情報学的視点による台風研究 および 台風画像データマイニング

国立情報学研究所 北本 朝展

本発表の概要

- 1. はじめに
- 2. 情報学的視点
- 3. 台風画像コレクション
- 4. 台風画像データマイニング
- 5. 気象学との接点
- 6. おわりに

研究の背景

- 私のこれまでの研究は、画像解析 やパターン認識といった情報学の研 究領域である。
- 2. 情報学の応用分野として気象学に興味をもっている。
- 3. 気象現象のような複雑な実世界の問題に貢献できるような、情報学的手法の研究を目的とする。

大規模データベースからの学習

世界には大量のデータが溢れている。

- ・地球観測データ、宇宙観測データ、医療・遺伝子データなどの科学的データ
- ・売上データ、金融データなどの経済データ

データの爆発的増大に伴い、これら大量のデータをデータベース化し、有用な情報を発見する技術が重要となる。

データマイニング

- ・大量のデータに潜む法則性を追究し、 データから情報、知識へ。
- 多様なデータマイニングアルゴリズム の組み合わせにより、システムと対話 しながら新しい知識を発見。
- ・データの取得から知識発見まで、 データの上流から下流までを体系的 に処理するという枠組み。

特に台風を対象とする理由

- 気象現象として、社会に与えるインパクトが最大である。
- ベストトラックなどの関連データがよく 整備されている。
- ・台風の雲パターンという時系列パターン情報の解析は、情報学の研究テーマとしても興味深い。

アプローチの比較

- 1. 気象学的アプローチ 気象現象の背後に潜む物理過程のモデル化
- 2. 人工知能的アプローチ(1980年代)気象観測データに対する人間の意思決定過程のモデル化
- 3. 情報学的アプローチ(本研究) 気象観測データに潜むデータの性質のモデル化

気象学における台風モデル

- 1. ドボラック法 予報官の長年の経験 に基づくヒューリスティックなモデル
- 2. 台風ボーガス 数値計算の都合に あわせた便宜的なモデル
- 3. シミュレーションモデル 気象学の理論体系に基づき数値計算にも適した物理過程のモデル(台風を明示的に表現するモデルではない)

本研究の台風モデル

- ・確率統計的モデルを基本とし、高レベルの処理として論理的構造も含む。
- ・データの出現やデータの共起関係、 データの状態遷移などが、完全にランダ ムではないと仮定し、それらの偏りを実 際の観測データから学習する。
- ・これらの性質を数学的に特徴づけ、確率統計的な意味での定量的な根拠をも つモデルを目指す。

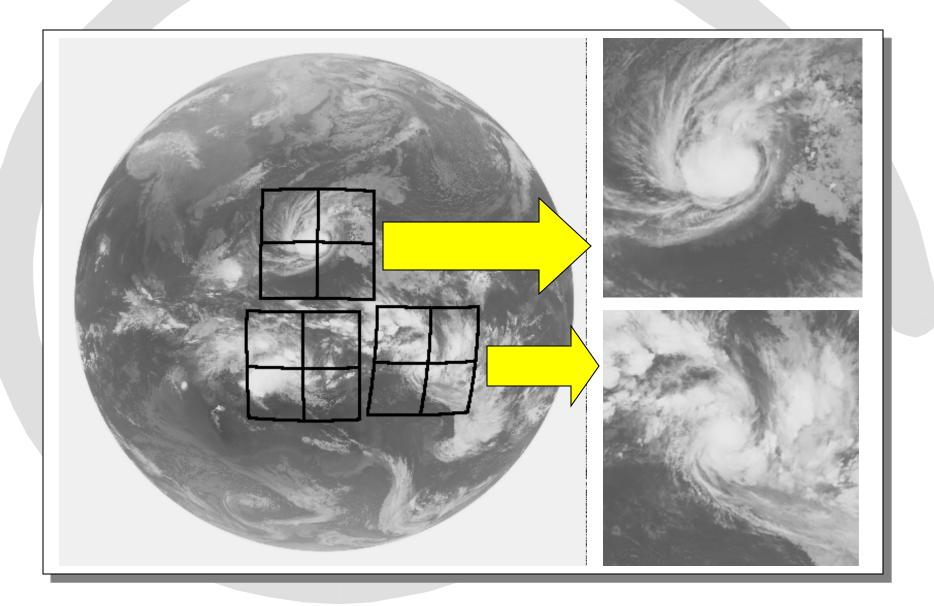
従来の研究

- ・流体力学に基づく動き解析(Zhou)
- Dynamic Link Architecture + Active Contours (Zhou)
- ・ 台風データマイニング(Lee)
- ・ファジー理論による台風予測(Hiraoka)
- ・いずれも情報学的手法の一応用例
- ・気象学的にはインパクトは無に近い?
- ・気象学と情報学の融合により、実際に インパクトを与える研究を目標とする

