Scaling and Time Warping in Time Series Querying

Ada Wai-chee Fu[1] Eamonn Keogh[2]
Leo Yung Hang Lau[1] Chotirat Ann Ratanamahatana[2]

[1]Department of Computer Science and Engineering
The Chinese University of Hong Kong

[2]Department of Computer Science and Engineering
University of California, Riverside

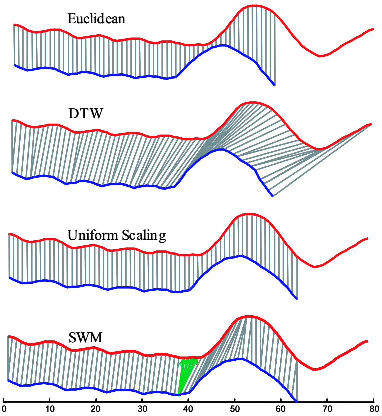
VLDB 2005

Outline

Introduction

- Euclidean Distance
  - No alignment
- Dynamic Time Warping (DTW)
  - Local alignment
- Uniform Scaling (US)
  - Global scaling
- Scaled and Warped Matching (SWM)
  - Both global scaling and local alignment are important!

Indexing Video (Sports Data)



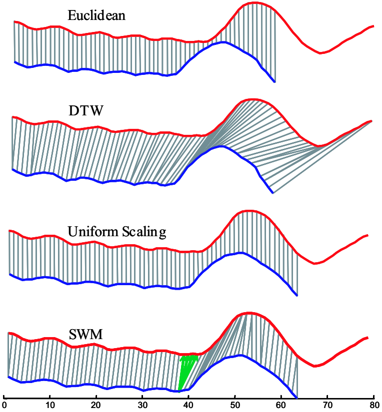
- Indexing sports data

  Sports fans Find particular types of shots or moves

  Coaches Analyze athletes' performance over time

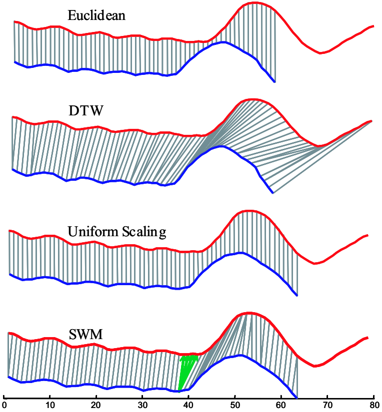
  - Video clips recording an athlete performing high jump
  - Collect the athlete's center of mass data from video (automatically)
  - Convert the data into a time series
  - Two examples of an athlete's trajectories aligned with various measures

Indexing Video (Sports Data)



- ✗ Euclidean Distance
  - Mapping part of the flight of one sequence to the takeoff drive in the other
- ✗ Dynamic Time Warping (DTW)
  - Trying to explain part of the sequence in one attempt (the bounce from the mat) that simply does not exist in the other sequence

Indexing Video (Sports Data)

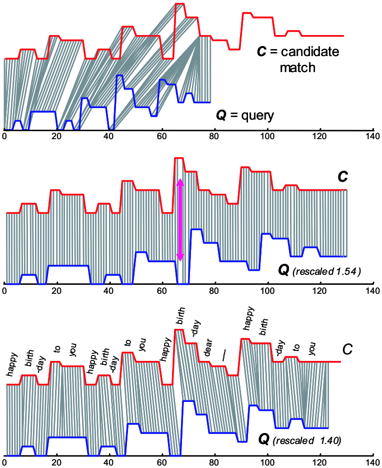


- ✗ Uniform Scaling (US)
  - Best match when we stretch the shorter sequence by 112%
  - Poor local alignment at takeoff drive ad up-flight
- ✓ Scaled and Warped Matching (SWM)
  - Global stretching at 112% allows DTW to align the small local differences

Query by Humming

- Search large music collections by providing an example of the desired content, by humming (or singing, or tapping) a snippet
- Humans cannot be expected to reproduce an exact fragment of a song
  - Query must be made invariant to key
  - Wrong tempo
  - Users may insert or delete notes
- Existing approaches
  - Do DTW multiple times, at different scalings
  - Do DTW with relatively short song snippets
    - Less sensitive to uniform scaling problem
    - Less discriminating power

Query by Humming



- *Happy birthday to you*
  - At very different tempos
- DTW doesn't produce the desired alignment
  - No global scaling
- US produces better global alignment, but serious local misalignments
  - No local alignment
- Only SWM produces the correct alignment
  - US aligns globally while DTW corrects the local misalignments

Problem Definition

- Given
  - A database D of M variable lengths data sequences
  - A query Q
  - A scaling factor l, $l \geq 1$
  - A time warping constraint r

Problem

Assume the data sequences can be longer than the query sequence Q. Find the best match to Q in database, for any rescaling in a given range, under the Dynamic Time Warping distance with a global constraint. By best match we mean the data sequence with the smallest distance from Q.

