

Real Time Tracking of Musical Performances

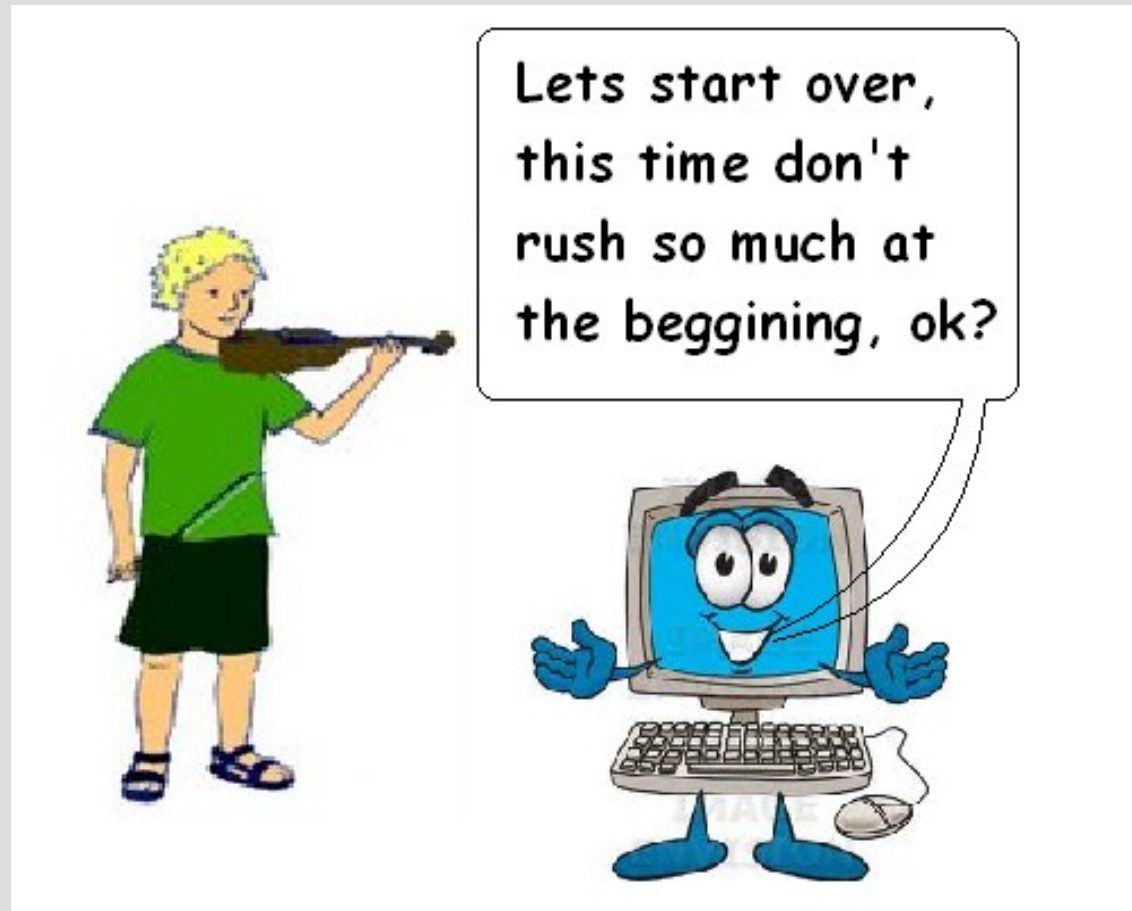
Antonio Camarena-Ibarrola & Edgar Chávez
{camarena,elchavez}@umich.mx

Universidad Michoacana de San Nicolás de
Hidalgo

Motivation

- Virtual Music Teacher
- Automatic accompaniment of single musicians or singers as well
- Automatic adding of special effects on live events

