International Society for Bayesian Analysis, 9^{th} World Meeting,
Hamilton Island, Australia, 2008.

Models and Inference for Musical Structure Analysis

Simon Godsill[1⋆], Taylan Cemgil[1], Paul Peeling[1] and Nick Whitely[1]

[1] University of Cambridge, Cambridge, UK
⋆ sjg@eng.cam.ac.uk

In this paper we describe recent advances in computer understanding of musical audio signals. The objectives of the work are to extract high-level and meaningful information from musical audio recordings in the form of such things as musical pitch, timbre, timing and instrument identity. These tasks are of use in themselves, but they also feed into other related tasks such as automated remixing, source separation and score-based alignment. This is a highly complex class of problems, and one which can currently only be performed accurately by trained musicians. In our research we propose Bayesian hierarchical models which represent (at the highest level) the musical score and at the lowest level generative models for the measurement of raw audio data, as captured and digitised through one or more microphones. The models attempt to capture the dynamics of the musical score in a generic and musically meaningful fashion, favouring both likely transitions over time and likely groupings of notes at a given time. These models are connected to lower level signal models, based on either expensive time-domain oscillator models or cheaper point process and Gaussian models in the time-frequency plane. Inference is of course a complex task and we discuss adaptations of MCMC, variational methods and sequential Monte Carlo work. For links to our recent work in this area see www-sigproc.eng.cam.ac.uk/∼ sjg and www-sigproc.eng.cam.ac.uk/∼ atc27.