

ネットワークに基づく 分散型地球環境データベースの構築

国立情報学研究所

北本 朝展

はじめに

地球環境データを多くの人々が容易に入手し活用するためにはどんな技術が必要だろうか？

ネットワーク技術

大量の地球環境データをネットワーク上で効率的に配信および検索するための技術

データベース技術

大量の地球環境データに埋もれた重要な情報を効果的に発見するための技術

概要

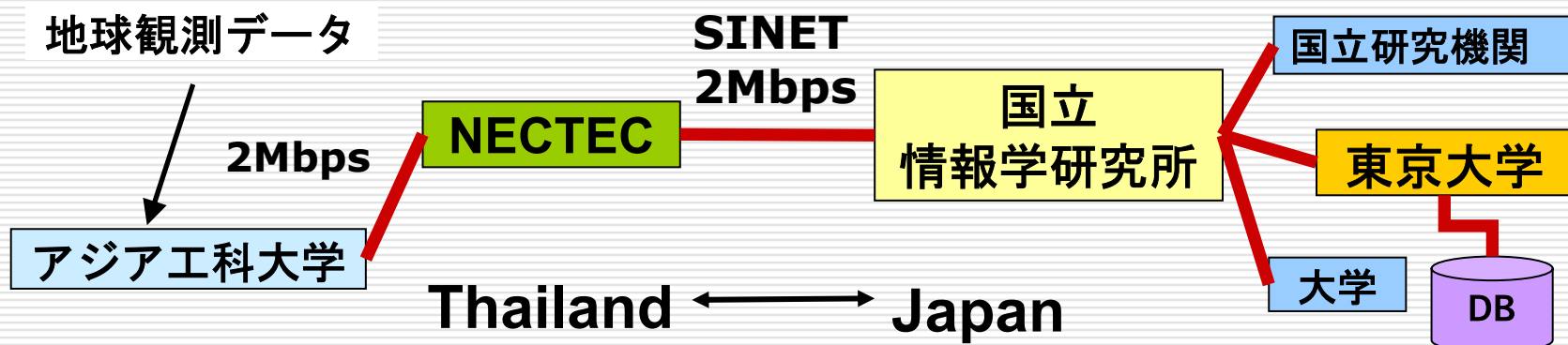
1. 地球環境データのネットワークング技術

1. **SINET**を用いた大規模衛星データの準リアルタイム交換実証実験
2. 分散データベースのメタサーバ・**XML**に基づくデータ定義・検索言語
3. 地球環境データへのアクセスパターンに基づくキャッシング
4. 多重解像度表現に基づく地球環境データのための画像符号化

2. 地球環境データの分散型データベース技術

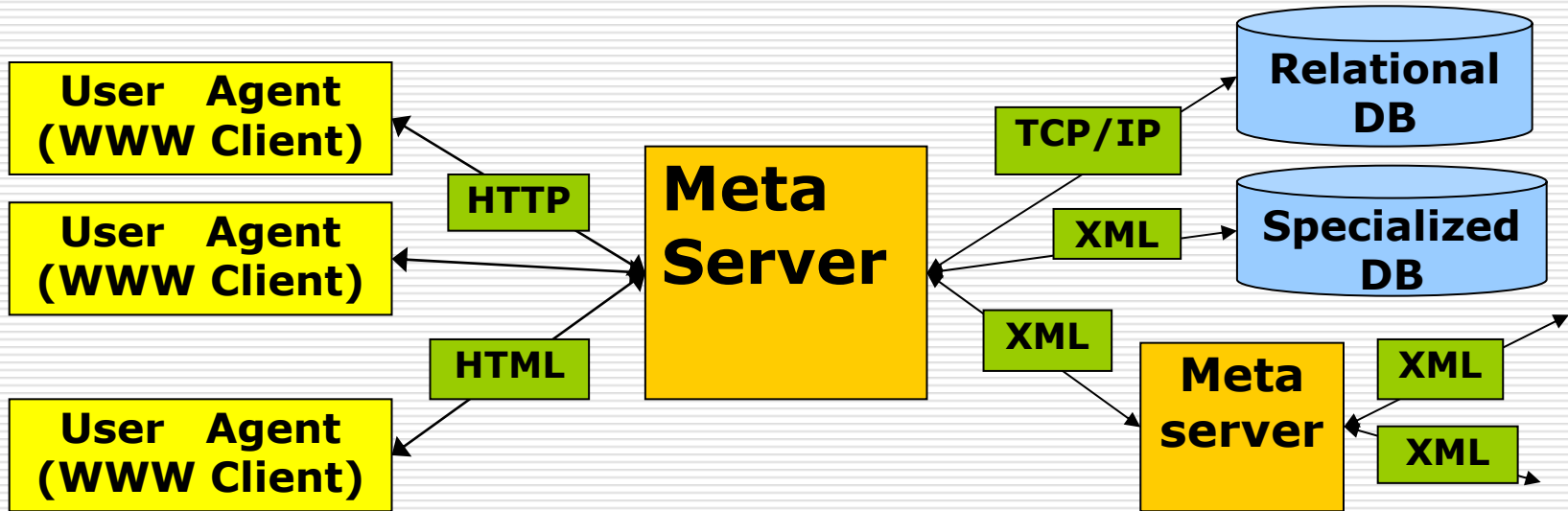
1. 地球環境データベースの内容検索技術
 2. 地球環境データベースのデータマイニング技術
-

