

# MPEG-7を利用したウェアラブルカメラ 映像の索引付け手法

兵 清弘<sup>†</sup>、天笠 俊之<sup>†</sup>、吉川 正俊<sup>†‡</sup>、植村 俊亮<sup>†</sup>

<sup>†</sup> 奈良先端科学技術大学院大学 情報科学研究科

<sup>‡</sup> 国立情報学研究所 ソフトウェア研究系

# 研究の背景

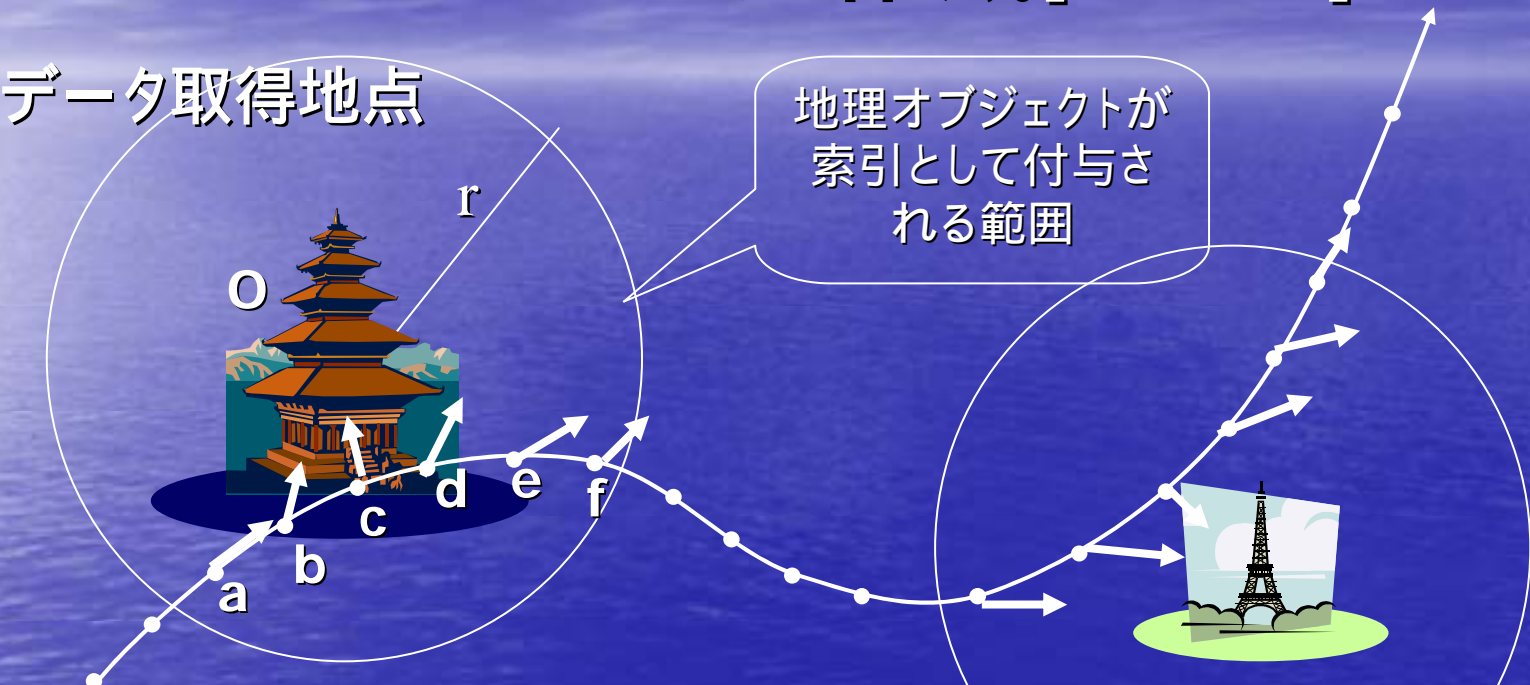
- ウェアラブルコンピュータ
- 記録媒体の小型化、大容量化
- 日常生活の常時撮影が可能となる
  - ビデオ日記、記憶補助、人生そのものも...



