DYNAMIC-SEGMENTATION-BASED FEATURE DIMENSION REDUCTION FOR QUICK AUDIO/VIDEO SEARCHING

Akisato Kimura, Kunio Kashino, Takayuki Kurozumi and *Hiroshi Murase
NTT Communication Science Laboratories, NTT Corporation
* Now, he is with Nagoya University.

SUMMARY #1

SUBJECT

- A quick audio/video signal search method based on **feature dimension reduction**

APPROACH (ðà #3)

Piecewise linear representation of feature trajectories

- **Dynamic segmentation**
  (our previous work [ICASSP02] : equi-partition)
- segment-based PCA

PROBLEM

- a huge amount of calculation for dynamic segmentation

SOLUTION (ðà #5)

- suboptimal partitioning of feature trajectories

RESULTS (ðà #6)

- **Compression performance :**
  almost the same as the optimal one
- **Calculation for compression : 1/5000** of the optimal one

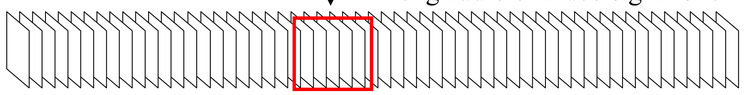
BACKGROUND #2

Objectives quick search for audio/video signals

query :
a known
audio or video signal

Matching
(based on a certain distance measure)

database :
a long audio or video signal stream



Applications

- retrieving for broadcasting music and movies
- checking copyright materials on the Internet

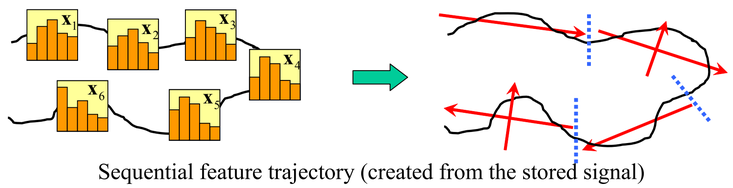
Problems high dimensionality of features

- Calculation for matching : large
- Indexing algorithm : not suitable

Feature dimension reduction is necessary

APPROACH #3

Piecewise linear representation of feature trajectories



Sequential feature trajectory (created from the stored signal)

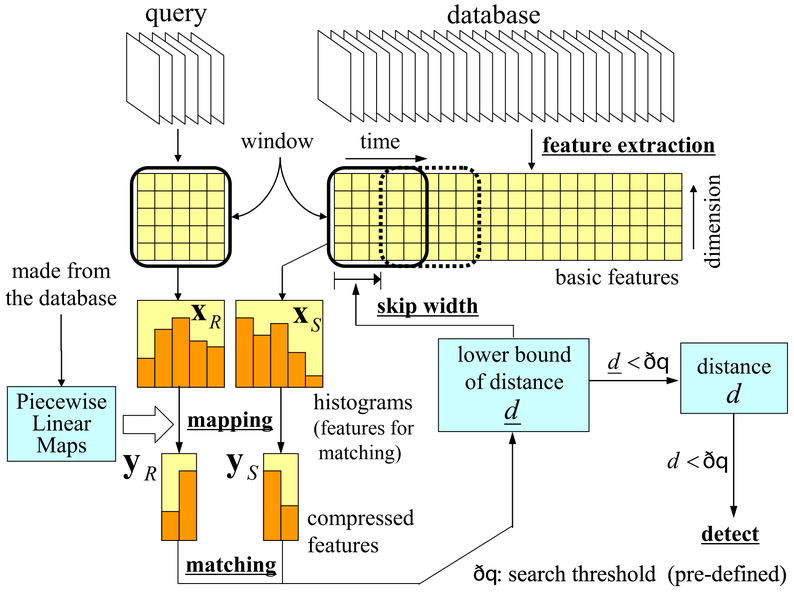
How to get the representation
(our previous work [ICASSP02])

- Equi-partitioning :
  partitioning feature trajectories into segments of equal length
- Segment-based PCA :
  perform PCA (Principal Component Analysis) in every segment

(this work)

- **Dynamic segmentation :**
  partitioning feature trajectories dynamically
  so as to minimize the average dimensionality
- Segment-based PCA

OUTLINE #4



DETAILS #5

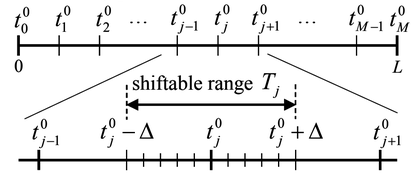
Dynamic segmentation

$$T^* = \{t_j\}_{j=0}^{M} = \arg\min_{t_j \in T_j \forall j} \frac{1}{L}\sum_{i=1}^{M}(t_j - t_{j-1})c(t_{j-1}, t_j, \delta s)$$

segment boundaries

shifable range of a boundary
$T_j = \{t \mid t_j^0 - \Delta \le t_j \le t_j^0 + \Delta\}$

average dimensionality per frame



→ needs a huge amount of calculation (cost = $(M\Delta^2)$)

To reduce calculation costs…

Suboptimal approach (cf. Proc.)

- Local search : (calculation cost → $(M\Delta)$)
  determine each boundary independently and in order of time
- Efficient pruning : (calculation cost → $\left(M\sqrt{K\Delta}\right)$)
  based on coarse-to-fine approach

Note : M =number of segments, Δ =width of shiftable range, K=constant (usually $K << \Delta$)

EXPERIMENTS #6

Conditions

- Database :
  a video recording of a 24-hour TV broadcasting
- Dimensions of histograms : 256
- Number of segments : 1000
- Required contribution ratio : 0.95
- The initial positions of boundaries : obtained by equi-partitioning

Results (width of shiftable range = 500)

Dimension reduction performance → almost the same as the others

	Equi-partitioning	DP	Local search	**proposed**
Average dimension	6.85034	unable	6.37024	**6.36971**

Calculation comparison →much less than the others
(1/6000 of DP, 1/15 of the local search)

	Equi-partitioning	DP	Local search	**proposed**
Number of operations	1.0×10^3	1.0×10^9	2.0×10^6	**2.2×10^5**