情報学科CSコース情報システム（3年後期）

講義ノート
ー第10回ー

マルチメディア情報検索

（画像検索，ビデオ動画像検索，
Geminiアルゴリズム）

田中克己

マルチメディア・インデキシング

内容（コンテンツ）によるマルチメディアデータの探索

- 予測・診断・教育・仮説検証・データ発掘などに有用

GEMINI（Generic Multimedia Object Indexing)

時系列データとイメージデータを対象

基本的アイデア(1)

マルチメディアオブジェクト間の距離

- マルチメディアオブジェクト　O_A, O_B
- 距離（非類似度）関数　$D(O_A, O_B)$

例：等長の時系列データ
ユークリッド距離

$$O_A = x_1 x_2 \ldots x_n$$

$$O_B = y_1 y_2 \ldots y_n$$

$$D(O_A, O_B) =$$

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

基本的アイデア(2)

問合せ(Queries)

- 問合せもマルチメディアオブジェクト　Q
- 全体一致 (Whole Match)
  Qからある距離ε内にあるオブジェクトを探索
- 部分（パターン）一致 (Subpattern Match)
  Qにマッチする部分をもつオブジェクトを探索

検索に関する要件

高速性
逐次スキャンに基づく距離計算は遅すぎる

正当性

- 正しい答えはすべて返す。
  （false dismissals（誤った棄却）がない！）
- 正しくない答 (false alarms（誤報）) は含んでも可。

小さなスペース。

動的：オブジェクトの追加・削除・更新が容易なこと。

全体一致質問のためのGEMINI

quick-and-dirty test
答えでないオブジェクトを排除

空間アクセス法の利用

例：株価の時系列データS, 質問Q
距離関数 D( )
オブジェクトの特徴 (Feature)：k個の数字
F( ) ：オブジェクトをk次元空間の点に写像
F(Q)から距離ε内にある点を検索
実際のオブジェクトとQの距離を計算し
false alarmsを排除

$$S = S[1]S[2]\ldots S[l], Q = Q[1]Q[2]\ldots Q[l]$$

$$D(S,Q) = \sqrt{\sum_{i=1}^{l}(S[i]-Q[i])^2}$$

全体一致質問のためのGEMINI

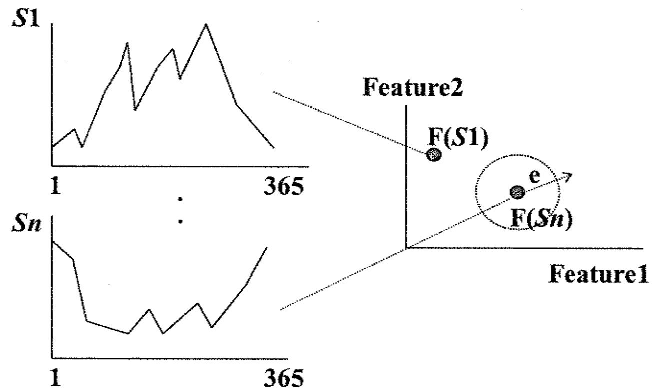
F-index (Feature Index)

- k次元空間の点集合の空間索引（R木やR*木を利用）

F-indexの探索

- (1) 質問オブジェクトQを特徴空間の点F(Q)に変換
- (2) 空間アクセス法をもちいて，F(Q)から距離ε内にある点を検索
- (3) 本当の距離を計算しfalse alarmsを排除

全体一致質問のためのGEMINI



データベース：時系列データ *S1, ...Sn*;
各時系列データは特徴空間上の点
許容度εの質問は，半径εの球

特徴空間上の距離 Dfeature

GEMINIが，全体一致質問に対してfalse dismissalsを起こさないためには，特徴関数 F( )は次式を満たす：

$$D_{feature}(F(O_1), F(O_2)) \leq D(O_1, O_2)$$

十分条件：

$$D_{feature}(F(O_1), F(O_2)) \leq D(O_1, O_2)$$

- 質問オブジェクト：Q，答えとなるオブジェクト：O
- 近傍質問の許容距離：ε
- 「OがQの答えならば，特徴空間でOはQの答えとなる」ことを証明
- すなわち，「D (Q,O) ≦ ε ならばDfeature(F(Q),F(O))≦ ε」を証明すればよい
- D (Q,O) ≦ ε と 条件 Dfeature(F(Q),F(O))≦D(Q,O)から，明らかにDfeature(F(Q),F(O))≦ ε が成立