SINETを用いた大規模衛星データの準リアルタイム交換実証実験



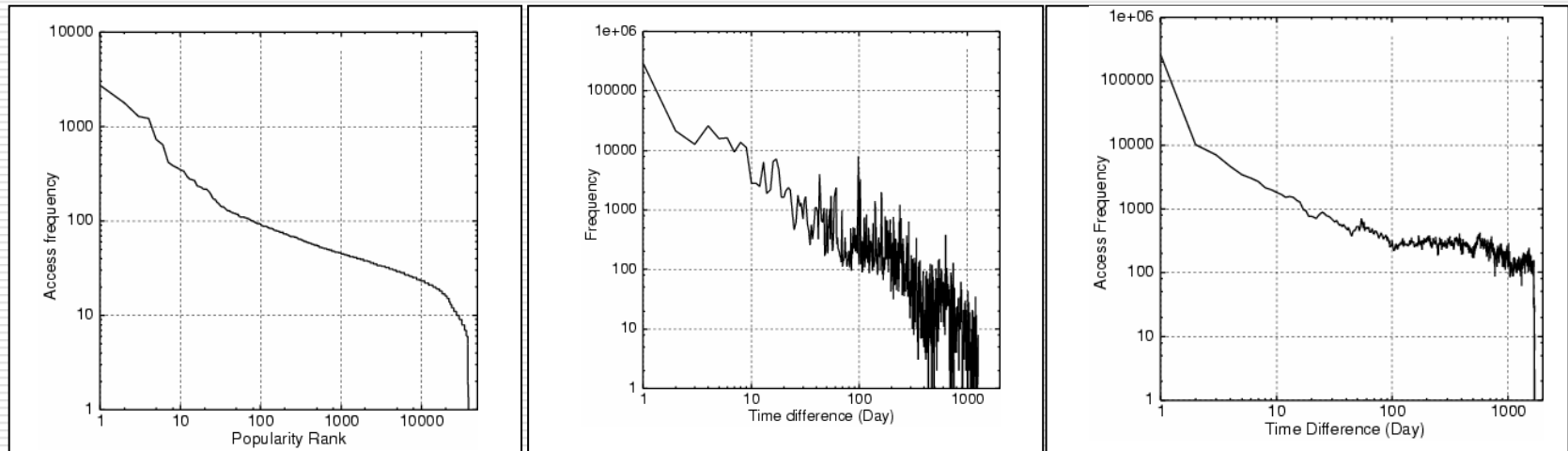
- アジア工科大学で受信するNOAA/MODIS衛星データを準リアルタイムに交換するためのネットワーク基盤を構築し、実証実験を開始。
- タイおよび日本国内のネットワーク設備を拡充。
- $130\text{KBytes/sec} = 11.2\text{GByte/day}$ の速度で地球環境データを国際的に交換共有する環境を実現。