Time Warping Distance (DTW)

Definition (Time Warping Distance (DTW))

Given two sequences $C = C_1, C_2, \cdots, C_n$ and $Q = Q_1, Q_2, \cdots, Q_m$, the time warping distance DTW is defined recursively as follows:

$$\mathrm{DTW}(\phi, \phi) = 0$$

$$\mathrm{DTW}(C, \phi) = \mathrm{DTW}(\phi, Q) = \infty$$

$$\mathrm{DTW}(C, Q) = D_{base}(\mathrm{First}(C), \mathrm{First}(Q)) + \min \begin{cases} \mathrm{DTW}(C, \mathrm{Rest}(Q)) \\ \mathrm{DTW}(\mathrm{Rest}(C), Q) \\ \mathrm{DTW}(\mathrm{Rest}(C), \mathrm{Rest}(Q)) \end{cases}$$

where $\mathrm{First}(C) = C_1$, $\mathrm{Rest}(C) = C_2, C_3, \cdots, C_n$, ϕ is the empty sequence, and D_{base} denotes the distance between two entries.

Warping Matrix

- An example warping matrix aligning the time series
  - $\{1, 2, 2, 4, 5\}$ and
  - $\{1, 1, 2, 3, 5, 6\}$

5	27	27	13	5	**1**	**2**
4	11	11	4	**1**	2	6
2	2	2	**0**	1	10	26
2	1	1	**0**	1	10	26
1	**0**	**0**	1	5	21	46
	1	1	2	3	5	6

$\mathrm{DTW}(\mathrm{Rest}(C), Q)$	$\mathrm{DTW}(C, Q)$
$\mathrm{DTW}(\mathrm{Rest}(C), \mathrm{Rest}(Q))$	$\mathrm{DTW}(C, \mathrm{Rest}(Q))$

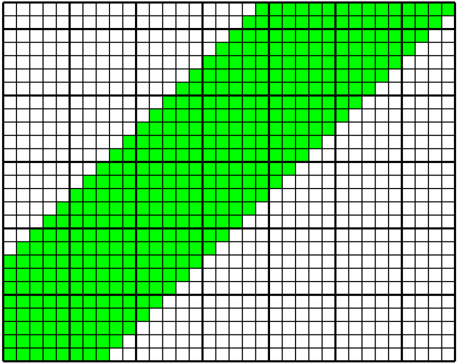
Warping Matrix

- An example warping matrix aligning the time series
  - $\{1, 2, 2, 4, 5\}$ and
  - $\{1, 1, 2, 3, 5, 6\}$

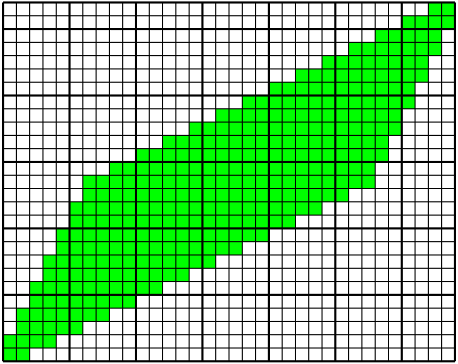
5	27	27	13	5	**1**	**2**
4	11	11	4	**1**	2	6
2	2	2	**0**	1	10	26
2	1	1	**0**	1	10	26
1	**0**	**0**	1	5	21	46
	1	1	2	3	5	6

- The highlighted entries denote the warping path.
- The DTW distance is 2. (the value at the top-right entry)

Constraints on the Warping Path



- Sakoe-Chiba Band



- Itakura Parallelogram

Constrained DTW (cDTW)

Definition (Constrained DTW (cDTW))

Given two sequences $C = C_1, C_2, \cdots, C_n$ and $Q = Q_1, Q_2, \cdots, Q_m$, and the time warping constraint r, the constrained time warping distance cDTW is defined recursively as follows:

$$\mathrm{Dist_r}(C_i, Q_j) = \begin{cases} D_{base}(C_i, Q_j) & \text{if } |i - j| \leq r \\ \infty & \text{otherwise} \end{cases}$$

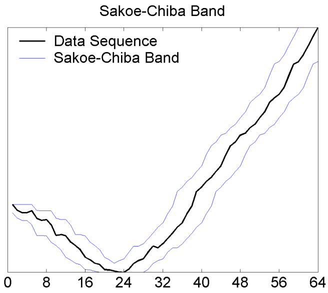
$$\mathrm{cDTW}(\phi, \phi, r) = 0$$

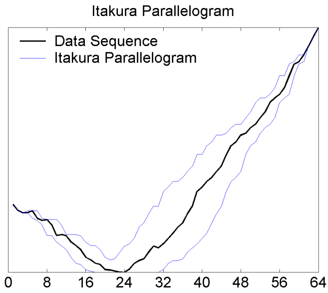
$$\mathrm{cDTW}(C, \phi, r) = \mathrm{cDTW}(\phi, Q, r) = \infty$$

$$\mathrm{cDTW}(C, Q, r) = \mathrm{Dist_r}(\mathrm{First}(C), \mathrm{First}(Q)) + \min \begin{cases} \mathrm{cDTW}(C, \mathrm{Rest}(Q), r) \\ \mathrm{cDTW}(\mathrm{Rest}(C), Q, r) \\ \mathrm{cDTW}(\mathrm{Rest}(C), \mathrm{Rest}(Q), r) \end{cases}$$

where ϕ is the empty sequence, $\mathrm{First}(C) = C_1$, $\mathrm{Rest}(C) = C_2, C_3, \cdots, C_n$, and D_{base} denotes the distance between two entries.