大量の映像から所望の場面をすばやく検索したい  
But 手作業での索引付けは困難

# 地理オブジェクトに基づく検索 およびダイジェスト作成[上田等]

- GPSデータ取得地点



[距離による重み、向きによる重み]と表現

$$W(a,O)=[0.1,0.5]=0.3$$

$$W(b,O)=[0.4,0.8]=0.6$$

$$W(c,O)=[0.9,0.9]=0.9$$

$$W(d,O)=[0.8,0.2]=0.5$$

$$W(e,O)=[0.4,0]=0.2$$

$$W(f,O)=[0.1,0]=0.05$$

地理オブジェクトO(五重塔)の重要度 合計 2.55

ジャイロ

角度データ

GPS

位置データ

メタデータが重要

カメラ

ビデオデータ

t

地理データ

重要度解析

重要シーンの特定



再生場所の決定



ダイジェストの作成、呈示



# MPEG-7の利用

- 映像のメタデータの記述形式を定めたもの
- XML schemaがベース言語
- ハイレベル、ローレベルな内容記述のための記述ツールが規定
- 国際標準規格

誰がどのように  
メタデータを作成  
す

メタデータを  
MPEG-7ファイルと  
して記述するには  
人手がかかる

- ・イチローがヒットを打ったシーンが見たい
- ・奈良公園で鹿と戯れる映像をもう一度

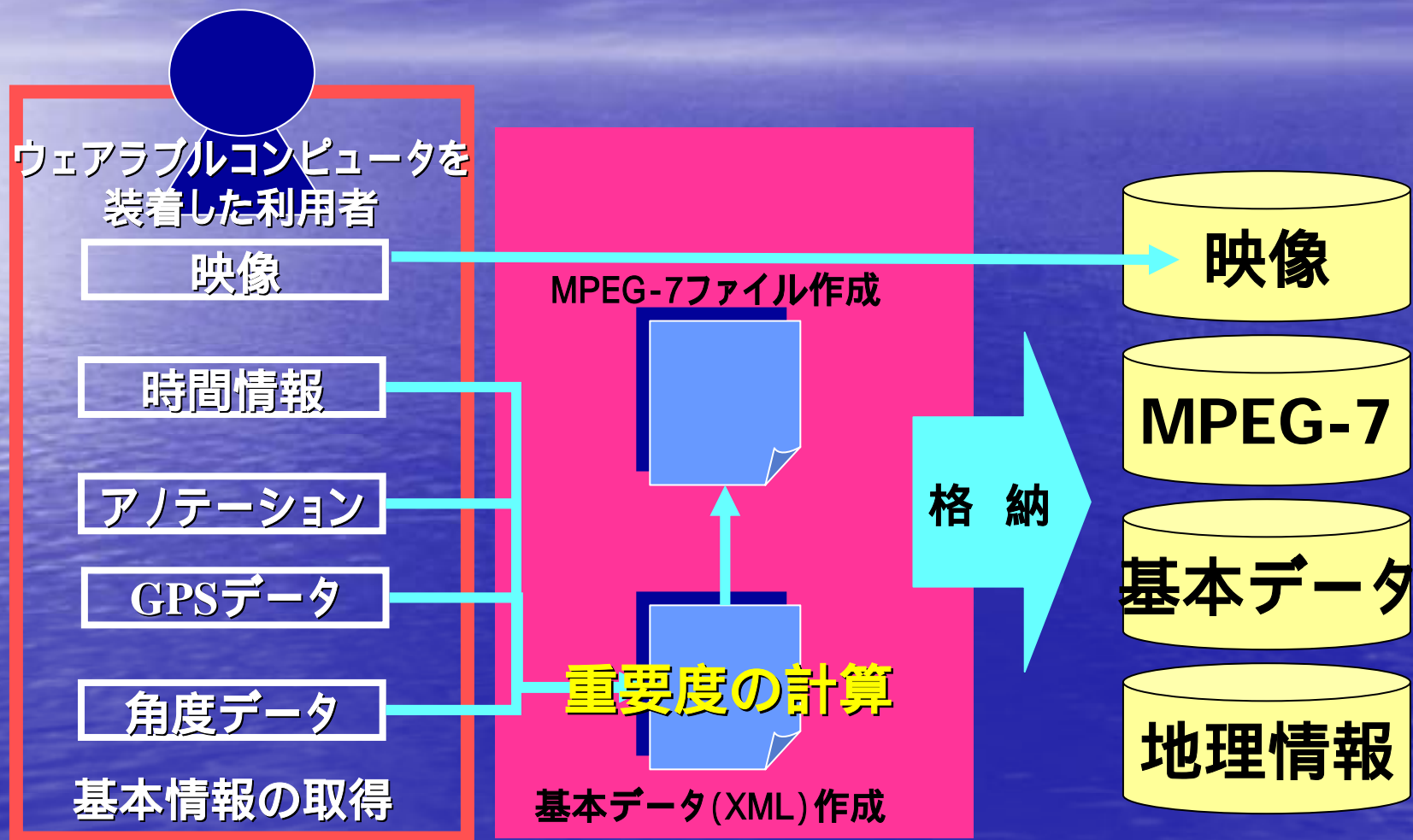
# 研究の目的

- ウェアラブルカメラ映像に対して自動で索引付けを行い、MPEG-7ファイルを生成
- 上記のデータを利用した検索システム  
およびダイジェスト作成
  - MPEG-7ファイルを動的に生成