分散データベースのメタサーバ・XMLを用いたデータ定義検索言語



- XML構文を用いたデータ定義・検索言語を、地球環境データに適した形に制定。
- 台風画像検索サーバとしてWWW上に実装。

地球環境データへのアクセスパターンに基づくキャッシング技術



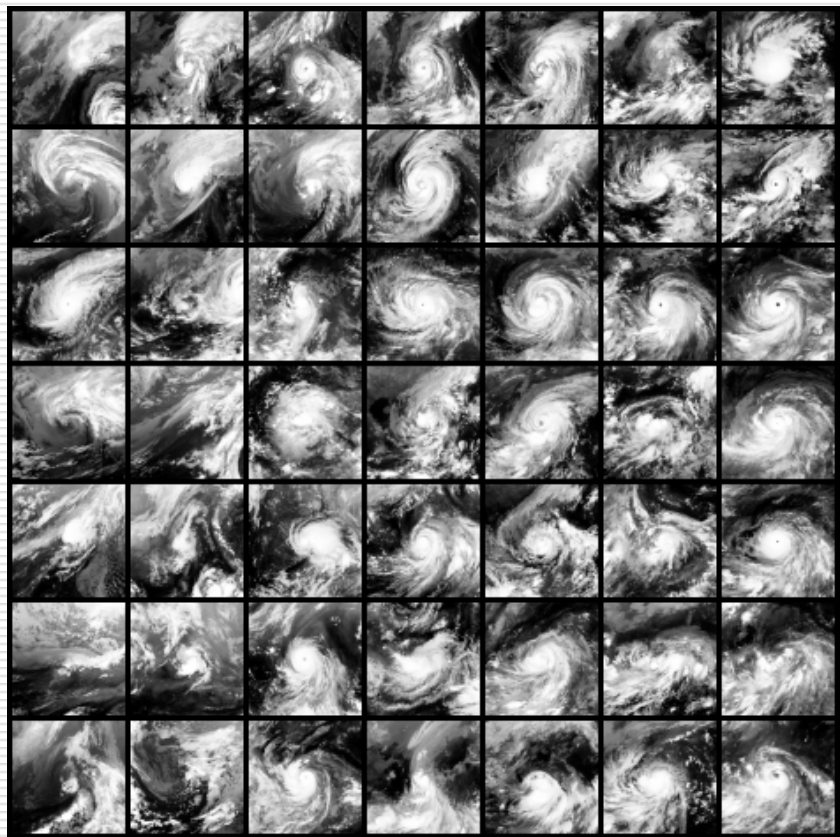
1. データに対するアクセス頻度の偏り
2. データに対するアクセス時刻間隔の偏り
3. データ受信時刻とアクセス時刻の間隔の偏り
4. データの解像度に対するアクセス頻度の偏り

地球環境データベースの内容検索技術

1. データの意味内容を画像解析で抽出し検索キーを生成
2. 気象学的・社会的に最重要の気象現象である「台風」衛星画像データベースを中心に研究

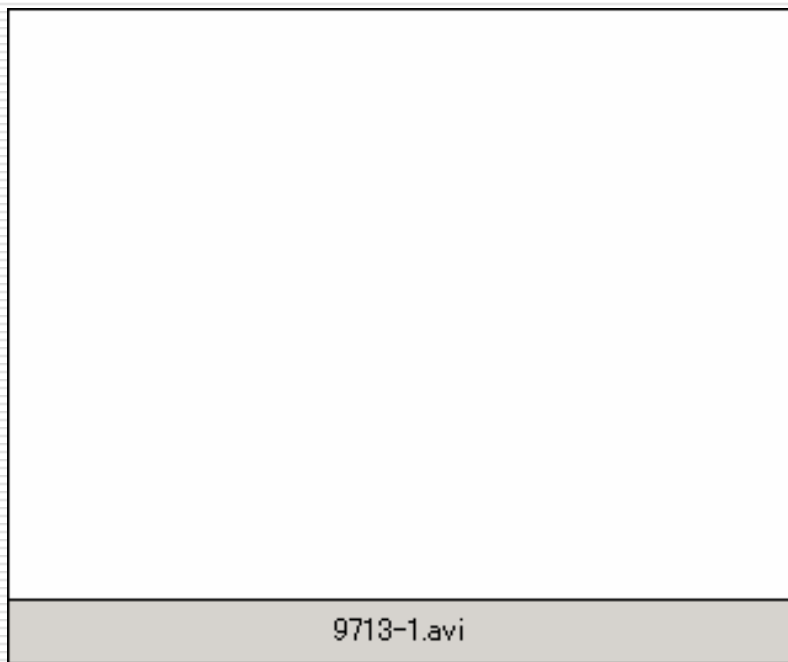
5 意味レイヤ	画像の意味内容
4 関係レイヤ	画像領域間の関係
3 領域レイヤ	画像領域単位の表現
2 画素レイヤ	画素単位の表現
1 観測レイヤ	画像観測の表現