Constraints and Enveloping Sequences





Constraints and Enveloping Sequences

Definition (Enveloping Sequences for DTW)

Let $UW = UW_1, UW_2, \cdots, UW_m$ and
$LW = LW_1, LW_2, \cdots, LW_m$,

$$UW_i = \max(C_{i-r}, \cdots, C_{i+r}) \text{ and}$$
$$LW_i = \min(C_{i-r}, \cdots, C_{i+r})$$

Considering the boundary cases, the above can be rewritten as

$$UW_i = \max(C_{\max(1,i-r)}, \cdots, C_{\min(i+r,n)}) \text{ and}$$
$$LW_i = \min(C_{\max(1,i-r)}, \cdots, C_{\min(i+r,n)})$$

Lower Bounding DTW

Definition (Lower Bounding DTW)

$$LB_W(Q,C) = \sum_{i=1}^{m} \begin{cases} (Q_i - UW_i)^2 & \text{if } Q_i > UW_i \\ (Q_i - LW_i)^2 & \text{if } Q_i < LW_i \\ 0 & \text{otherwise} \end{cases}$$

Uniform Scaling (US)

Definition (Uniform Scaling (US))

Given two sequences $Q = Q_1, \cdots, Q_m$ and $C = C_1, \cdots, C_n$ and a scaling factor bound $l, l \geq 1$. Let $C(q)$ be the prefix of C of length q, where $\lceil m/l \rceil \leq q \leq lm$ and $C(m, q)$ be a rescaled version of $C(q)$ of length m,

$$C(m, q)_i = C(q)_{\lceil i \cdot q/m \rceil} \text{ where } 1 \leq i \leq m$$

$$\mathrm{US}(C, Q, l) = \min_{q=\lceil m/l \rceil}^{\min(lm, n)} \mathrm{D}(C(m, q), Q)$$

where $\mathrm{D}(X, Y)$ denotes the Euclidean distance between two sequences X and Y.

Lower Bounding US

Definition (Enveloping Sequences for US)

We create two sequences $UC = UC_1, \cdots, UC_m$ and $LC = LC_1, \cdots, LC_m$, such that

$$UC_i = \max(C_{\lceil i/l \rceil}, \cdots, C_{\lceil il \rceil})$$
$$LC_i = \min(C_{\lceil i/l \rceil}, \cdots, C_{\lceil il \rceil})$$

Definition (Lower Bounding US)

$$LB_S(Q, C) = \sum_{i=1}^{m} \begin{cases} (Q_i - UC_i)^2 & \text{if } Q_i > UC_i \\ (Q_i - LC_i)^2 & \text{if } Q_i < LC_i \\ 0 & \text{otherwise} \end{cases}$$

Scaling and Time Warping (SWM)

Definition (Scaling and Time Warping (SWM))

Given two sequences $Q = Q_1, \cdots, Q_m$ and $C = C_1, \cdots, C_n$, a bound on the scaling factor $l, l \geq 1$ and the Sakoe-Chiba Band time warping constraint r which applies to sequence length m. Let $C(q)$ be the prefix of C of length q, where $\lceil m/l \rceil \leq q \leq \min(lm, n)$ and $C(m, q)$ be a rescaled version of $C(q)$ of length m,

$$C(m,q)_i = C(q)_{\lceil i \cdot q/m \rceil} \text{ where } 1 \leq i \leq m$$

$$\mathrm{SWM}(C, Q, l, r) = \min_{q=\lceil m/l \rceil}^{\min(lm,n)} \mathrm{cDTW}(C(m,q), Q, r)$$

