JOINT ESTIMATION OF CHORDS AND DOWNBEATS

Hélène Papadopoulos
IRCAM / CNRS-STMS
Sound Analysis/Synthesis Team, Paris - France
papadopo@ircam.fr

Geoffroy Peeters
IRCAM / CNRS-STMS
Sound Analysis/Synthesis Team, Paris - France
peeters@ircam.fr

1 INTRODUCTION

This document describes a submission to the Music Information Retrieval Evaluation eXchange in the Audio Chord Recognition task. We propose a new technique for joint estimation of the chord progression and the downbeats from an audio file. Musical signals are highly structured in terms of harmony and rhythm. Integrating knowledge of mutual dependencies between chords and metric structure allows us to enhance the estimation of these musical attributes. We propose a specific topology of hidden Markov models that allows us to model chord dependence on metric structure. We consider pieces with complex metric structures such as beat addition, beat deletion or changes in the meter. With the proposed system, downbeat positions of a music piece can be estimated in terms of its harmonic structure and conversely the chord progression estimation benefits from considering the interaction between the metric and the harmonic structures.

2 OVERVIEW OF THE MODEL

We consider an ergodic $I * K$-states HMM where each state s_{ik} is defined as an occurrence of a chord c_i, $i \in [1; I]$ occurring at a "position in the measure" (position of a beat or tatum inside a measure) pim_k, $k \in [1; K]$:

$$s_{ik} = [c_i, pim_k].$$

Here, our chord lexicon is composed of $I = 24$ Major and minor triads (C Major, ..., B Major, C minor, ..., B minor). The chord sequence and the downbeat positions are simultaneously estimated. We consider pieces in predominant $3/4$ and in predominant $4/4$ meter. In both cases, the transition matrix will allow 4 beat positions in the measure. When a song is analyzed, the algorithm chooses between two metrical structure situations. The first one corresponds to the case of songs in 4/4 meter with ternary passages. In this case, we favour measures of 4 beats but transitions to parts with measures of 3 beats are allowed. The second one corresponds to the case of songs in 3/4 meter with passages in 4/4. In this case, we favour measures of 3 beats but transitions to parts with measures of 4 beats are allowed. Each state in the model generates with some probability an observation vector $\mathbf{O}(t_m)$ at time t_m. This is defined by the observation probabilities. Given the observations, we estimate the most likely chord sequence over time and the downbeat positions in a maximum likelihood sense.

3 FLOWCHART OF THE MODEL

The flowchart of the system is represented in Fig. 1.

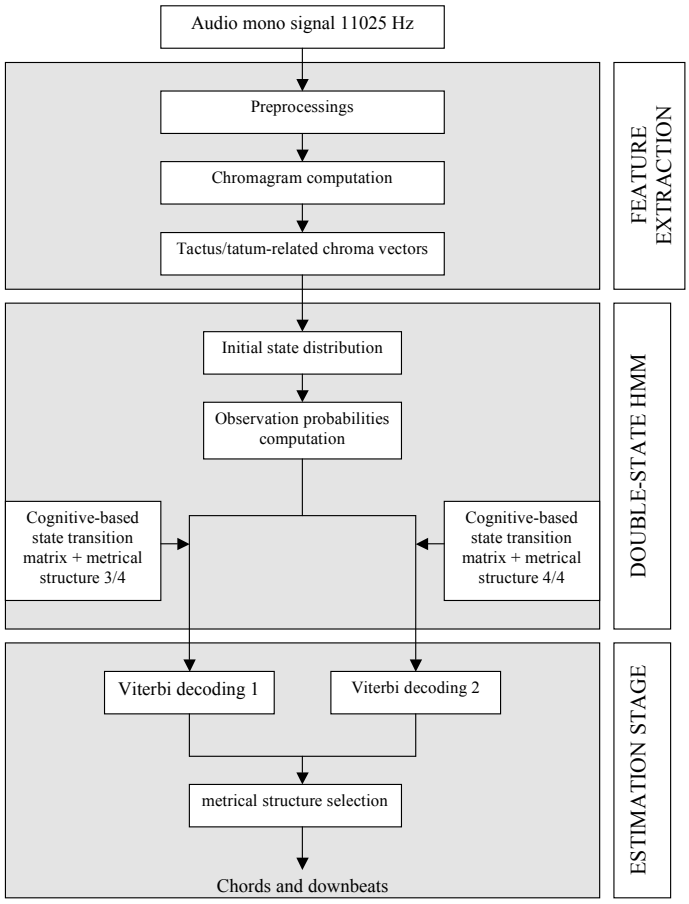

Figure 1. Simultaneous chord progression and downbeat estimation system.