Virtual Music Teacher



Automatic accompaniment of single musicians or singers



Automatic adding of special effects on live events

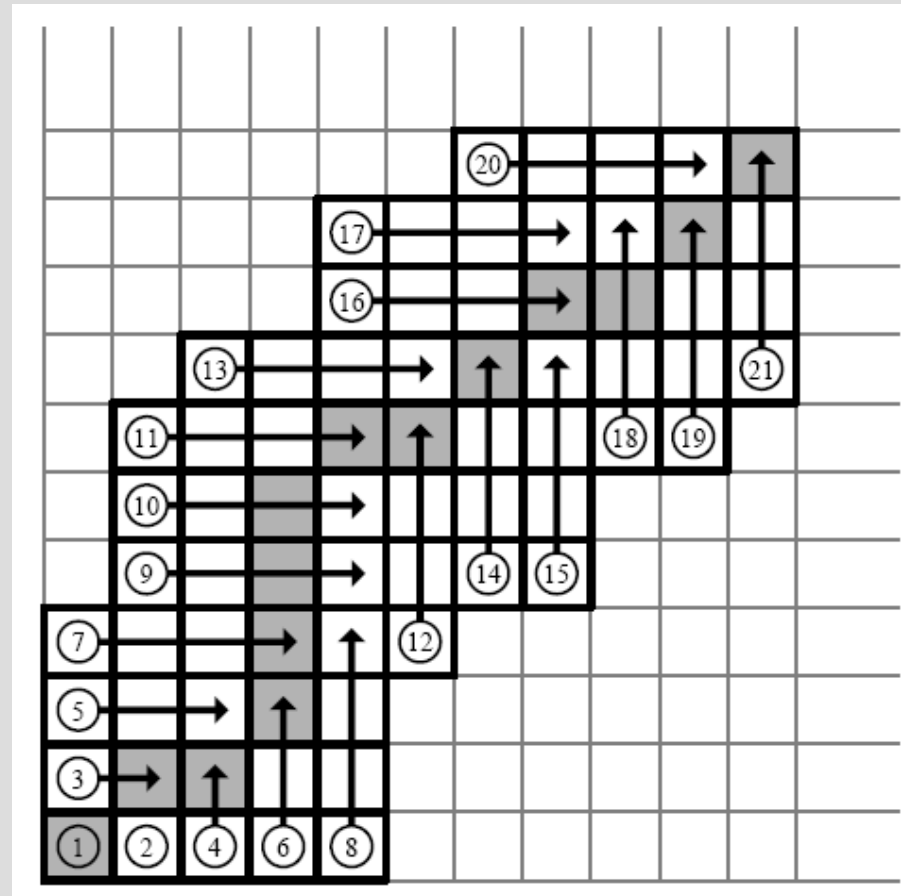


It's time for a little action !!!