Enveloping Sequences for SWM

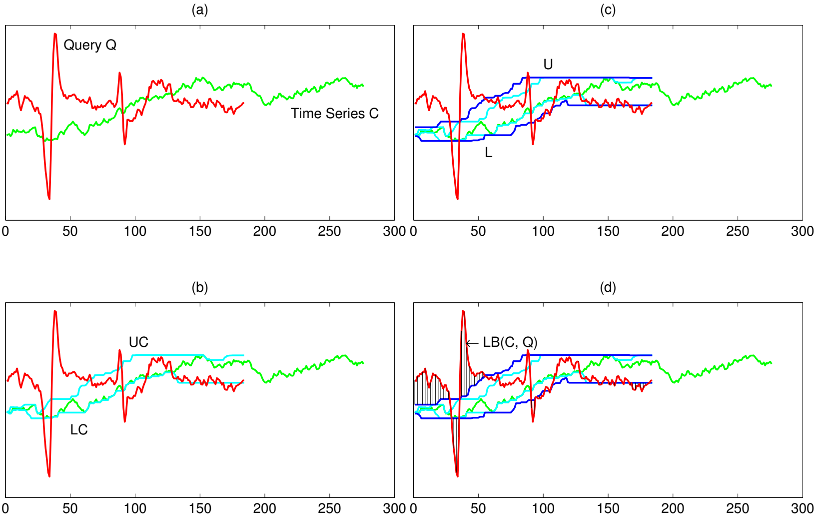
Definition (Enveloping Sequences for SWM)

$$U_i = \max(C_{\max(1,\lceil i/l \rceil - r')}, \cdots, C_{\min(\lceil il \rceil + r', n)})$$
$$L_i = \min(C_{\max(1,\lceil i/l \rceil - r')}, \cdots, C_{\min(\lceil il \rceil + r', n)})$$

Definition (Lower Bounding SWM)

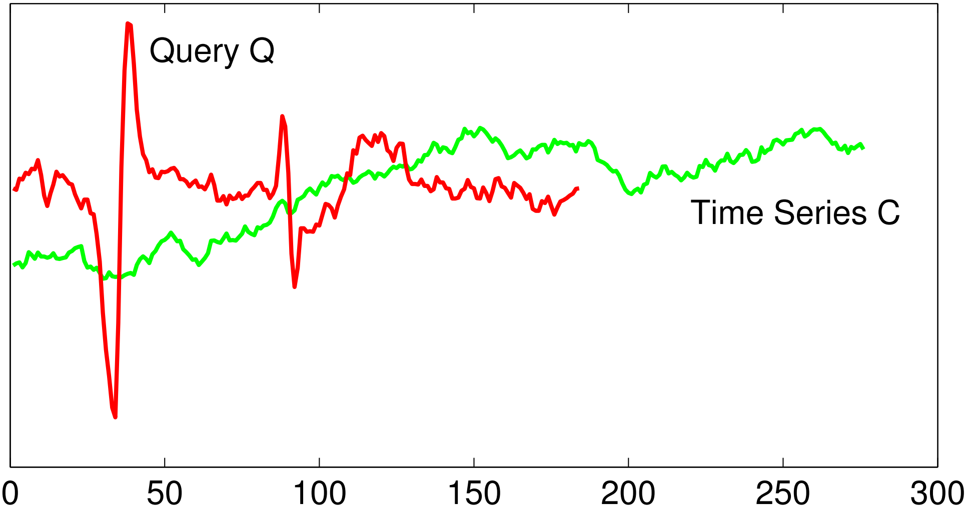
$$LB(Q, C) = \sum_{i=1}^{m} \begin{cases} (Q_i - U_i)^2 & \text{if } Q_i > U_i \\ (Q_i - L_i)^2 & \text{if } Q_i < L_i \\ 0 & \text{otherwise} \end{cases}$$

An illustration of the SWM envelopes

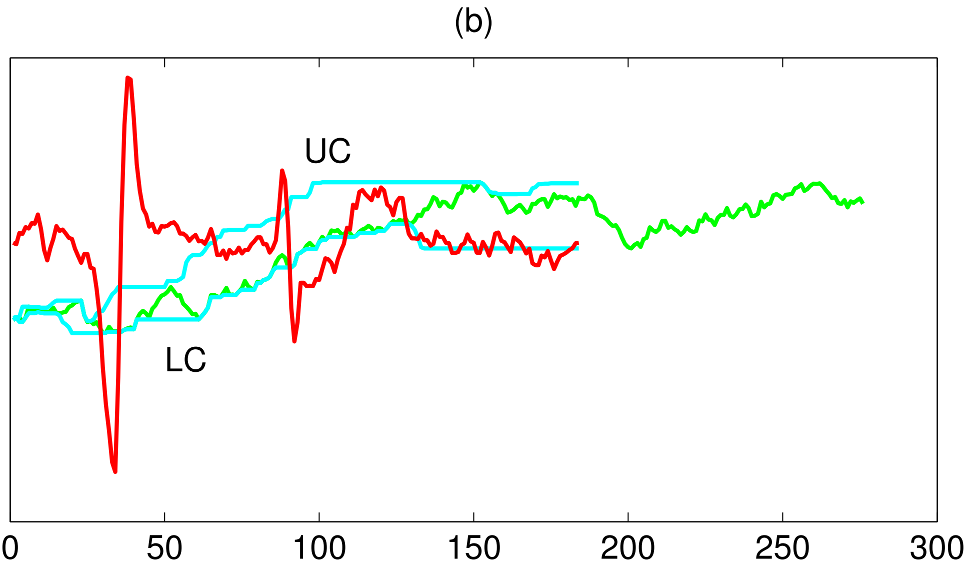


An illustration of the SWM envelopes

(a)

