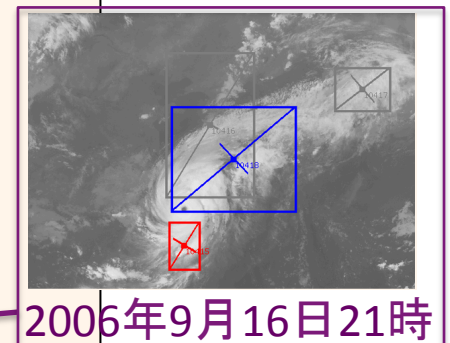
ID	オブジェクト名	発生時刻	消滅時刻
1307	06072109.d_01_0_0_1_1_0_0_0_0_0_0_0_1_1	2006-08-02 09:00:00	2006-08-07 06:00:00
1325	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_0_0	2006-08-08 00:00:00	2006-08-08 00:00:00
1331	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_0_1	2006-08-07 18:00:00	2006-08-10 15:00:00
1738	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_0	2006-08-06 00:00:00	2006-08-08 21:00:00
1741	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_0	2006-08-06 21:00:00	2006-08-07 21:00:00
1742	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_0_0_1_1_1_0_0_0	2006-08-08 00:00:00	2006-08-08 09:00:00
1743	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_0_1_1_1_0_0_1	2006-08-08 00:00:00	2006-08-08 18:00:00
1746	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_0_1_1_0_0_1_1_1_0_1	2006-08-06 21:00:00	2006-08-07 09:00:00
1747	06072109.d_01_0_0_1_1_0_0_0_1_1_1_0_1_1_1_0_1_1_0_0_1_1_1_1	2006-08-06 09:00:00	2006-08-08 03:00:00

別条件で検索

# 検索結果例 -生命線図-



2006年9月15日0時から9月18日0時までの検索結果

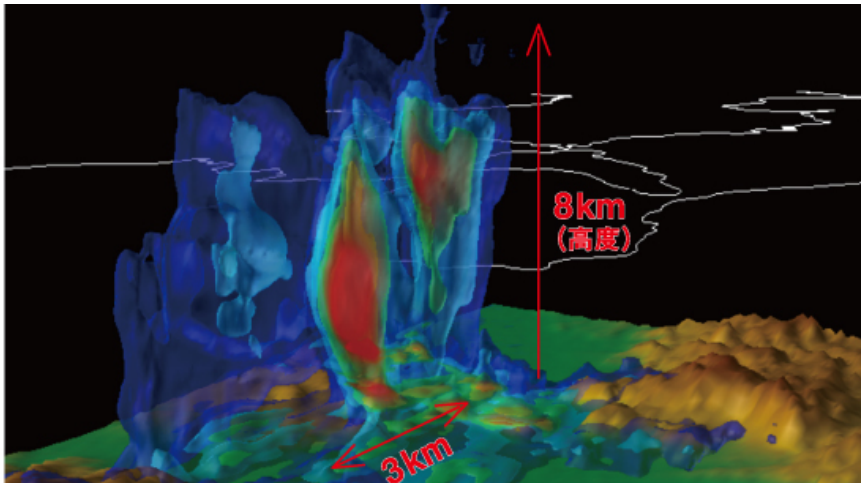


# 問題点と課題

- 意味的ラベルの精度が低く、意味的な知識発見が困難
  - →画像特徴だけでなく、形状、発生時期を属性に加え、教師つき学習を実施
- オブジェクト追跡過程の不安定性
  - 解のリセット(追跡の中断)
  - オブジェクトの新規発生への取り扱い
  - オブジェクト同士の融合
    - オブジェクト抽出の前処理、後処理の調整プロセスを付加

# 3Dフェーズドアレイ 気象レーダデータへの適用

- 2012年5月、大阪大学吹田キャンパスに設置
- 豪雨や竜巻をもたらす雨雲の詳細な3次元構造を10～30秒間隔で観測→直前予測への期待
- NICT佐藤氏、村田氏との共同研究
  - 3次元ホットスポット抽出手法開発(済み)
    - 強度も利用( $\propto$ 確率密度)度を利用→中心の精密決定



2012年7月26日 17:38:16の3次元降水分布 (NICT, 大阪大学、東芝)

<http://www.nict.go.jp/publication/NICT-News/1301/02.html>

# まとめ

- 時系列からホットスポット(オブジェクト)を抽出、要約することによって時空間知識発見を行うためのシステムを構築→気象画像、3Dフェーズドアレイ気象データに適用
  - オブジェクトの発達やオブジェクト間の相互作用を表す情報の要約、可視化を実現
  - 情報の精度を増すために、オブジェクト抽出、ラベル付け過程についてさらなる頑健性の向上が必要
  - また高次の時空間変動パターンの自動抽出にむけた検討が必要