台風画像コレクションの概観

	北半球	南半球
ベストトラック		
観測機関	気象庁	Australia気象局
緯度範囲	赤道以北	赤道以南
経度範囲(東経)	100~180	90~170
台風画像コレクション		
台風シーズン数	6シーズン	5シーズン
台風系列数	136系列	62系列
総台風画像数	24500画像	9400画像
系列当り画像数	53~433	25~480

「ひまわり」からの台風画像生成



台風中心の位置合わせ

オイラー的表現

地球に対する固定的な座標から眺める

(例えばテレビの天気予報の雲画像)

ラグランジュ的表現

台風中心と共に動く座標系から眺める

台風中心位置の推定=ベストトラックで代用

ベストトラック

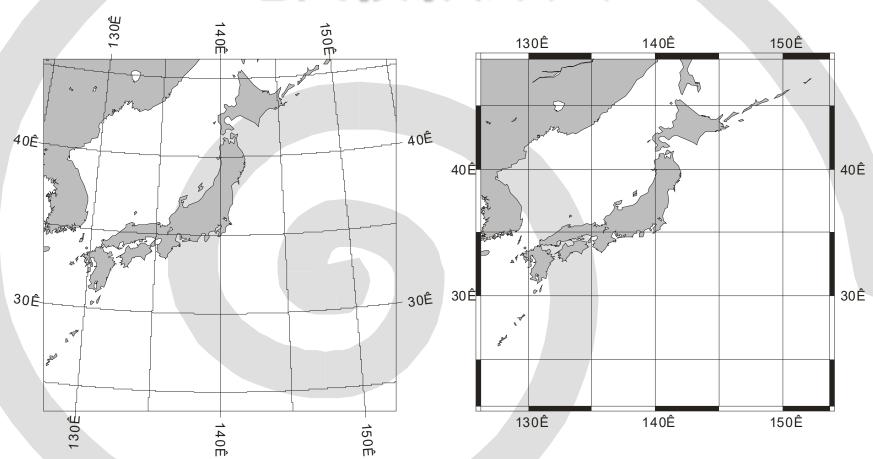
- ✓台風経路の最終解析結果。
- √3~6時間ごとの台風の中心位置、中心 気圧等が記録される。
- ✓専門家による検討を経た「正解データ」、 あるいは近似的な「グランドトルースデータ」とみなせる。
- ✓日本の気象庁およびオーストラリア気象 局のデータを活用する。

地図投影法(1)

ランベルト天頂等積図法 Lambert azimuthal equal-area projection

- 1. 台風の地理的な移動に関わらず台風の見かけの大きさは変化しない
- 2. 歪みが投影画像中心からの半径にしたがって増加するため、台風のような円形に近い物体では歪みが小さい

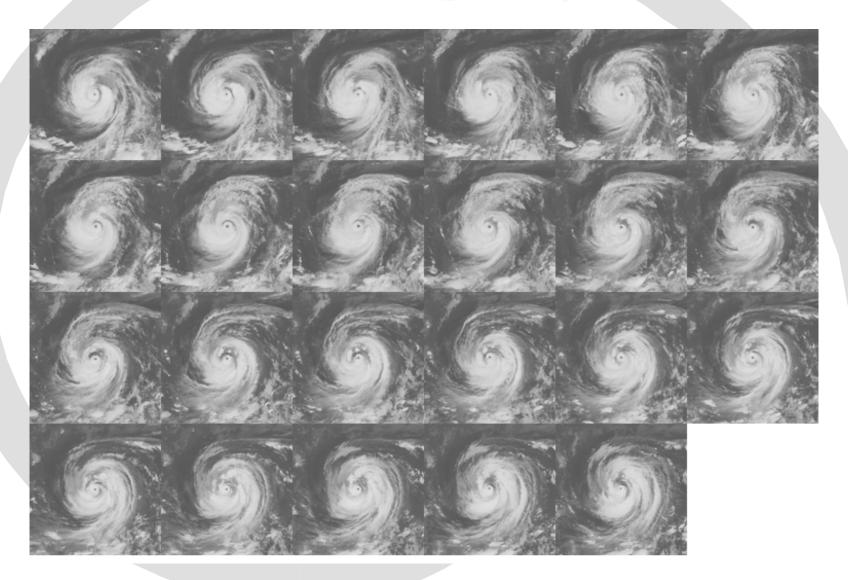
地図投影法(2)