Related Work

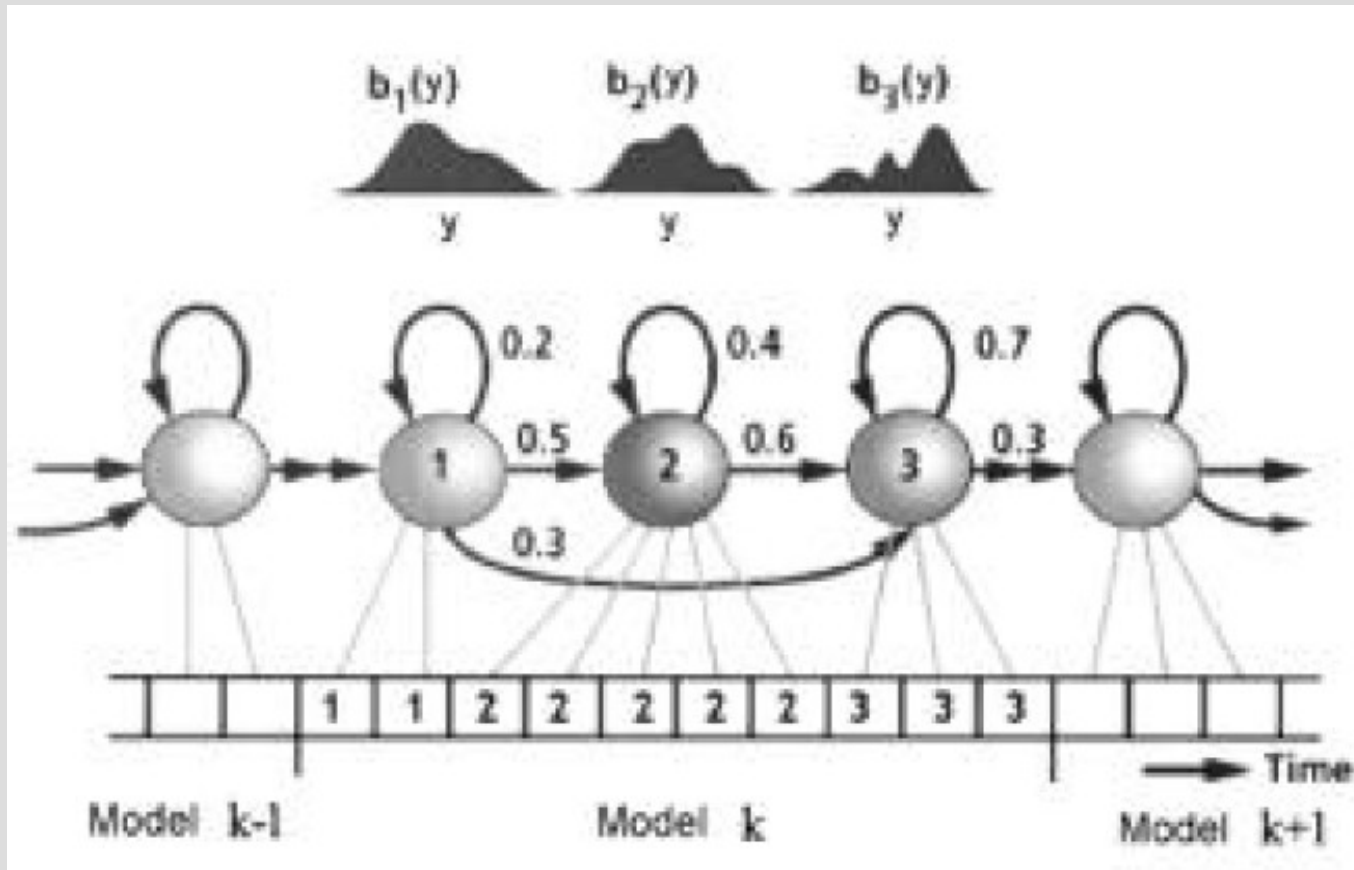
On-line time warping [1]



[1] Dixon, S. "Live tracking of musical performances using on-line time warping". 8th International Conference on Digital Audio Effects (DAFx 2005), Austrian Research Institute for Artificial Intelligence, Vienna (September 2005)

Related Work.

Hidden Markov Models [2-3]



[2] Cano, P., Loscos, A., Bonada, J.: Score-performance matching using hmms. InICMC 1999, Audiovisual Institute, Pompeu Fabra University, Spain (1999)

[3] Orio, N., D'echelle, F.: Score following using spectral analysis and hidden Markov models. In: Proceedings of the ICMC, pp. 151–154 (2001)

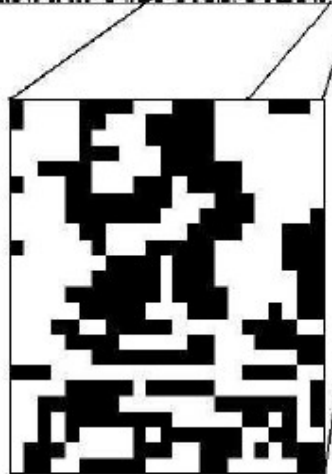
Drawbacks of state-of-the-art approaches

- Both approaches (HMM and on-line time warping) have the markov property
- These methods assume that the tracking up to the current time is correct
- Only local information is used at each step of the tracking process
- Since errors are not contemplated, recovery is not even considered
- Errors are cumulative, such systems may loose track completely
- Design issues both in HMM and in on-line time warping