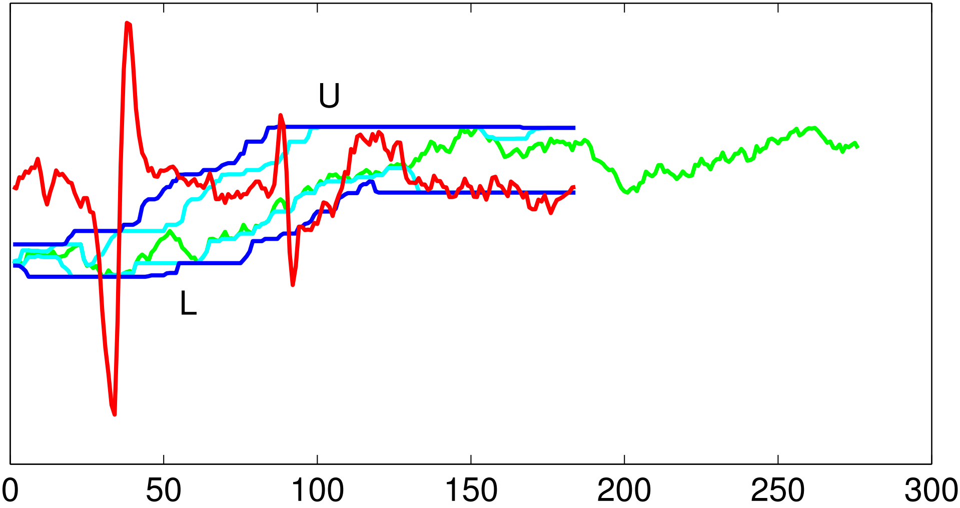


An illustration of the SWM envelopes

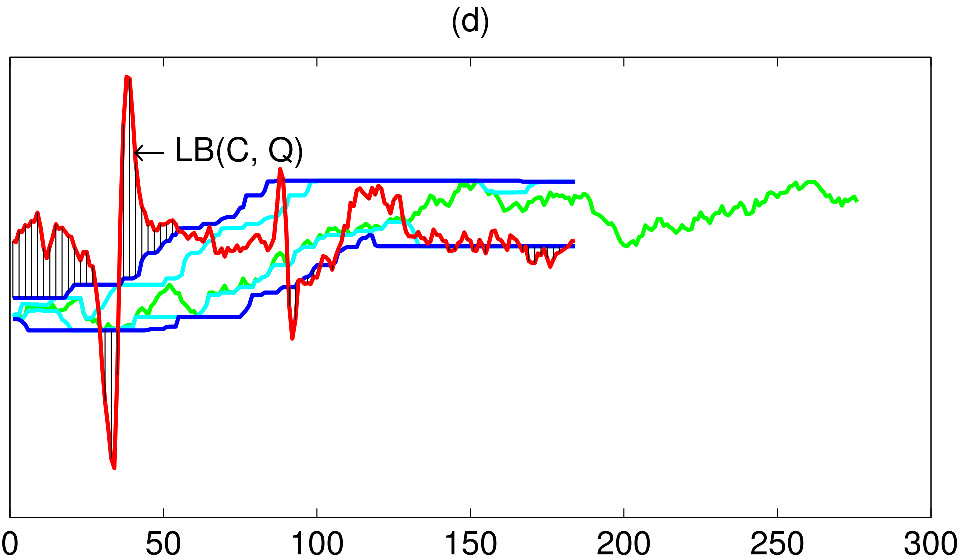


An illustration of the SWM envelopes

(c)



An illustration of the SWM envelopes



The Lower Bounding Lemma

Lemma (Lower Bounding Lemma)

For any two sequences Q and C of length m and n respectively, given a scaling constraint of $\{1/l, l\}$*, where* $l \geq 1$*, and a Sakoe-Chiba Band time warping constraint of* r' *on the original (unscaled) sequence C, the value of* $LB(Q, C)$ *lower bounds the distance of* $SWM(C, Q, l, r')$*.*

Proof Sketch.

1. The matching warping path $w_k = (i, j)_k$ defines a mapping between the indices i and j. Each such mapping constitutes term $t = (Q_i, C_j)^2$ to the required distance.
2. We can show that the i-th term t_{lb} in our lower bounding distance $LB(Q, C)$ can be matched with the term t resulting in a one-to-one mapping, with $t_{lb} \leq t$. □

Tightness of Lower Bounds

Definition

Consider a lower bound $LB(Q,C)$ for a distance $D(Q,C)$ of the form

$$LB(Q,C) = \sum_{i=1}^{m} \begin{cases} (Q_i - U_i)^2 & \text{if } Q_i > U_i \\ (Q_i - L_i)^2 & \text{if } Q_i < L_i \\ 0 & \text{otherwise} \end{cases}$$

We say that the lower bound is tight, if there exists a set of sequence pairs so that for each pair $\{Q,C\}$ in the set,

1. $D(Q,C) = LB(Q,C)$, and
2. The U_i and L_j values for some i, j are used (in the $(Q_i - U_i)^2$ or $(Q_j - L_j)^2$ term) at least once in computing the lower bounds in the set.