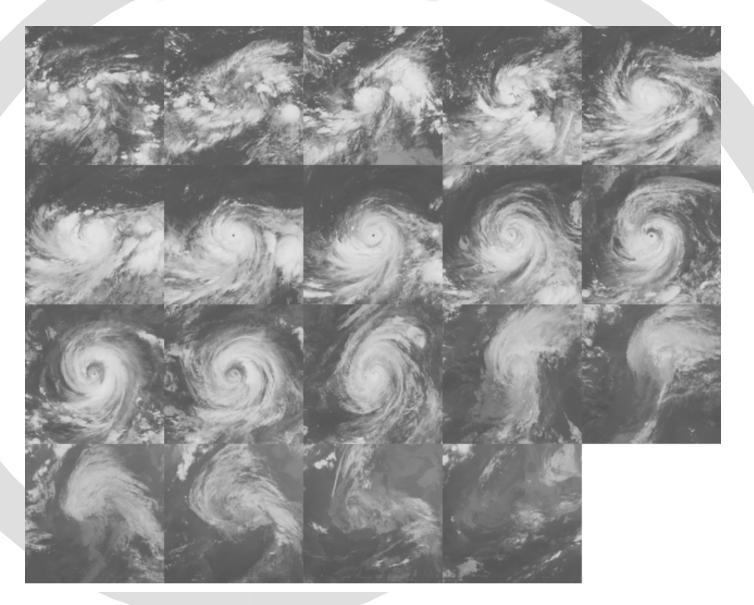
Well-Framedな台風画像生成

- ・台風画像中心が常に台風中心と一致・雲システム全体の
 - ・雲システム全体の 移動に関係なく、台 風雲パターンの変 化のみを注視
 - 北半球と南半球の 平均台風画像は、 南北反転となる

台風9713号のある1日



台風9713号の一生



台風9713号の一生

ひまわり画像の雲分類

GMS-5の赤外3チャネルを組み合わせた 雲画素の抽出・分類処理

- 1. 赤外画像を輝度温度に変換しチャネル間の差分も計算
- 2. 赤外の波長帯における水蒸気の吸収率の差を利用して雲領域を分離
- 3. 海面地上温度をロバスト推定し、雲の 輝度温度を高度に換算、雲の識別を おこなう

雲分類画像の時間変化

可視画像の利用

- ・台風画像コレクションでは基本的に可 視画像は利用しない。
- ・たとえ可視画像の利用で分類精度が 向上しても、昼間と夜間で分類精度 がばらつくという別の問題が生じる。
- ・ 画像コレクションに蓄積するデータの 質を一定に保つことが重要。

他の衛星データの利用

- 1. NOAA 東京およびバンコクの観測データが入手可能であり、「ひまわり」のデータを補完できる。「ひまわり」が稼動不能となる場合の備えとして重要?
- 2. TRMM 個別の台風研究に極めて有効だが、網羅的なデータセットとしては観測頻度が問題である。
- 3. GOES NESDIS/CIRAが同様のハリケーン画像アーカイブを構築している。

台風画像データマイニング

- 1. 空間的データマイニング
- 2. 時間的データマイニング
- 3. 時空間的データマイニング

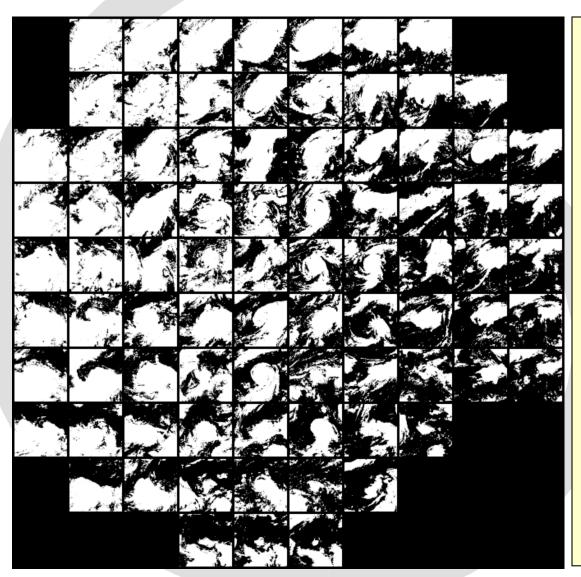
これらの例として、台風雲パターンの成分表現、台風の典型的雲パターンの解析、台風の日変化の解析、台風の一生の解析などについて述べる。

台風雲パターンの成分表現

- l. 64x64の(縮小)雲量台風画像を生成し、相関係数行列を計算
- ||. 主成分分析(PCA/EOF)の固有ベクトルを固有台風画像として用いる。
- |||. 多重解像度解析の適用