Our Proposal

- Take the problem to the area of searching in metric spaces
- Consider the musical performance as a sequence of short excerpts of one second.
- Search every incoming excerpt inside the score
- The search is performed among all possible segments of the score
- Use a proximity index for fast searching (Ej BK-tree)
- Query the index for the K-nearest neighbors, choose among them based on heuristics
- This approach allows for recovery in case of tracking errors (errors are not cumulative)
- The tracking could be done even if the musical performance has already started (no synchronization required)

Extracting the Multi-Band Espectral Entropy Signature (MBSES)



(a) Original



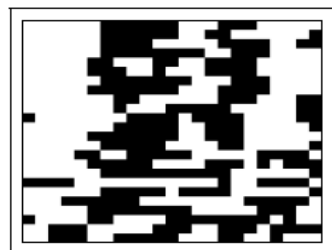
(b) Equalized



(c) Low-pass filtered



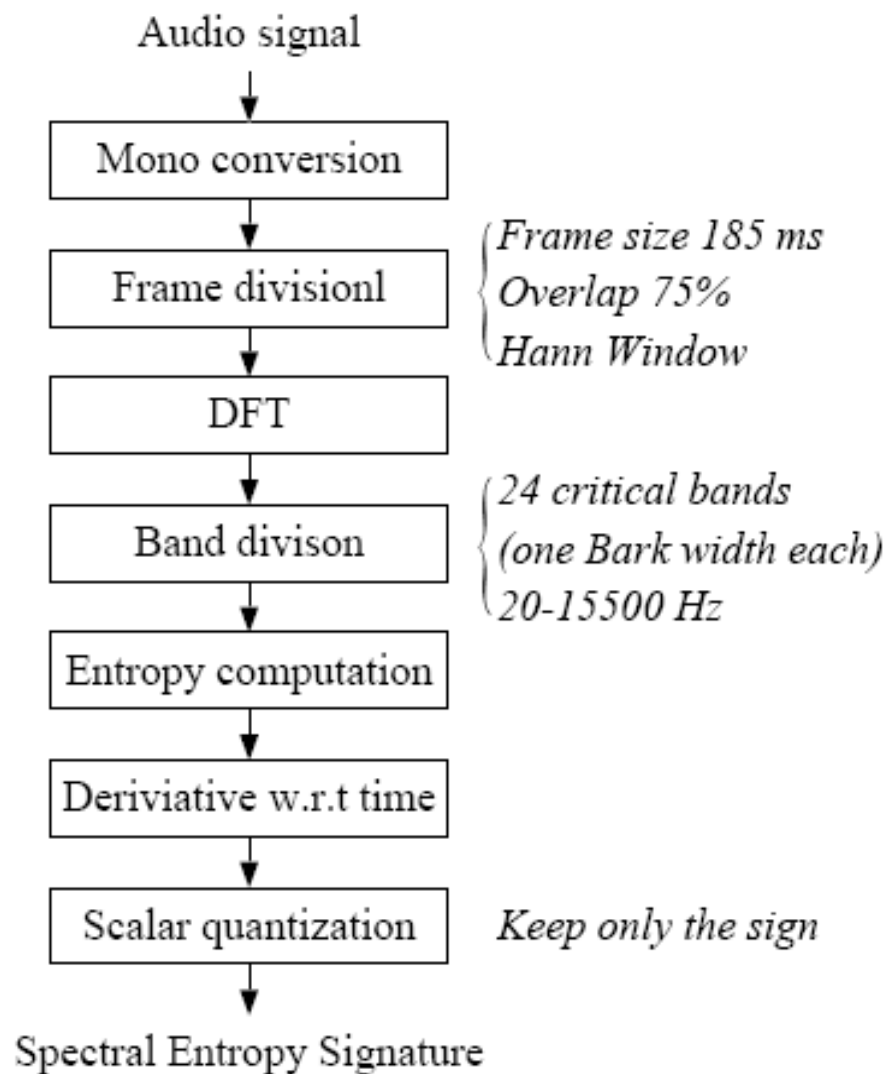
(d) Noisy



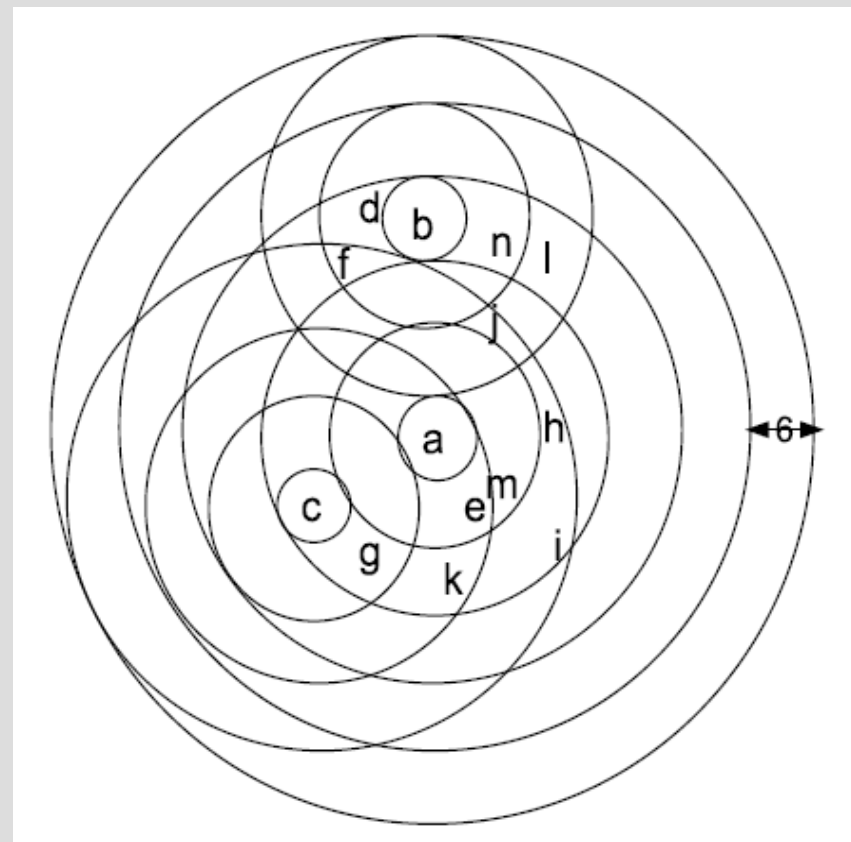
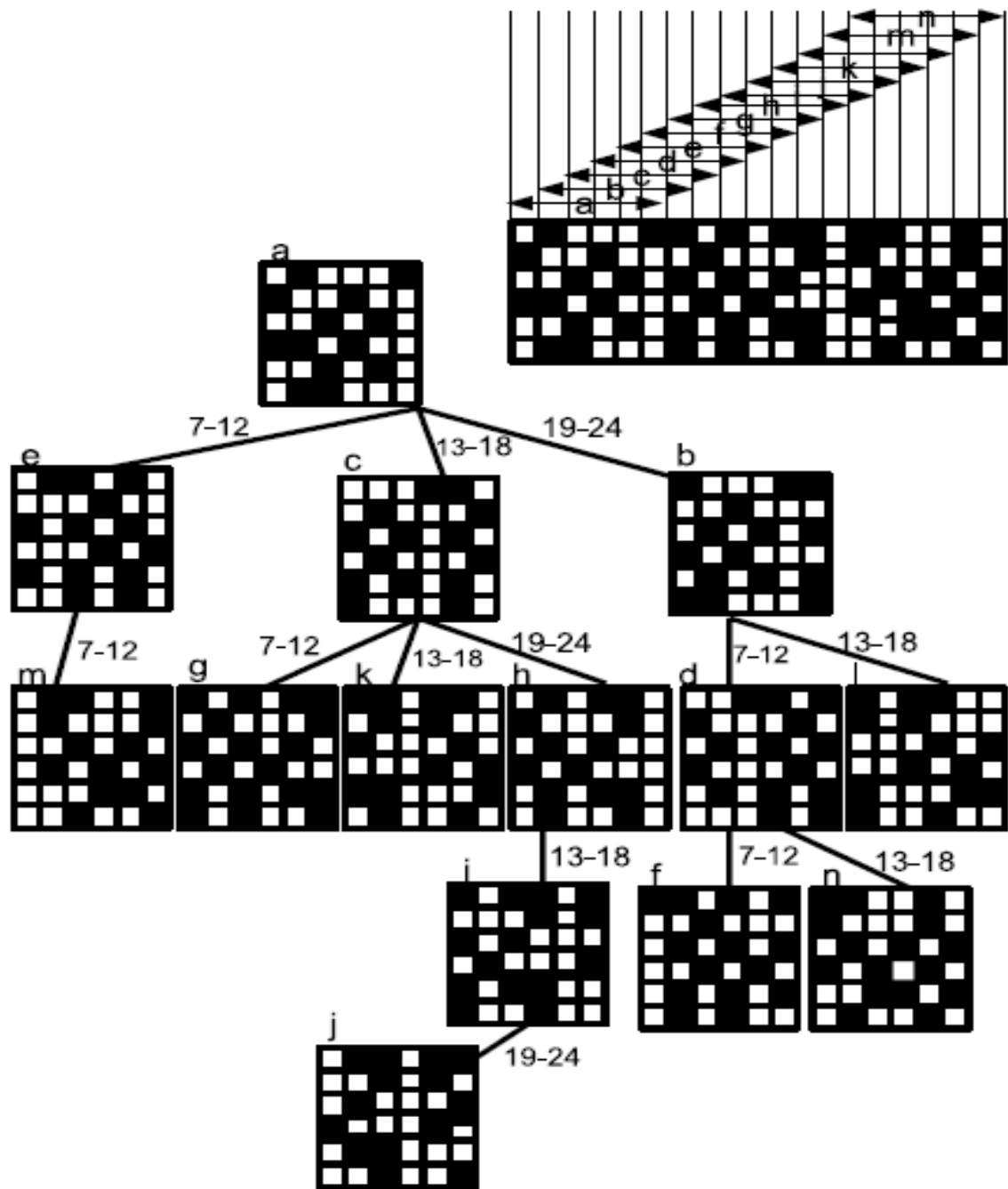
(e) Lossy compressed