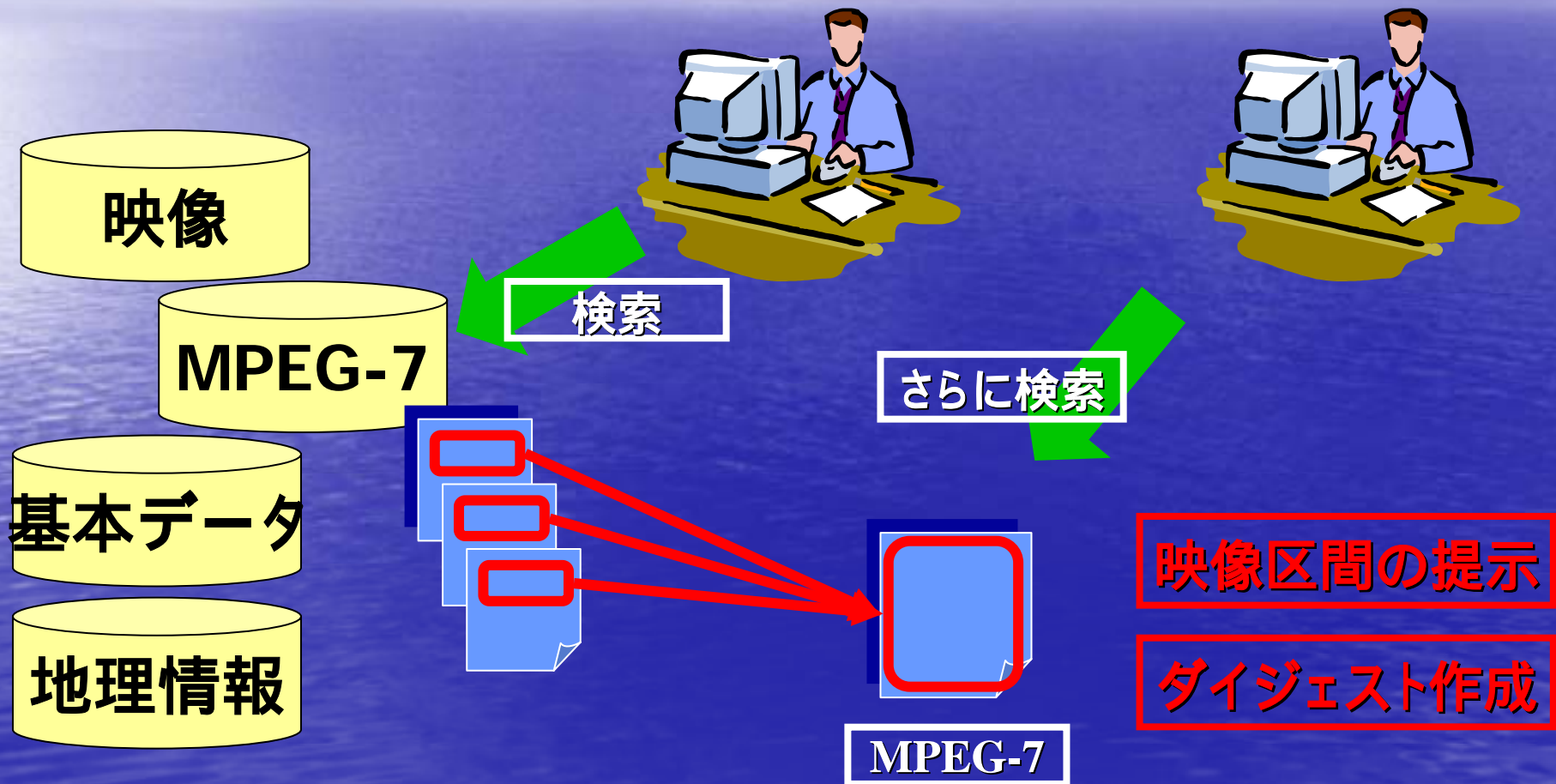


あのときの  
鹿の映像は？

# 本研究の流れ(索引付け)



# 本研究の流れ(検索・ダイジェスト作成)





# まとめと今後の課題

- MPEG-7を利用したウェアラブルカメラ映像の索引付け手法を提案
- MPEG-7を利用した検索手法の提案
- 実装および実験
- ユーザインタフェースの構築
  - 検索方法及び結果の提示方法
- 領域、高さを持った地理オブジェクトへの対応