GEMINI

- オブジェクト間の距離関数D( )を決定
- 1つ（又はそれ以上の）特徴抽出関数 F( )を見つける
- 特徴空間での距離が以下のようになることを証明
  $$D_{feature}(F(O_1), F(O_2)) \leq D(O_1, O_2)$$
- 空間アクセス法によってk次元特徴ベクトルを扱う

1次元時系列データのためのGEMINI

- データベース：等長の時系列データの集合
- 質問：1つの時系列データ
- 答え：質問に類似する時系列データの集合
- 時系列データS, Q
  - Len(S): Sの長さ
  - S[I:J]: Sの I番目からJ番目までの部分系列
  - S[I] : Sの I 番目の要素
  - D(S,Q) : SとQの距離

1次元時系列データのための GEMINI距離関数

ユークリッド距離

- 多くの場合に有用
- 多くの類似度は，適当な変換の後，特徴ベクトルのユークリッド距離に帰着可能

他の距離関数（Time-Warpingなど）

ユークリッド距離をlower-boundする特徴ベクトル

条件

- 元の距離関数をlower-boundする
- false alarmsができるだけ少ない

特徴抽出関数

- 平均
  全体平均，前半平均，後半平均，最初の1/4の平均，．．．（アダマール変換の係数に類似）
- 離散フーリエ変換 (Discrete Fourier Transform)

離散フーリエ変換(DFT)

- 信号 ***x***
- ***x***のn点離散フーリエ変換（X_fは複素数）
- 逆変換

$$\vec{x} = [x_0, x_1, \ldots, x_{n-1}]$$

$$\vec{X} = [X_0, X_1, \ldots, X_{n-1}]$$

$$X_f = 1/\sqrt{n}\sum_{i=0}^{n-1} x_i \exp(-j2\pi fi/n)$$

$$f = 0,1,.., n-1$$

$$j = \sqrt{-1}$$

$$x_i = 1/\sqrt{n}\sum_{f=0}^{n-1} X_f \exp(j2\pi fi/n)$$

$$i = 0,1,\ldots, n-1$$

Parsevalの定理

$|x_0|^2 + |x_1|^2 + \ldots\ldots + |x_{n-1}|^2 =$
$|X_0|^2 + |X_1|^2 + \ldots\ldots + |X_{n-1}|^2$

DFTは線形変換なので，D(**x**, **y**) = D(**X**, **Y**)．
すなわち，DFTはユークリッド距離を保存する．

DFTの最初のk個の係数を特徴とすると
Dfeature(F(**x**),F(**y**)) ＝
$|X_0|^2 + |X_1|^2 + \ldots\ldots + |X_{k-1}|^2 \leqq$
$|X_0|^2 + |X_1|^2 + \ldots\ldots + |X_{n-1}|^2 =$
$|x_0|^2 + |x_1|^2 + \ldots\ldots + |x_{n-1}|^2 =$ D(**x**, **y**)

ユークリッド距離を保存する変換

任意の正規直交変換

- 離散コサイン変換(DCT)
- ウエーブレット変換など

DFT

DFTパッケージ
（Mathematicaパッケージ，Cパッケージなどあり）

DFTの性能の良さ

- 最悪のケース：白色雑音の場合．つまり各xiがxi-1, xi+1に対して完全に独立な場合．
- 実際の信号：random walks (brown noise)など前の値に依存．
- 実験結果：k=2または3で十分な結果を示す．

サブパターン・マッチング（1D時系列の場合）

N個の時系列データ集合{S1, S2, ..., SN}
長さは可変

質問：部分系列Q（長さLen(Q)≧w≧1），許容度ε

答え：(Si, k)
D(Q, Si[k: k+Len(Q)-1])≦ε
D：ユークリッド距離

逐次スキャン法よりもスペースの少ない方法

ST Index (Subtrail Index)