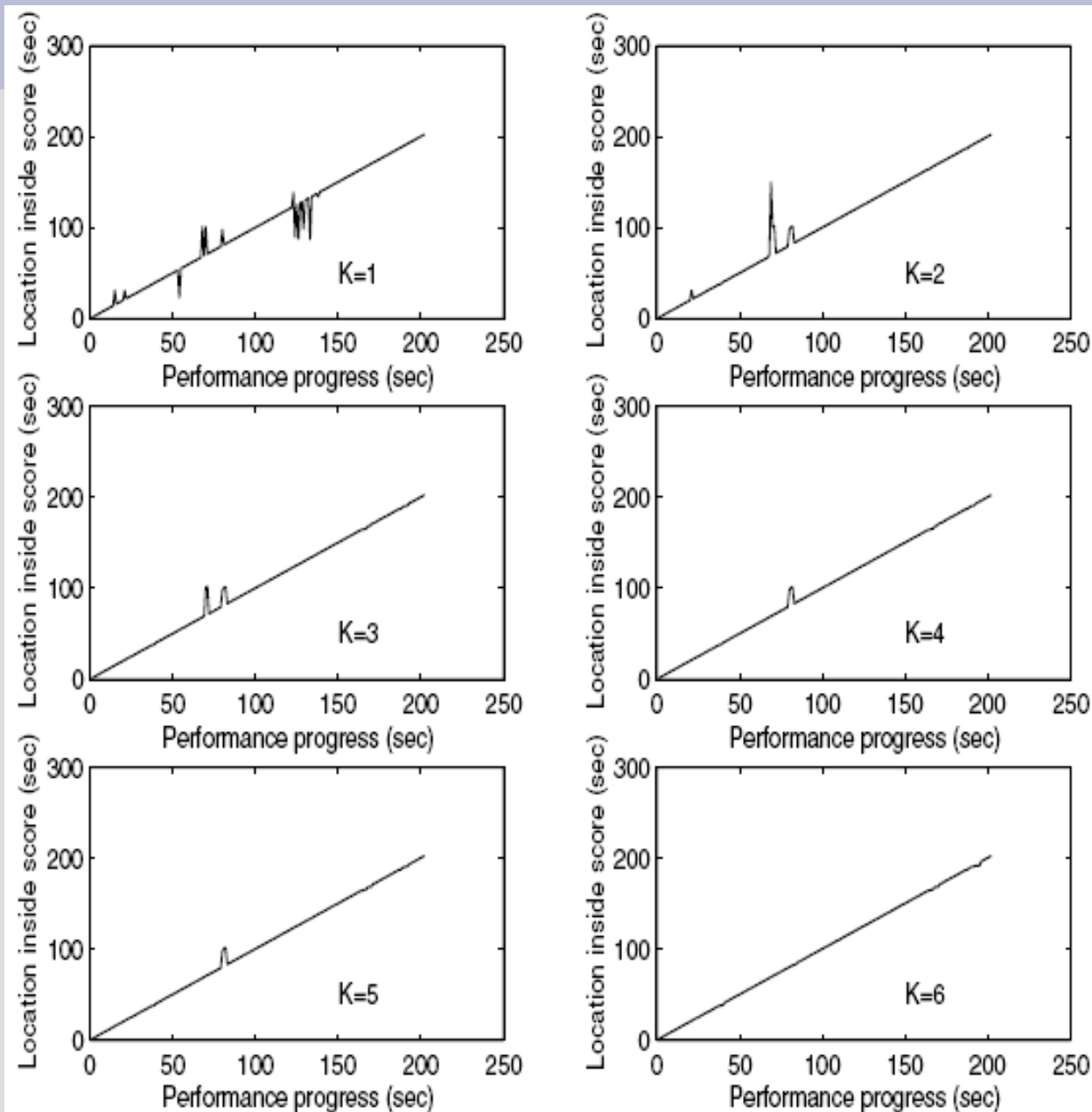
(f) Ls-Mic



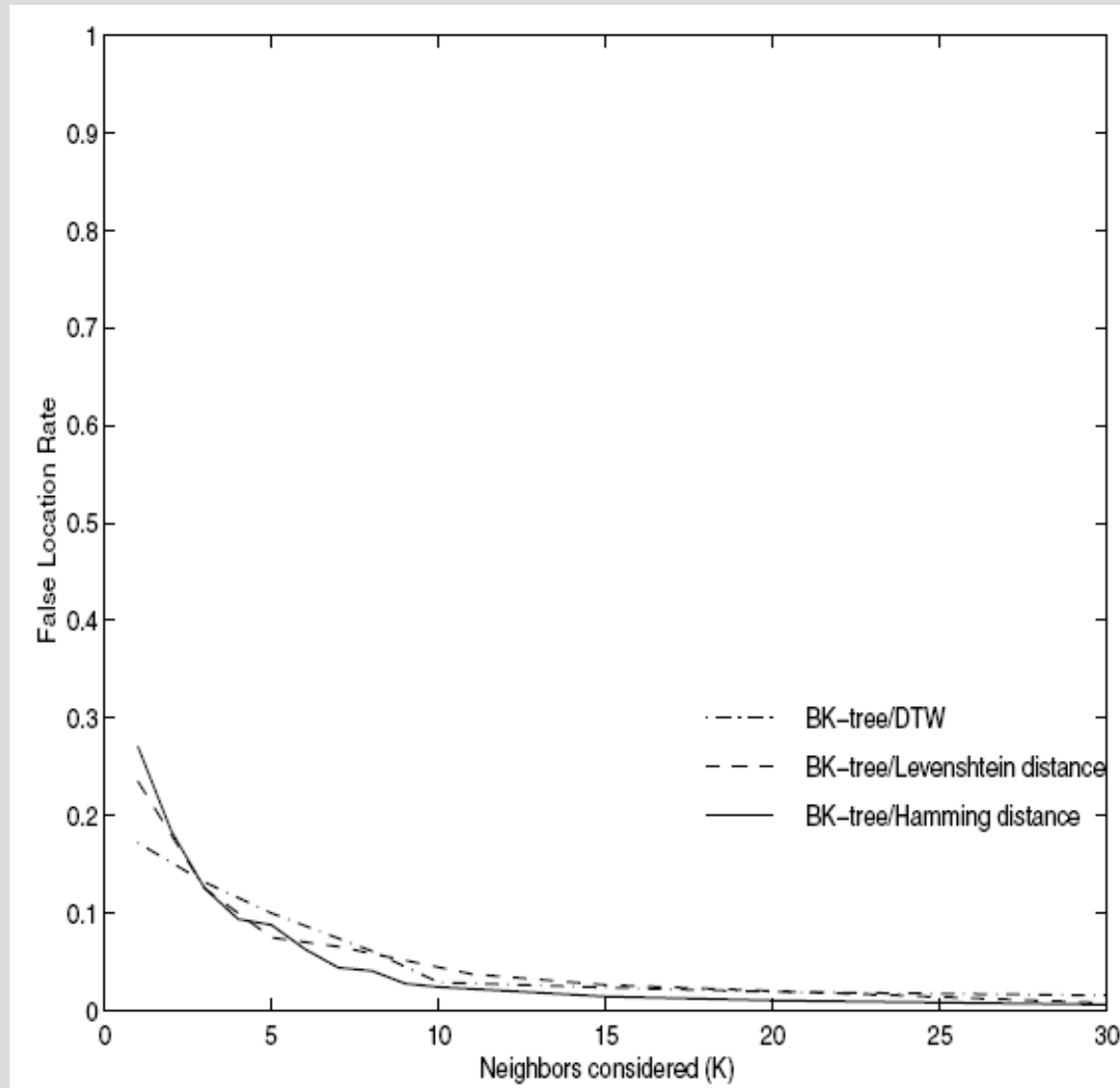
BK-tree



Tracking of a Musical Performance



Location errors vs K



Avoiding False Locations

- False locations occurrences should not lead to execute actions.
- False locations appear due to the fact that none of the K nearest neighbors occur near the current tracking position in time.
- When none of the K nearest neighbors occur near the current tracking position then the tracking position in the score should not move as if the last short segment of audio had not been received.
- With this modification peaks no longer occur, not even for $K=1$.

Time Analysis

- Two steps are required to find the location inside the score of a **one-second** segment of audio taken from the musical performance under tracking
- 130 ms for determining the audio-fingerprint of the audio signal
- Only 10 ms for searching the K nearest neighbors of using a BK-tree proximity index
- Dual core pentium laptop with 1GB of RAM
- Building the index of the score of a song about 4 min long means indexing about 5300 audiofingerprints, this takes about 10 seconds, however, it is an off-line process.

Conclusions

- It is possible to track a musical performance through excerpt searching instead of traditional alignment techniques
- Unlike HMMs there's no need for training (parameters estimation)
- Unlike the “on-line time warping” technique, the fact that the score is known a-priori is exploited
- No Markov assumption
- Errors are not cumulative (recovery is automatic)
- Tracking may start at any time not only at the beginning of the musical performance

Future Work

- Try alternative proximity indexes
- Use POMDPs (Partially Observable Markov Decision Processes)
- Automatic closed-captioned labelling of films for the hearing impaired

Thanks!

- camarena@umich.mx
- elchavez@umich.mx

Automatic closed-captioned labelling of films for the hearing impaired