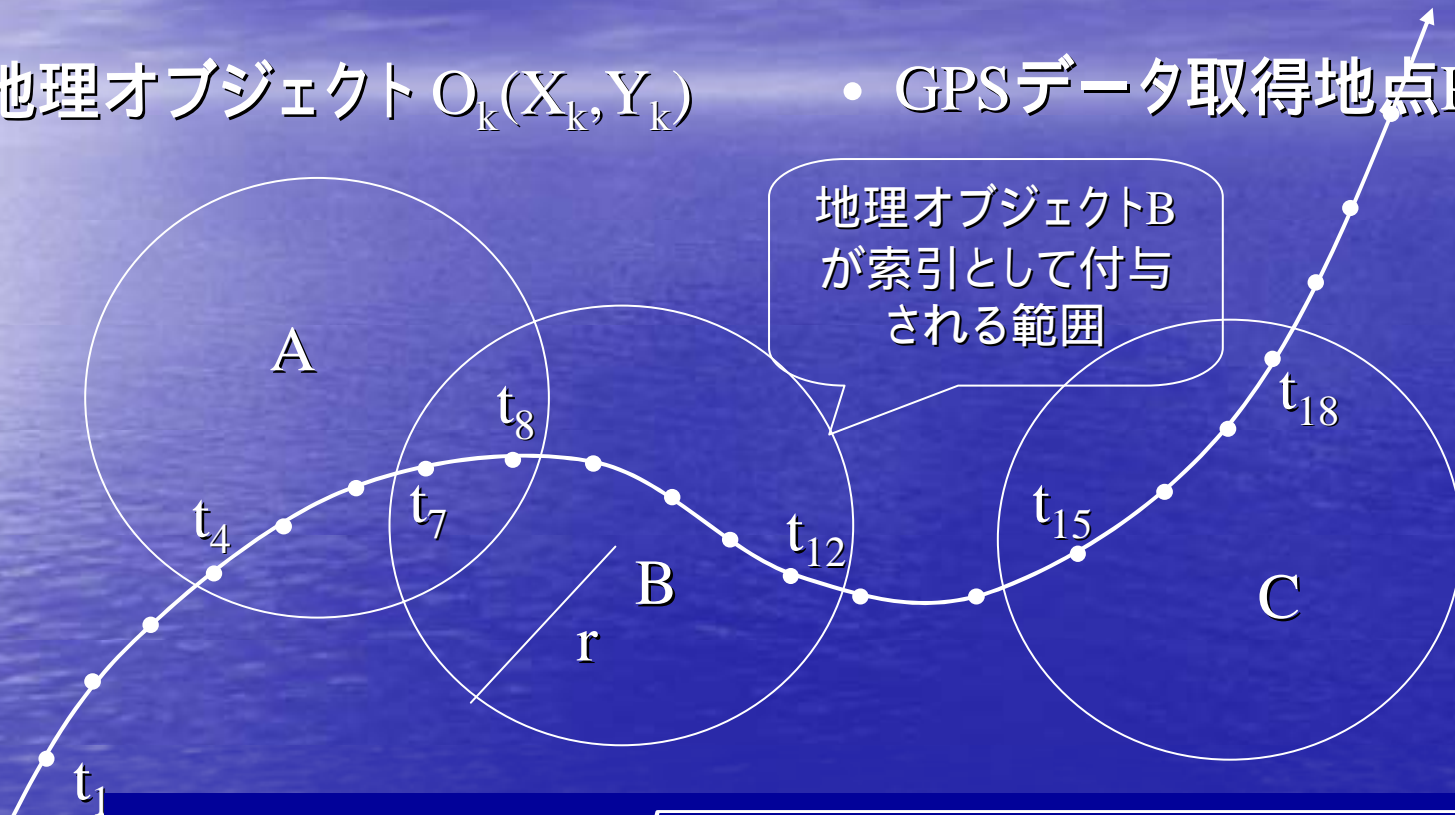
# 地理オブジェクトの重要度

- 距離による重みの計算
  - 地理オブジェクトが利用者とどれだけ離れているか
- 向きによる重みの計算
  - 利用者が地理オブジェクトの方向を向いているか
- 重みの統合
  - ある時刻 $t$ における地理オブジェクトの重要度
- “位置情報と地理情報を用いたウェアラブルカメラ映像のダイジェスト作成” [上田等]

# 地理オブジェクトの重み (距離)

地理オブジェクト  $O_k(X_k, Y_k)$

• GPSデータ取得地点  $P_t(x_t, y_t)$



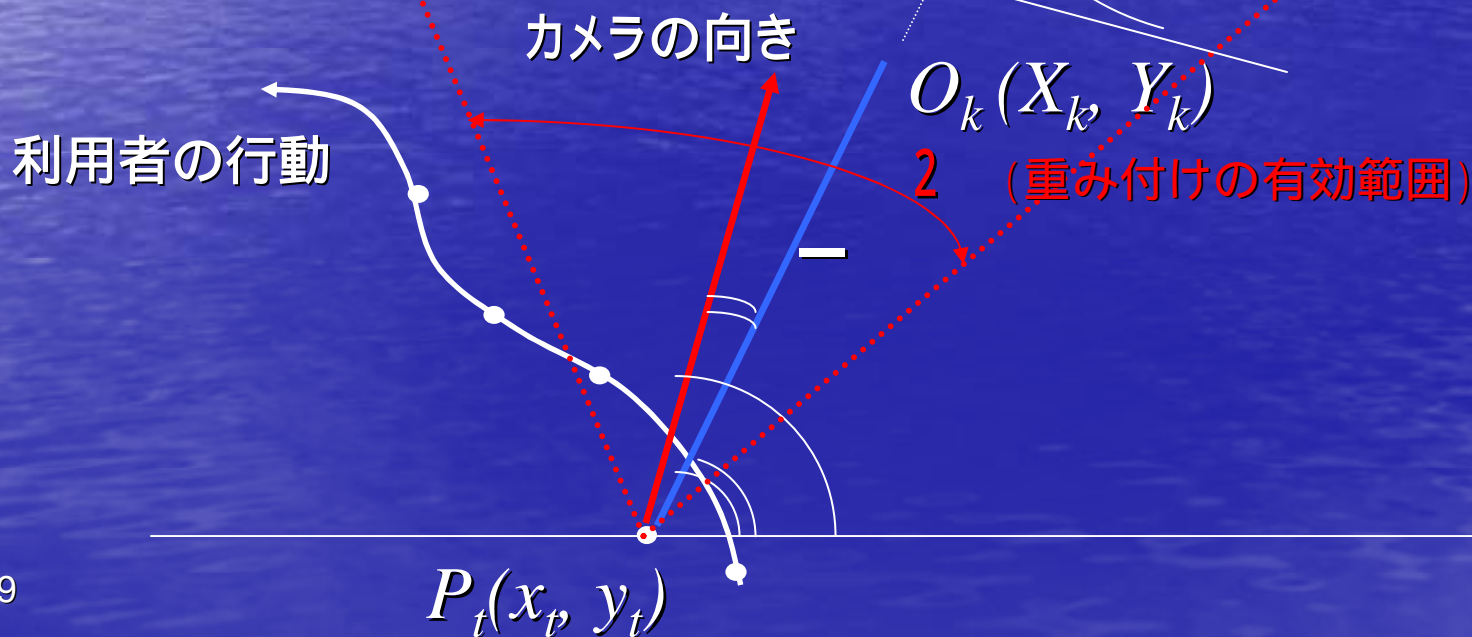
$$d(P_t, O_k) = \sqrt{(X_k - x_t)^2 + (Y_k - y_t)^2}$$