I-naive法（Index naive法）

時系列Siを長さwの部分系列に（重複を許して）分割

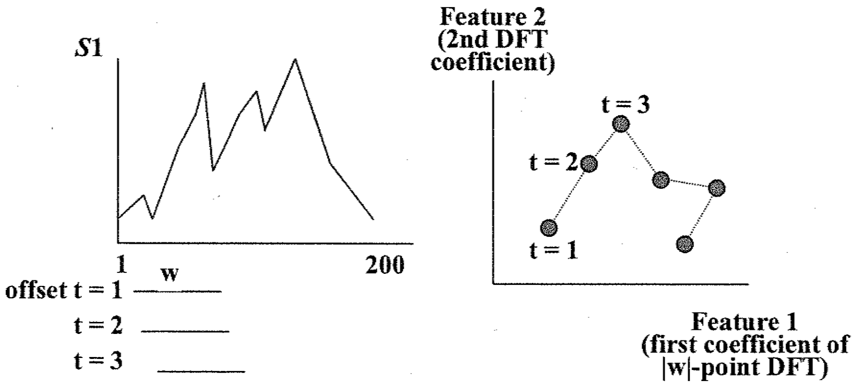
長さwの各部分系列を，w点DFT変換．

W点DFTの係数k個を特徴関数値として利用

時系列Siは，特徴空間上で，trail（小径）として表現される．

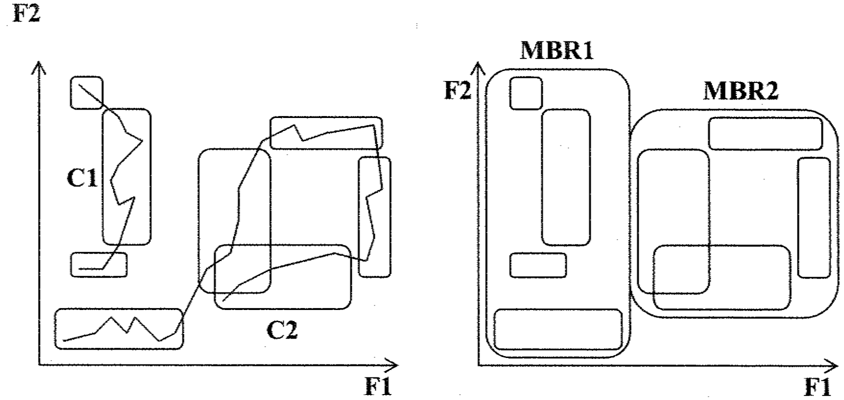
trailの各点をすべて記憶（R*木などを利用）

I-naive法



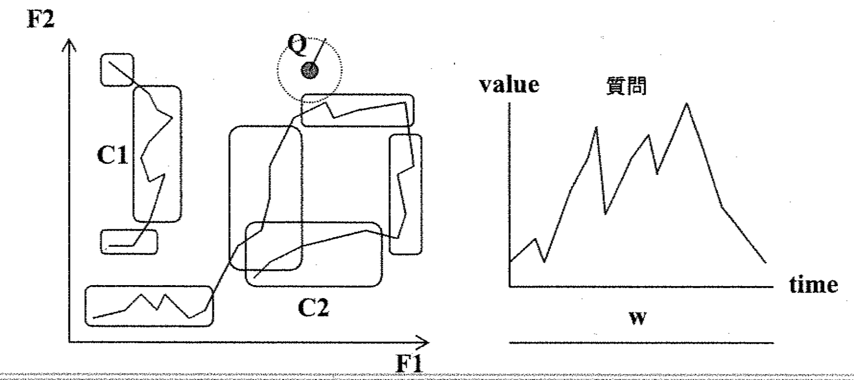
ウインドウ幅 w
時系列データ S1を，最初から，k個ずつに区切る.
長さkの部分系列を特徴空間に写像(DFTの利用)
特徴空間で，S1は，点列(小径)であらわされる.

ST Index



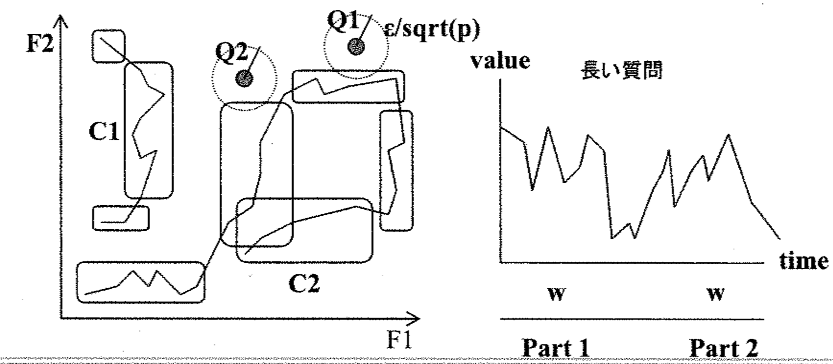
Trailをsubtrailに分割．各subtrailからMBRを形成．
MBRを階層的にグループ化しR*木に格納．

ST Index:short queryの場合



質問：長さwの時系列データ．特徴空間では点Q．
Qから半径εの円と重なるMBRを検索．
False alarmsを取り除く．

ST Index: long queryの場合



長い質問Q：長さwのp個の質問Q1,Q2 (p=2)に分割．
Q1, Q2から半径ε/sqrt(p) の円と重なるMBRを検索．
検索されたMBRの和集合をとる．False alarmsを除く．