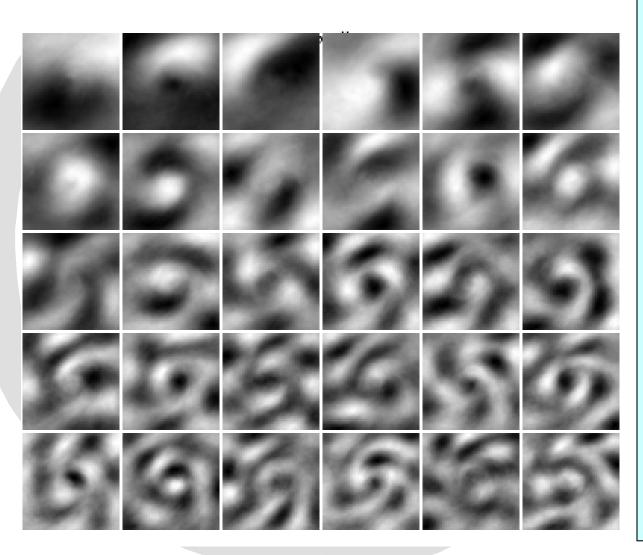
台風中心雲に対応する精円形と、中心から伸びるバンド構造に対応するらせん形が、台風雲パターンの主要形状

主成分分析(PCA)



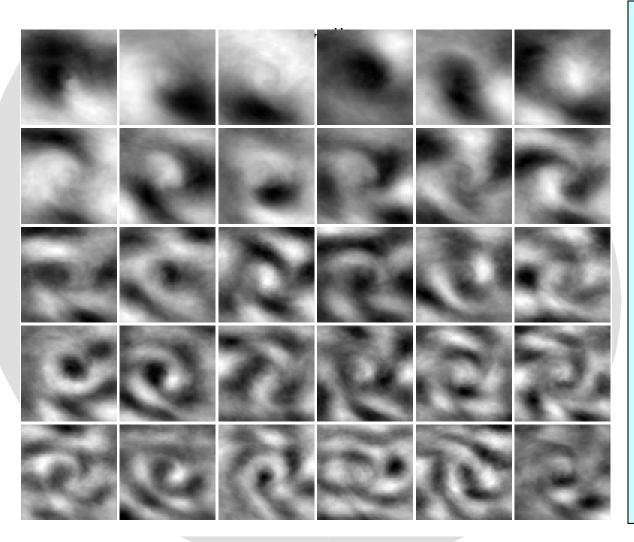
- ・分散が最大となる 方向を見出すため の数学的手法
- ・2つの主要な主成 分軸で張られる平 面上での雲パター ンの分布
- ・緯度方向の構造に関する多様なパターンが出現している。

北半球の固有台風画像



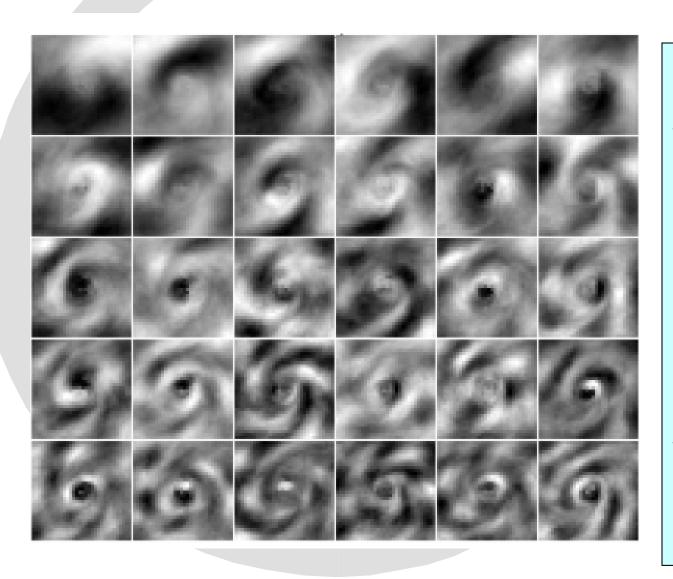
- ・第1~第3 主成分は、南 北方向の雲量 傾度を表現
- ・第4〜第8 主成分は台風 中心付近の雲 塊を表現
- ・第8主成分以降は曲率をもつバンド構造を表現