台風画像コレクションとは？



1. 台風の中心位置に基づく高品質の画像コレクション。
2. 北半球・南半球の衛星画像計**34000**件の大規模画像データベース。
3. 衛星画像データマイニングの基礎データ。

台風画像コレクション



「ひまわり」赤外画像



画素単位雲分類画像

分散型台風画像内容検索システム

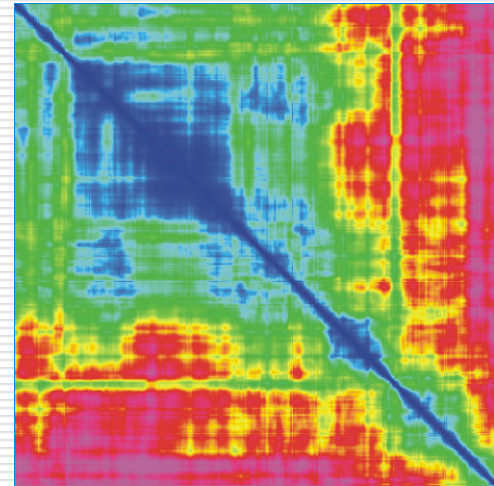
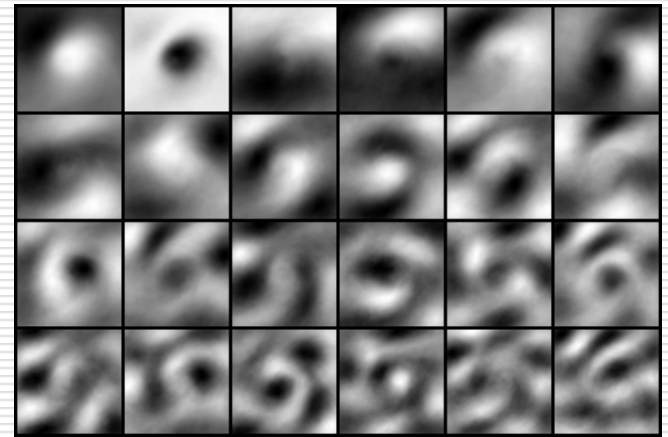


- 画像内容検索技術とXML構文によるデータ検索言語の研究とを融合し、WWWインタフェースによる検索システムを実現。
- 地球環境データベースにおいても、画像内容に基づく検索システムは重要となる。

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

地球環境データベースのデータマイニング

- 台風画像コレクションという巨大科学データベースに潜む規則性や不規則性を発見。
- クラスタリング・自己組織化マップ・主成分分析・カオス時系列分析・サポートベクターマシンなどの機械学習手法を画像解析手法と融合。
- 気象学的に意味のある結果を得ることが課題。



研究の今後の展開

ネットワーク技術

1. 地球環境分散メタサーバシステム
2. XML構文を用いた地球環境データに適したデータ検索プロトコルの制定
3. システムの実用化

データベース技術

1. 地球環境データマイニング技術
 2. 地球環境データおよび関連ソフトウェアの公開
 3. 災害対策など実社会問題との連携
-

おわりに

- 本研究では、地球環境データを容易に入手し活用するための技術として、**ネットワーク技術**および**データベース技術**に関する研究を進めた。
- 特にデータベース技術については、台風画像データマイニングを中心に大きな成果を得た。
- またネットワーク技術との融合により、分散地球環境画像データベースのプロトタイプを実現した。
- 5年間にわたる支援を賜りました通信放送機構および担当者の方々に深く感謝いたします。