$$W_{\text{distance}}(P_t, O_k) = \frac{1}{d(P_t, O_k)}$$

# 地理オブジェクトの重み(向き)

- 地理オブジェクト
- GPSデータ取得地点

正規分布に  
当てはめた  
重み付け



# 地理オブジェクトの重要度

- 向きによる重みの計算

$$w_{\text{direction}}(P_t, O_k) = \frac{1}{1-C} \times \exp\left(-\frac{(\alpha_t - \theta(P_t, O_k))^2}{2\sigma^2}\right)$$
$$-\frac{C}{1-C} \left( \text{ただし } C = \exp\left(\frac{\beta^2}{2\sigma^2}\right) \right)$$

- 重みの統合

(ある時刻 $t$ における地理オブジェクトの重要度)

$$w(P_t, O_k) = zw_{\text{distance}}(P_t, O_k) + (1-z)w_{\text{direction}}(P_t, O_k)$$

$z(0 \leq z \leq 1)$ : 利用者が指定

# 地理オブジェクトを基にした映像分割

- 重要度および取得したデータをXML形式で記述(基本データ)

```
<?xml version="1.0" encoding="Shift_JIS"?>
<Point time="2002-02-08T14:35:04">
  <GeographicPoint longitude="135.442086"
latitude="34.438085" direction="90"/>
  <Place value="100">奈良先端科学技術大学院大学</Place>
  <Place value="90">NEC</Place>
</Point>
:
<Point time="2002-02-08T14:35:24">
  <GeographicPoint longitude="135.442099"
latitude="34.438214" direction="120"/>
  <Place value="90">奈良先端科学技術大学院大学</Place>
  <Place value="70">高山サイエンスプラザ</Place>
  <Place value="65">NEC</Place>
</Point>
```

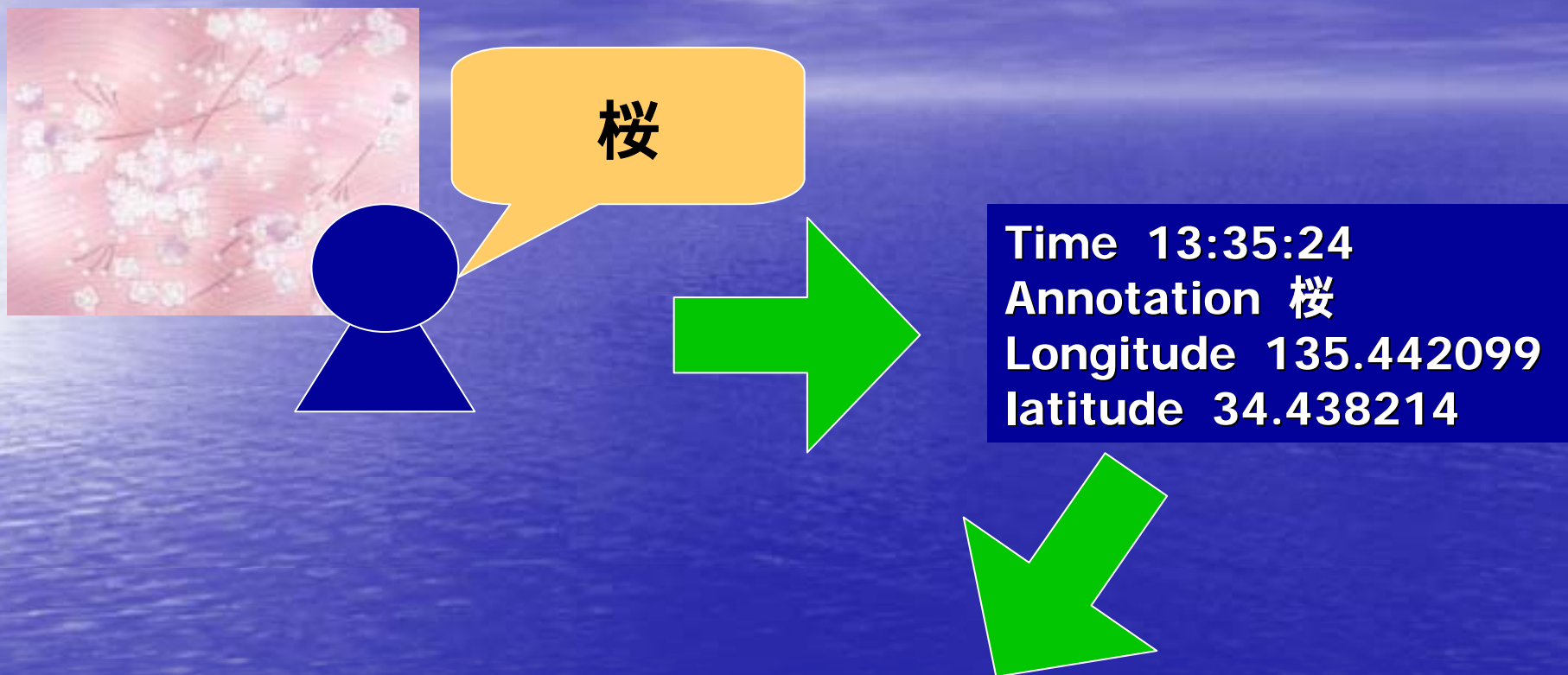
# 地理オブジェクトを基にした映像分割

- 地理オブジェクト毎をまとめてMPEG-7ファイルへ記述

```
<AudioVisualSegment id="place1">  
  <GeographicPointType longitude="135.442099"  
latitude="34.438214" />  
  <PointOfView viewpoint="奈良先端科学技術大学院大学">  
    <Importance><Value>重要度</Value></Importance>  
  </PointOfView>  
  <MediaTime>  
    <MediaTimePoint>  
      2002-02-08T14:35:24:0F30  
    </MediaTimePoint>  
    <MediaIncrDuration timeUnit="PT1S">  
      374  
    </MediaIncrDuration>  
  </MediaTime>  
</AudioVisualSegment>
```