南半球の固有台風画像



- ・北半球の場合と同様の傾向
- バンド構造 が巻き込む 向きは北半 球とは逆
- ・両者は本質的に同一の気象現象

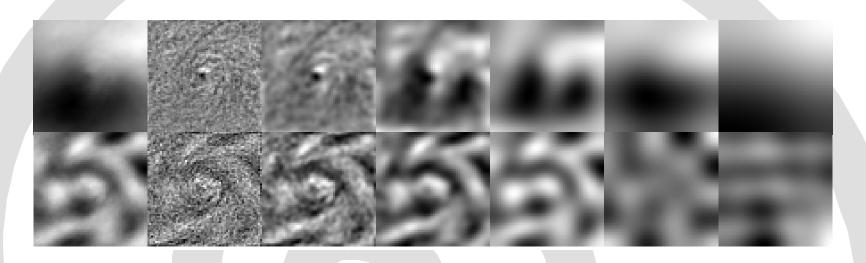
強い台風に対する固有台風画像



- ・眼付近の微 細構造がよ り明瞭
- ・バンド構造 の曲率がよ り大きい
- ・南北方向の雲量傾度が大きいのは同じ

固有台風の時間変化

固有台風画像の多重解像度解析



- 気象現象においてはスケールが重要
- Wavelet変換の一種A Trous変換を活用した固有画像の多重解像度解析
- ・上段は1000km以上のスケール、下段は~300kmスケールの現象

その他の成分表現

このような成分を用いた表現が流行しており、様々な手法が提案されている。

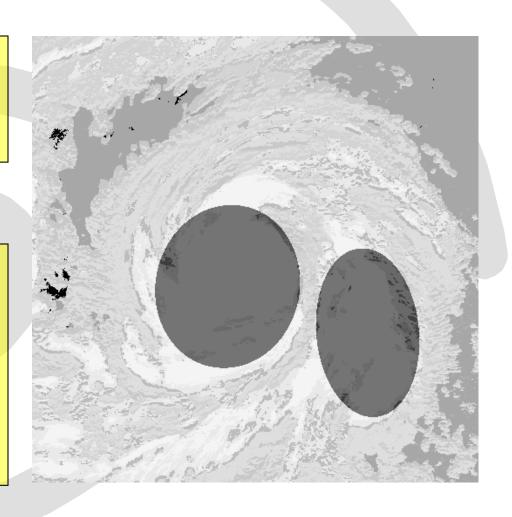
- ·独立成分分析(ICA)
- · 非負行列分解(NMF)
- ・カーネル主成分分析(KPCA)

これらの新しい手法を用いて、台風を成分の重ねあわせとして表現する。

台風のシンボリックな表現

台風の基本的形状の一つは楕円

精円を基本構成要素とし、台風雲パターンを精円の集合体として表現

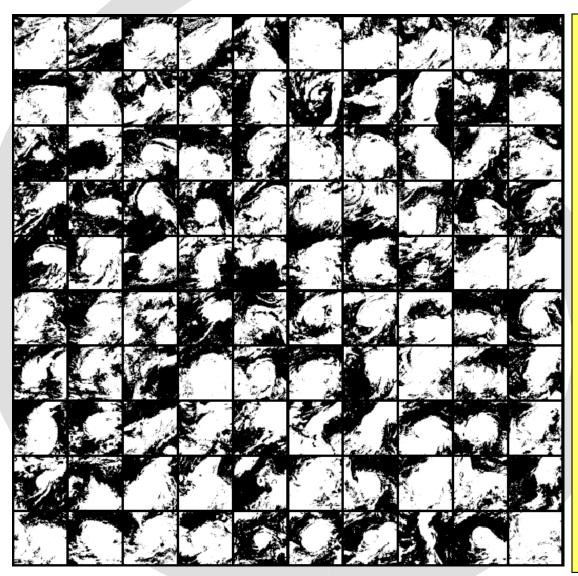


シンボリック表現の時間変化

台風の典型的パターン

- ・ドボラック法は「台風の典型的パターン」という考え方に基づく手法である。
- ・台風の雲パターンの確率密度関数が 特徴空間上で一様分布でなければ、 典型的なパターンが存在しうる。
- 特徴空間上のクラスタリング手法