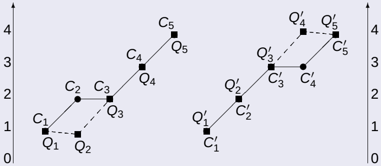
Tightness of LB_W

Lemma (Tightness of LB_W)

The lower bound $LB_W(Q, C)$ for the DTW distance with the Sakoe-Chiba Band constraint is tight.

Proof.

Consider DTW with a Sakoe-Chiba Band constraint of $r = 1$.
Hence in the warping path entry (i, j), $j - 1 \leq i \leq j + 1$.



It is easy to see that $D(Q, C) = LB_W(Q, C)$, and $D(Q', C') = LB_W(Q', C')$.
For Q, C, $Q_2 < LW_2$ and hence LW_2 is used in the computation of $LB_W(Q, C)$.
For Q', C', $Q'_4 > UW'_4$, hence UW'_4 is used in the computation of $LB_W(Q', C')$. □

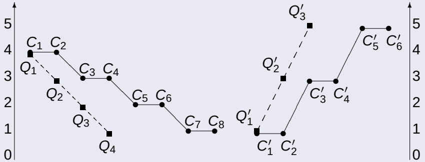
Tightness of LB_S

Lemma (Tightness of LB_S)

The lower bound $LB_S(Q, C)$ for the distance between Q, C with a scaling factor between $1/l$ and l is tight.

Proof.

Consider scaling between 0.5 and 2.
Hence $l = 2$.



It is easy to see that $D(Q, C) = LB_S(Q, C)$, and $D(Q', C') = LB_S(Q', C')$.
For Q, C, $LC_i > Q_i$ and all LC_i are used in the computation of $LB_S(Q, C)$.
For Q', C', $UC'_i < Q'_i$ and all UC'_i are used in the computation of $LB_S(Q', C')$. □

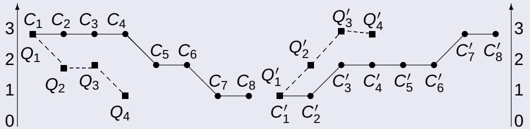
Tightness of LB

Lemma (Tightness of LB)

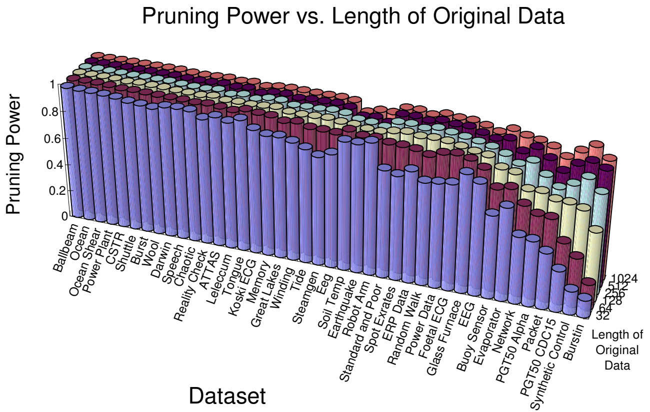
The lower bound $LB(Q,C)$ for the distance between Q,C with a scaling factor bound l and time warping with the Sakoe-Chiba Band constraint r' is tight.

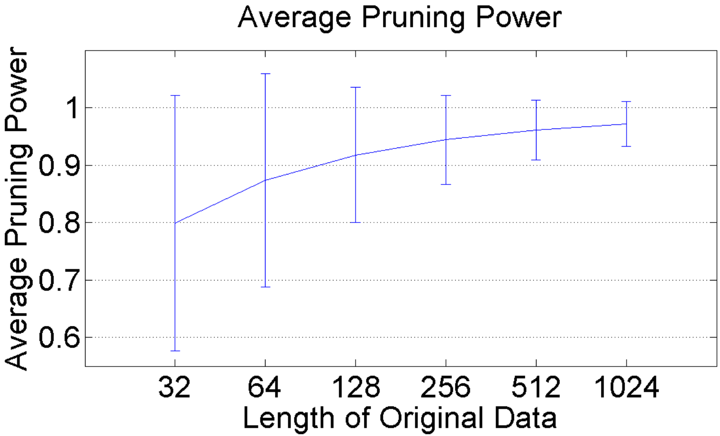
Proof.