# アノテーションを基にした映像分割

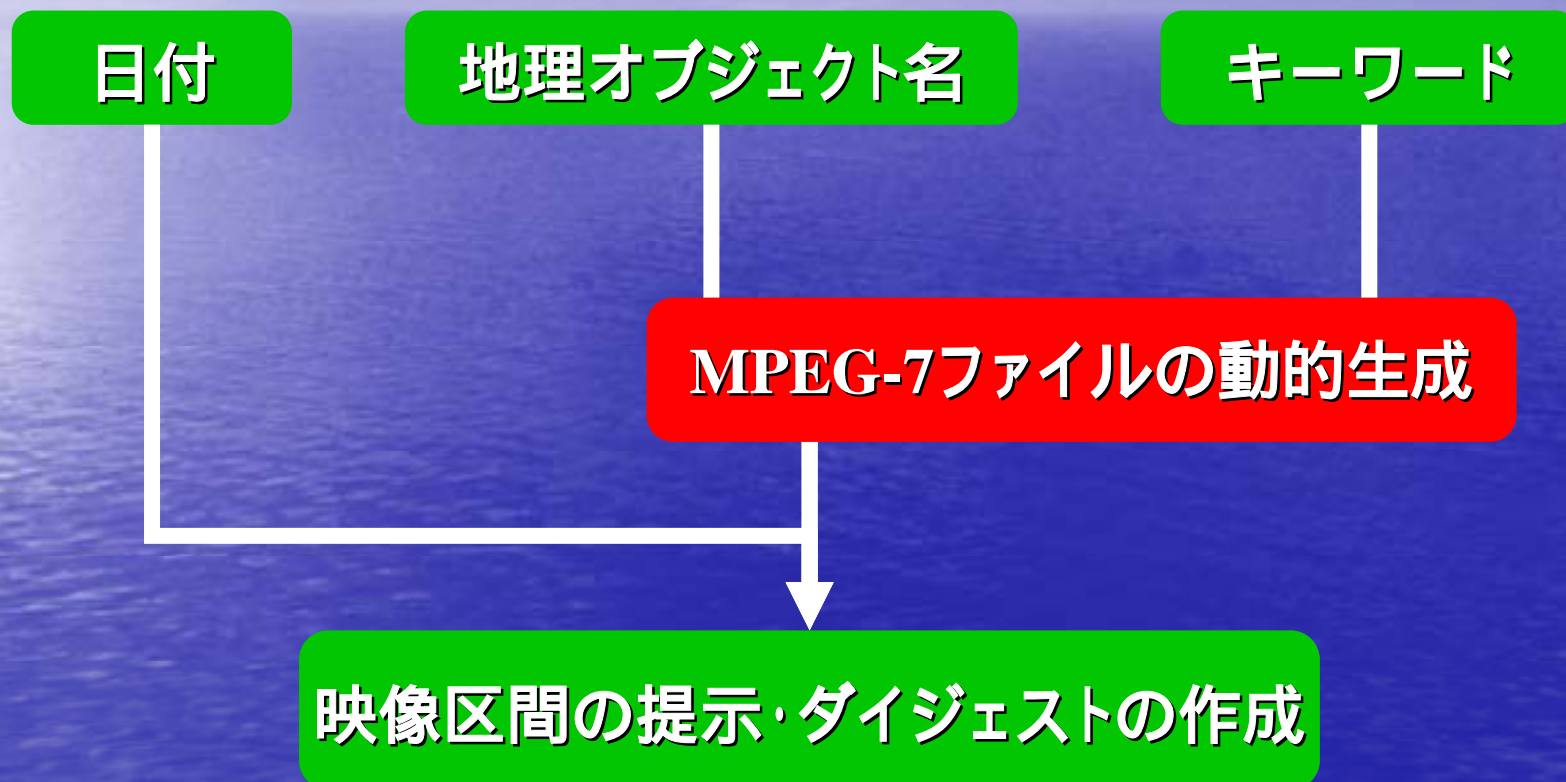


# アノテーションを基にした映像分割

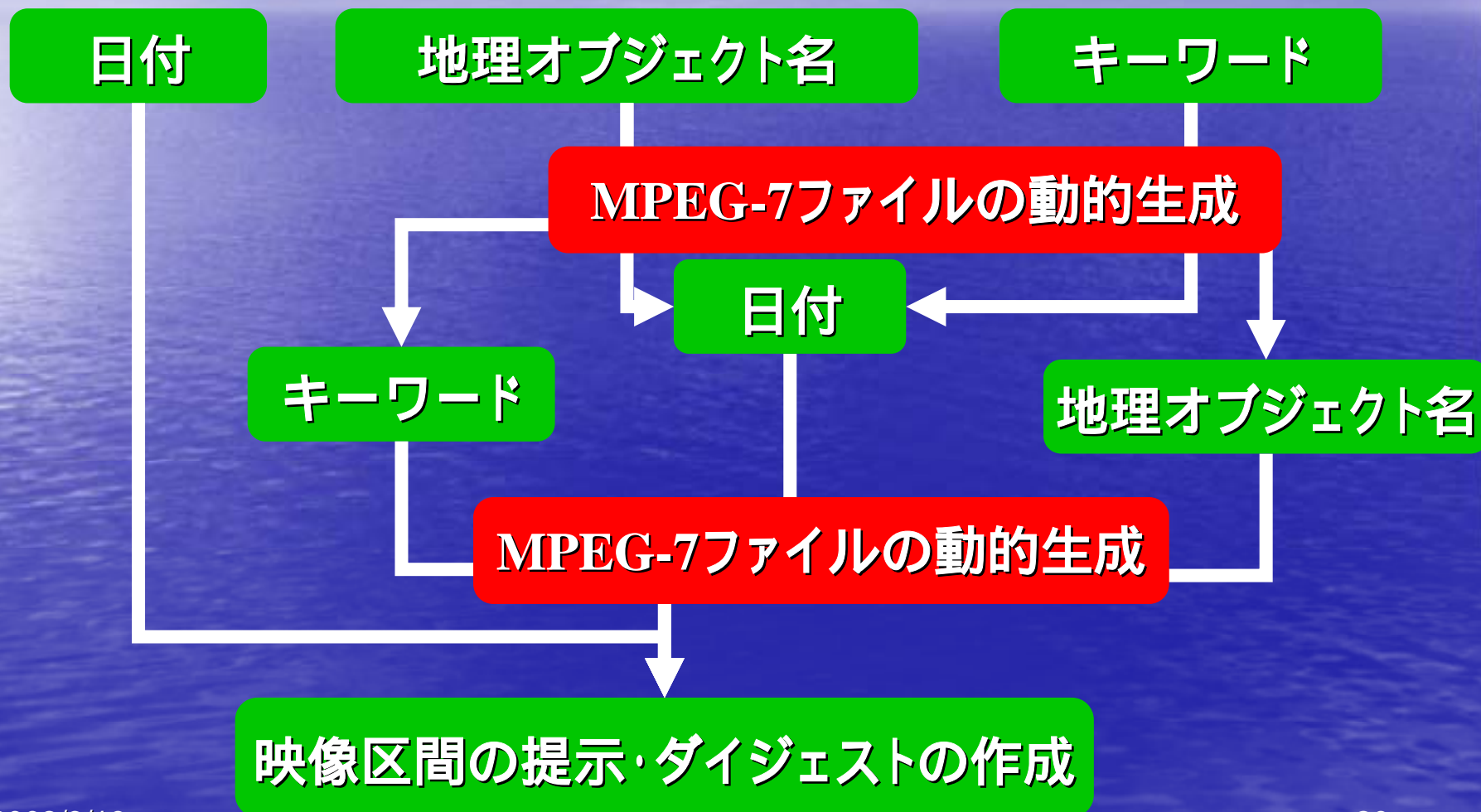
- 仮想の地理オブジェクト毎をまとめてMPEG-7ファイルへ記述

```
<AudioVisualSegment id="annotation1">  
  <GeographicPointType longitude="135.442099"  
latitude="34.438214" />  
  <PointOfView viewpoint="桜">  
    <Importance><Value>重要度</Value></Importance>  
  </PointOfView>  
  <MediaTime>  
    <MediaTimePoint>  
      2002-02-08T13:35:24:0F30  
    </MediaTimePoint>  
    <MediaIncrDuration timeUnit="PT1S">  
      374  
    </MediaIncrDuration>  
  </MediaTime>  
</AudioVisualSegment>
```

# MPEG-7ファイルを用いた検索



# MPEG-7ファイルを用いた検索



# MPEG-7ファイルの動的生成

```
<AudioVisualSegment id='place1' >  
  <GeographicPointType longitude="135.442099" latitude="34.438214" />  
  <PointOfView viewpoint="奈良先端科学技術大学院大学">  
    <Importance><Value>重要度</Value></Importance>
```

```
<AudioVisualSegment id='place2' >  
  <GeographicPointType longitude="135.442125" latitude="34.438639" />  
  <PointOfView viewpoint="高山サイエンスプラザ">  
    <Importance><Value>重要度</Value></Importance>
```

```
<AudioVisualSegment id="annotation1">  
  <GeographicPointType longitude="135.442099" latitude="34.438000" />  
  <PointOfView viewpoint="桜">  
    <Importance><Value>重要度</Value></Importance>
```

```
<AudioVisualSegment id="annotation2">  
  <GeographicPointType longitude="135.442125" latitude="34.438639" />  
  <PointOfView viewpoint="工事中">  
    <Importance><Value>重要度</Value></Importance>
```