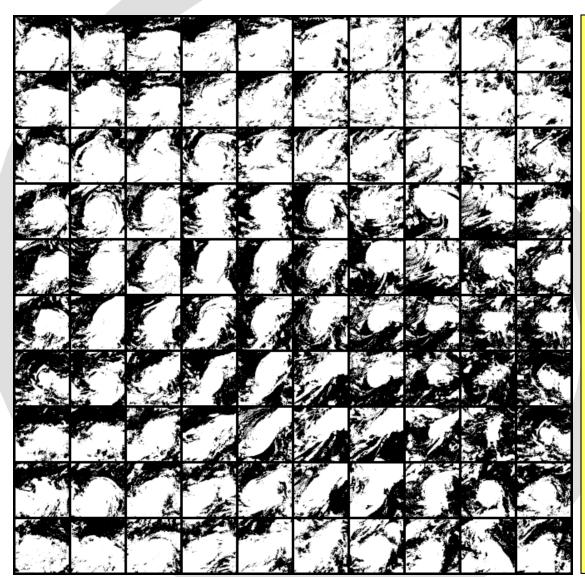
クラスタリング手法



・K平均法はクラスタリングのための簡便な手法である。

・雲パターンの 潜在的な構造 に関する直観 的理解は難しい。

自己組織化マップ



- ・クラスタ間に 位相構造が定 義されたクラス タリング手法。
- ・雲パターンの 連続的な遷移 に基づく潜在 的構造の理解 への手がかり。

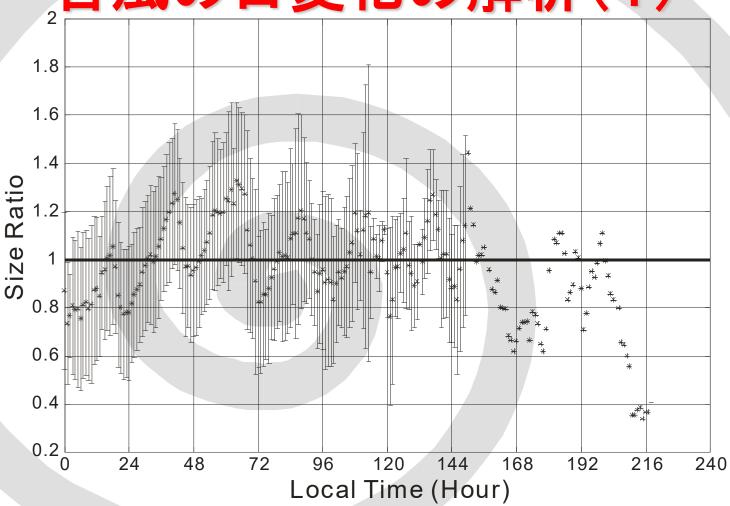
高次元特徵空間

- ・台風を高次元の特徴空間で表現する。
- データの分類 ニューラルネットワーク からカーネル法へ。
- ・低次元空間への射影 主成分分析からクラスタリングや生成的位相空間射 影法へ。
- ・近年の研究の進展によって、有望な方法が多数生まれてきている。

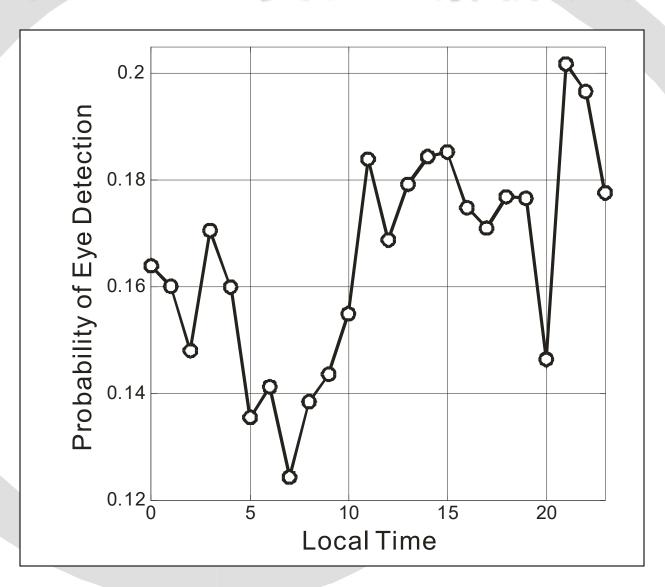
時間的データマイニング

- 1. これまで述べてきたのは、主に台風 の空間的データマイニング手法。
- 2. しかし時間的データマイニング、さらには時空間的データマイニング、た発展させることが本質的に重要。
- 3. 時空間的データマイニングは台風解析と台風予測をつなぐ掛け橋。

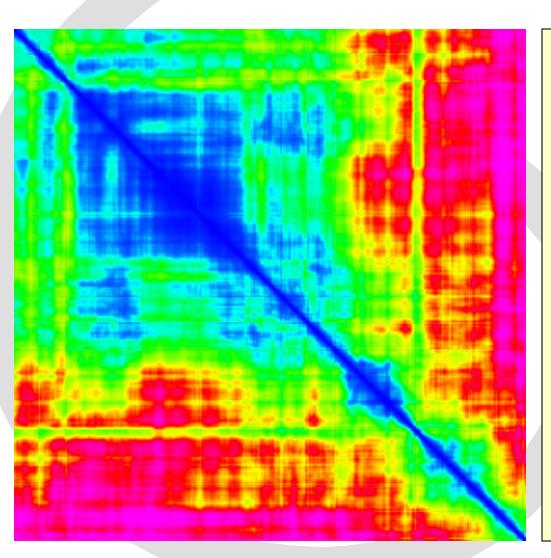
台風の日変化の解析(1)