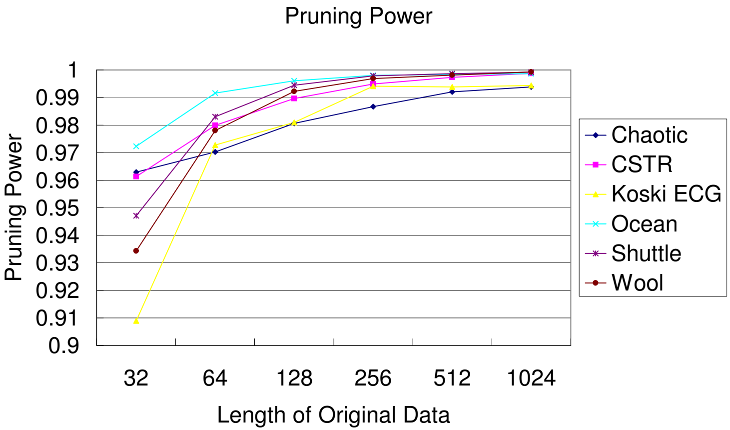
Consider a Sakoe-Chiba Band constraint of $r' = 1$ and a scaling factor between 0.5 and 2. Hence $l = 2$.

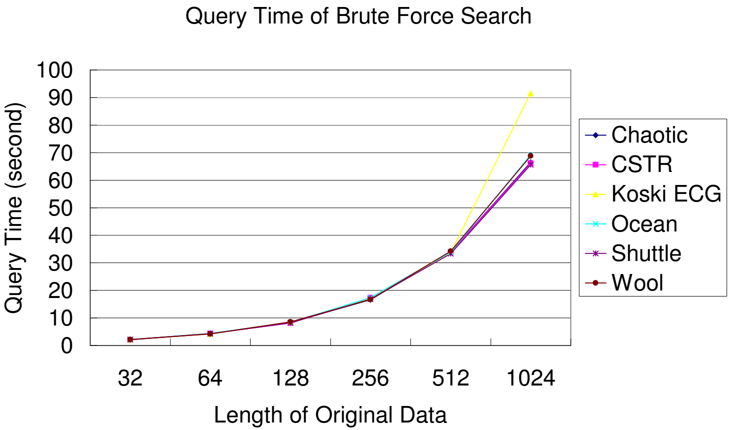


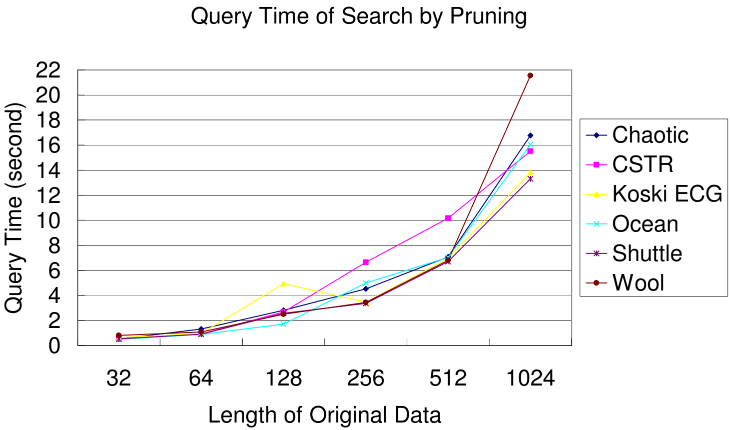
It is easy to see that $SWM(Q,C,l,r') = LB(Q,C)$, and $SWM(Q',C',l,r') = LB(Q',C')$.
For Q,C, $Q_2 < L_2$ and L_2 is used in the computation of $LB(Q,C)$.
For Q',C', $Q'_3 > U'_3$ and U'_3 is used in the computation of $LB(Q',C')$. □

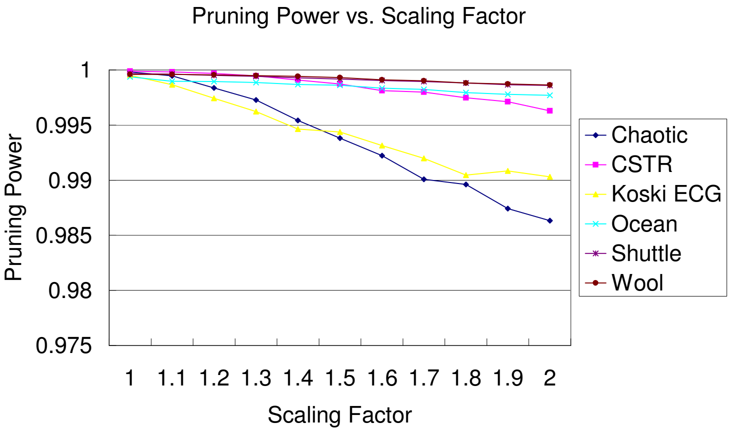

Pruning Power vs. Length of Original Data
Pruning Power
0
0.2
0.4
0.6
0.8
1
Dataset
Ballbeam
Ocean
Ocean Shear
Power Plant
CSTR
Shuttle
Burst
Wool
Darwin
Speech
Chaotic
Reality Check
ATTAS
Leleccum
Tongue
Koski ECG
Memory
Great Lakes
Winding
Tide
Steamgen
Eeg
Soil Temp
Earthquake
Robot Arm
Standard and Poor
Spot Exrates
ERP Data
Random Walk
Power Data
Foetal ECG
Glass Furnace
EEG
Buoy Sensor
Evaporator
Network
PGT50 Alpha
Packet
PGT50 CDC15
Synthetic Control
Burstin
Length of Original Data
32
64
128
256
512
1024