# MPEG-7ファイルの動的生成

```
<AudioVisualSegment id="place1">  
  <GeographicPointType longitude="135.442099" latitude="34.438214" />  
  <PointOfView viewpoint="奈良先端科学技術大学院大学">  
    <Importance><Value>重要度</Value></Importance>
```

```
  </PointOfView>  
</AudioVisualSegment>  
.  
<AudioVisualSegment id="place2">  
  <GeographicPointType longitude="135.442125" latitude="34.438639" />  
  <PointOfView viewpoint="高山サイエンスプラザ">  
    <Importance><Value>重要度</Value></Importance>
```

```
  </PointOfView>  
</AudioVisualSegment>  
.  
<AudioVisualSegment id="annotation1">  
  <GeographicPointType longitude="135.442099" latitude="34.438000" />  
  <PointOfView viewpoint="桜">  
    <Importance><Value>重要度</Value></Importance>
```

```
  </PointOfView>  
</AudioVisualSegment>  
.  
<AudioVisualSegment id="annotation2">  
  <GeographicPointType longitude="135.442125" latitude="34.438639" />  
  <PointOfView viewpoint="工事中">  
    <Importance><Value>重要度</Value></Importance>
```

# 映像区間の提示

14:35:24

14:41:38

place1: 奈良先端科学技術大学院大学

14:38:00

14:45:00

place2: 高山サイエンスプラザ

14:34:00

14:37:30

14:40:00

14:43:00

annotation1: 桜

annotation1: 工事中

# ダイジェストの作成

- 地理オブジェクトの選択
- 再生時間の指定
- 地理オブジェクト毎の重要度に基づき再生時間を振り分け
- 地理オブジェクト毎の再生区間を特定



# ダイジェストの作成

「桜」の映像区間に割り当てられた時間を4秒と仮定

14:35:24	桜	重要度: 7.5	35.0
14:35:25	桜	重要度: 8.0	37.5
14:35:26	桜	重要度: 9.5	38.5
14:35:27	桜	重要度: 10.0	37.5
14:35:28	桜	重要度: 10.0	
14:35:29	桜	重要度: 9.5	
14:35:30	桜	重要度: 8.0	

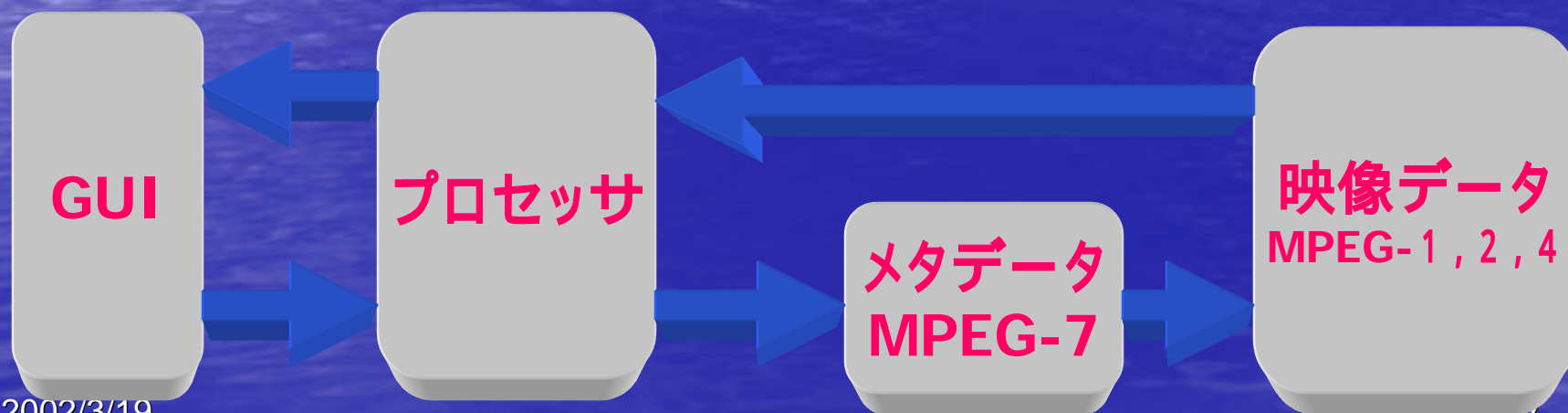
基礎データ内の情報

ダイジェストの再生時間



# MPEG-7とは？

- Multimedia Content Description Interface
  - マルチメディアコンテンツに対し内容を記述
  - MPEG-1,2,4とは異なる(付帯情報として併用)
- XML Schemaをベース言語



# MPEG-7記述例

```
<Mpeg7 type="complete" ... >  
  <ContentDescription xsi:type="ContentEntityType">  
    <MultimediaContent xsi:type="AudioVisualType">  
      <AudioVisual id="news1">  
        <MediaLocator> ... </MediaLocator>  
        <MediaTime>  
          <MediaRealTimePoint>PT0S </MediaRealTimePoint>  
          <MediaDuration>PT20M </MediaDuration>  
        </MediaTime>  
        <TemporalDecomposition>  
          <AudioVisualSegment id="introduction">  
            <MediaTime>  
              <MediaRealTimePoint>PT0S </MediaRealTimePoint>  
              <MediaDuration>PT1M </MediaDuration>  
            </MediaTime>  
          </AudioVisualSegment>  
        </TemporalDecomposition>  
      </AudioVisual>  
    </MultimediaContent>  
  </ContentDescription>  
</Mpeg7>
```

映像全体の情報

1番目の映像区間

# MPEG-7応用例

- MPEG-7を使用したダイジェスト映像生成
  - NTTドコモ、日本IBM (2001.9.18)
  - 移動体通信に向けての動画配信技術を開発