台風の日変化の解析(2)

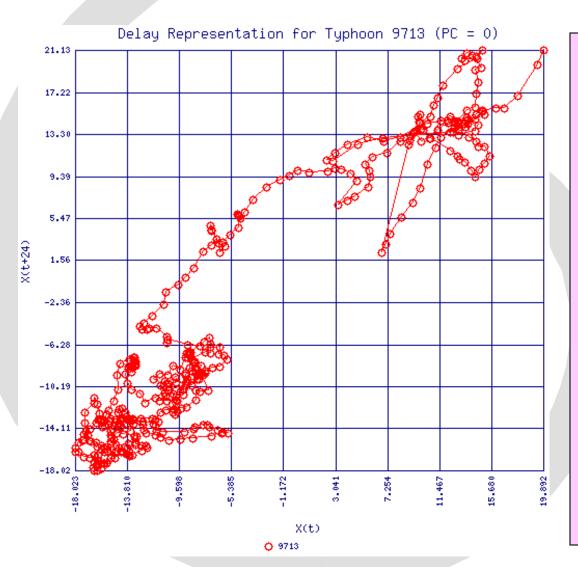


台風の一生の解析(1)



- リカーレンスプロットを用いた 白風の一生の可視化。
- 誕生後しばらく して雲パターン が安定。
- 終末期は雲パ ターンが急速 に変化。

台風の一生の解析(2)



- 台風雲パター ンの時間変化 を時間遅れ座 標系で可視化。
- パターンが安 定期から変化 を開始するそ のきっかけは 何?

類推法に基づく予測

ある台風画像を検索キーとして提示し、台風データベースから類似した画像を検索

課題

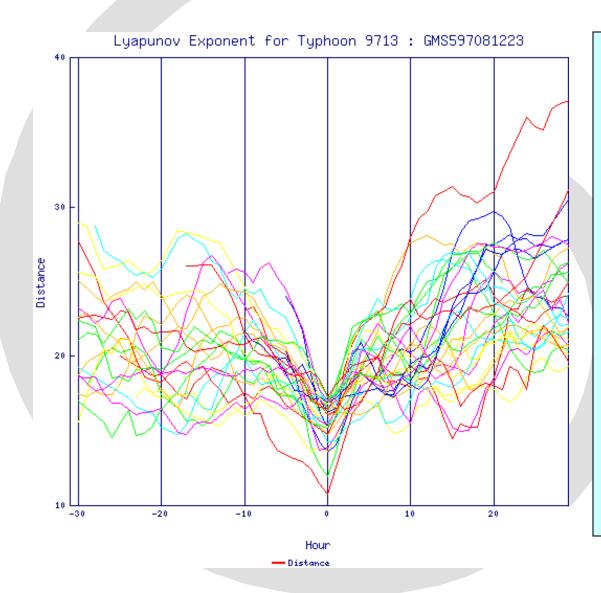
- 1. 検索キーと類似した雲パターンをもつ台風の画像は検索できるか?
- 2. 過去の類似した雲パターンに基づく台風予測は可能だろうか?

類推法(Method of Analogues)

気象パターンの類似度(誤差)が時間の経過でどのように減少(拡大)するか?

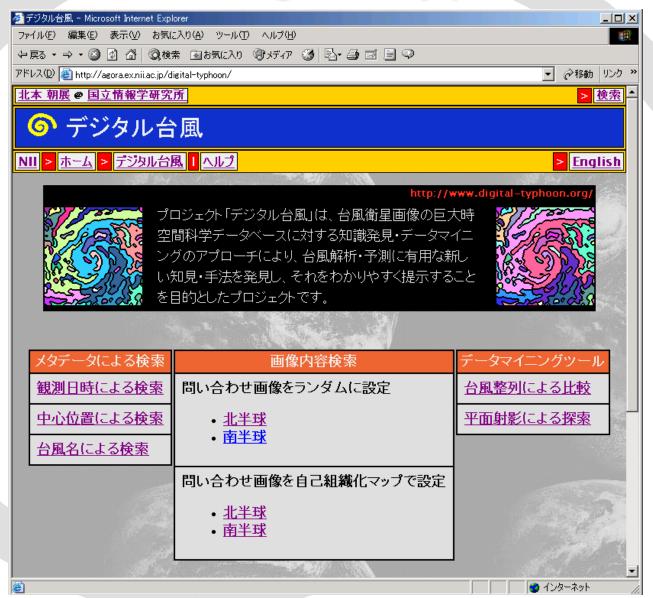
- ・誤差が2倍に拡大するのは、2.5日から8日=予測可能性の目安
- ・大気力学系の非線形性=カオス
- 類似パターンの誤差が拡大する速度=リヤプノフ指数

