

Doğaç Başaran

Thesis Supervisor: Prof. Emin Anarım
Thesis Co-supervisor: Assoc. Prof. Ali Taylan Cemgil

Model Based Multiple Audio Sequence Alignment

Abstract

It is increasingly more common that an occasion is recorded by multiple individuals with the proliferation of recording devices such as smart phones. When properly aligned, these recordings may provide several audio and visual perspectives to a scene which leads to several applications in restoring, remastering and remixing frameworks in various fields.

In this study, we interpret the problem of aligning multiple unsynchronized audio sequences in a probabilistic framework. In this manner, we propose a novel, model based approach where we define a template generative model. We define 6 different generative models using this template covering basically all kinds of features (real valued, positive, binary and categorical). Proper scoring functions that evaluates the quality of an alignment are derived from each model where we are able to penalize non-overlapping alignments and alignment of a single sequence against a pre-aligned sequences. Having defined a cost or score function, a heuristic sequential search algorithm and a Gibbs sampler approach are proposed to find the optimum alignment of sequences on the surfaces defined by derived score functions. In addition we propose a multi resolution alignment algorithm where we combine Sequential Monte Carlo (SMC) samplers and proposed sequential search method. The models and appropriate features are exhaustively evaluated with artificial and real-life data sets. The simulation results suggest that the approach is able to handle difficult, ambiguous scenarios and partial matchings where simple baseline methods such as correlation, fail.

PUBLICATIONS

Journal Papers

1. D. Basaran, A. T. Cemgil and E. Anarım, "A Model Based Approach for Multiple Audio Sequence Alignment", IEEE Transactions on Audio, Speech and Language Processing, Vol. 23, No. 7, pp, 1160-1171, 2015.

Conference Papers

1. D. Basaran, A. T. Cemgil and E. Anarım, "SMC Samplers for Multiresolution Audio Sequence Alignment", Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pp. 201-205, IEEE, 2013.
2. D. Basaran, A. T. Cemgil and E. Anarım, "Belirlenimci Benzerlik Yöntemlerine Dayalı Model Tabanlı Ses Dizisi Hizalaması", in Signal Processing and Communications Applications Conference, IEEE 21th, 2013.
3. D. Basaran, A. Cemgil and E. Anarım, "Model Tabanlı Ses Dizisi Hizalanması" in Signal Processing and Communications Applications (SIU), 2011 IEEE 19th Conference on, pp. 606 -609, April 2011.

4. D. Basaran, A. T. Cemgil and E. Anarim, "Model Based Multiple Audio Sequence Alignment", Applications of Signal Processing to Audio and Acoustics (WASPAA) 2011 IEEE Workshop on, pp. 13-16, IEEE, 2011.

Defense Jury Members

- | | |
|-----------------------------------|---------------------|
| 1. Prof. Emin Anarım | Boğaziçi University |
| 2. Assoc. Prof. Ali Taylan Cemgil | Boğaziçi University |
| 3. Prof. Aşşın B. Ertüzün | Boğaziçi University |
| 4. Assoc. Prof. Murat Saraçlar | Boğaziçi University |
| 5. Assoc. Prof. Barış Bozkurt | Koç University |

Defense Date: 05.06.2015