Average Pruning Power
Average Pruning Power
1
0.9
0.8
0.7
0.6
32
64
128
256
512
1024
Length of Original Data

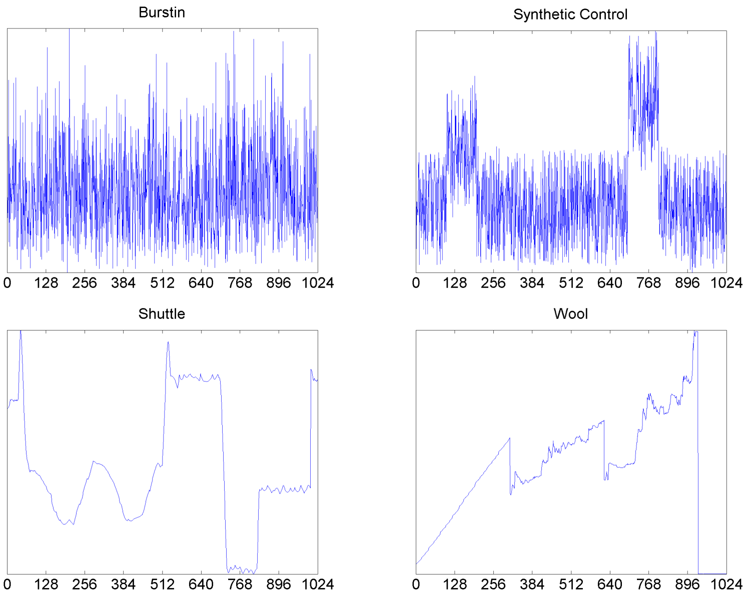

Pruning Power
Pruning Power
1
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0.92
0.91
0.9
32
64
128
256
512
1024
Length of Original Data
Chaotic
CSTR
Koski ECG
Ocean
Shuttle
Wool

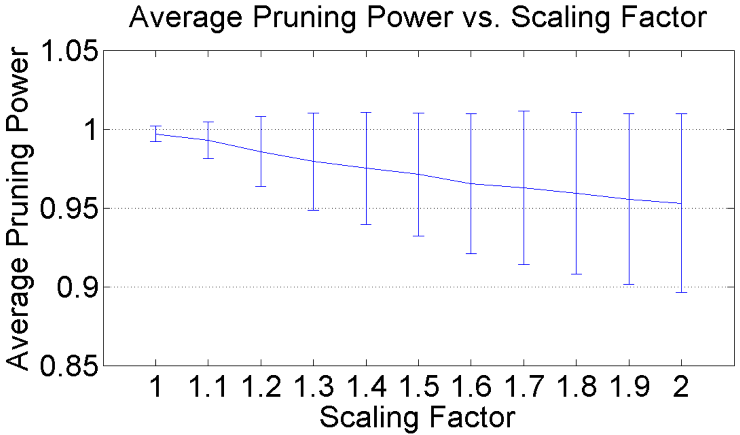

Query Time of Brute Force Search
Query Time (second)
0
10
20
30
40
50
60
70
80
90
100
32
64
128
256
512
1024
Length of Original Data
Chaotic
CSTR
Koski ECG
Ocean
Shuttle
Wool


Query Time of Search by Pruning
Query Time (second)
0
2
4
6
8
10
12
14
16
18
20
22
32
64
128
256
512
1024
Length of Original Data
Chaotic
CSTR
Koski ECG
Ocean
Shuttle
Wool



Pruning Power vs. Scaling Factor
Pruning Power
1
0.995
0.99
0.985
0.98
0.975
1
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2
Scaling Factor
Chaotic
CSTR
Koski ECG
Ocean
Shuttle
Wool


Burstin
0
128
256
384
512
640
768
896
1024
Synthetic Control
0
128
256
384
512
640
768
896
1024
Shuttle
0
128
256
384
512
640
768
896
1024
Wool
0
128
256
384
512
640
768
896
1024


Average Pruning Power vs. Scaling Factor
Average Pruning Power
1.05
1
0.95
0.9
0.85
1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2
Scaling Factor

Conclusion

1. Reviewed existing time series similarity measures
2. Showed that these measures are inappropriate or insufficient for many applications.
3. Proposed Scaled and Warped Matching (SWM)
4. Derived a lower bounding function for SWM
5. Experimentally showed the effectiveness of the lower bounding function