台風の予測可能性



- 基準時刻でパターンが類似していても、その類似性は急速に減少。
- 雲パターンの 類似性は予 測において本 質的か?

http://www.digital-typhoon.org/

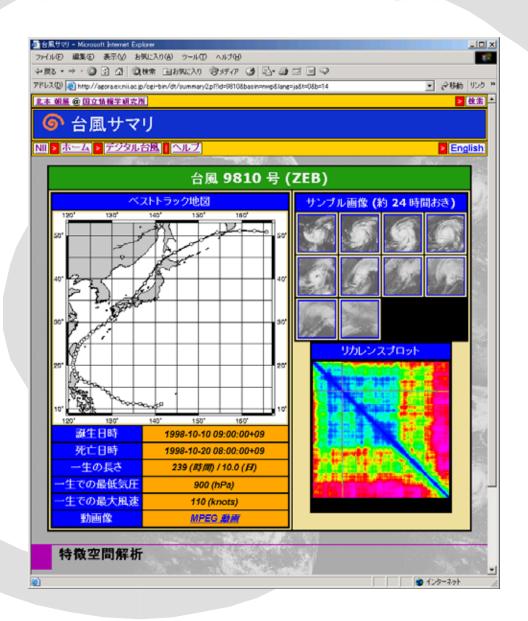


台風画像検索



- ・固有空間の次元数は86(累積寄与率80%)
- ・ロバスト性に優れる二奇妙な検索結果は比較的 少ない
- ・台風眼などの微細構造を検索対象とするのはほぼ不可能

台風の一生の表示



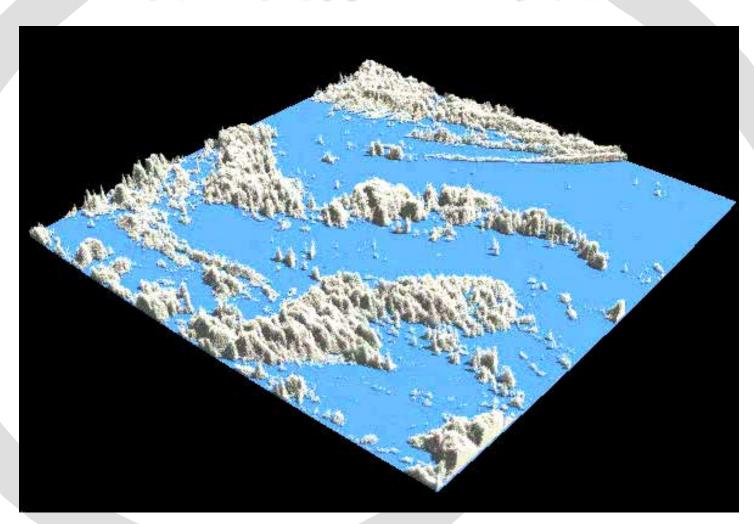
雲パターンの平面への射影



複数の台風の整列



台風画像の3D表示



気象学との接点

- これらの知見は、気象学にすぐに活用できるような知識にはなっていない。
- ・ 気象学の枠内では解決が難しく、情報学的視点が有効となるような問題とは何だろうか?
- ・予報官の経験を、どのように学習システムに組み込むことができるか?

台風に関する難問

急発達の解析と予報: 雲パターンの特徴から、急発達の予兆を発見できないか。

迷走台風の解析と予報: 雲パターンの非 対称性などを活用して、迷走台風の進路 予測を向上できないか。

台風発生の解析と予測: 熱帯収束帯に散 在する大量の雲塊から、台風に発達しそ うな雲塊群を予想できないか。

研究の方向性と評価

- 1. 台風予測の精度の向上という形で成果を計測できるのが理想。
- 2. 台風に関する新しい視点を提供するだけでも、それが真に新しければ価値がある。
- 3. 台風に関する専門家支援システムとしての有効性を計測する。
- 4. 一般の人々に対する台風情報提供システムとしての有効性を計測する。

おわりに

- 1. 3万件以上の高品質な台風衛星画像コレクションを構築した。
- 2. 情報学的視点による台風画像データマイニングの試みについて紹介した。
- 3. 台風雲パターンを表現する適切な数学的モデルはまだ研究途上である。
- 4. 情報学的視点で眺めても、時空間データの典型例である台風は興味深い研究対象であり、今後も研究を続けたい。