ISE575/CSCI575

Week 5
February 10, 2005

Mohammad Anwar Hossain
Student ID: 8813190224

Analysis of the Meter of Acoustic Musical Signals

By - Anssi P. Klapuri, Antti J. Eronen, Jaakko T. Astola

In one sentence, what the paper is about?

The paper proposed a method to estimate musical meter (basic pattern of beats in a piece of music) for a wide range of genres and also at 3 different hierarchical levels - tatum pulse level, tactus pulse level and musical measure level.

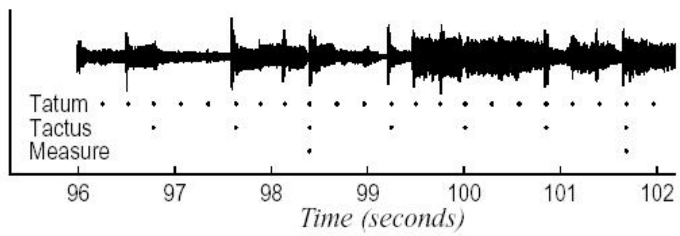
Why Meter analysis?

Meter analysis is an essential part of understanding music signals and innate cognitive ability of humans even without musical education. Meter analysis could framework facilitates the cut-and-paste operations and editing of music signals. It enables synchronization with light effects, video, or electronic instruments, such as a drum machine. In a disc jockey application, metrical information can be used to mark the boundaries of a rhythmic loop or to synchronize two or more percussive audio tracks. Meter estimation for symbolic (MIDI) data is required in time quantization, an indispensable subtask of score typesetting from keyboard input.

Background on the topics and the previous efforts:

Out of the 3 different hierarchical levels of musical meter, the most prominent level is the tactus,often referred to as the foot tapping rate or the beat.

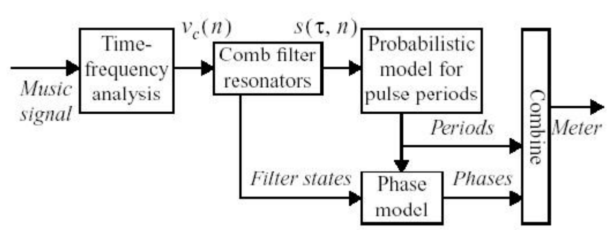
The tatum pulse has its name stemming from "temporal atom". The period of this pulse corresponds to the shortest durational values in music that are still more than incidentally encountered.The musical measure pulse is typically related to the harmonic change rate or to the length of a rhythmic pattern.



Some approaches taken earlier were – rule based system by Dixon, probabilistic generative model by Cemgil/Kappen, meter tracking using multi-agents by Goto/Muraoka and so on. Most of the previous work on meter estimation has concentrated on symbolic (MIDI) data and typically analyzed the tactus pulse only. Even though some of the systems could be extended to process audio signal by employing onset detectors, but they will not be robust on diverse acoustic material.

Proposed Methods:

The motivation of the proposed method is to utilize metrical information in further signal analysis and classi?cation, more exactly, in music transcription. The diagram on the right side shows the overview of the method. For time-frequency analysis part, a new technique is proposed which aims at measuring the degree of accentuation in acoustic signals. The technique is robust to diverse acoustic material and and is actually a synthesis and generalization of two earlier state-of-the-art methods by Goto & Scheirer. Feature extraction for the pulse period and phase analysis is performed using comb ?lter resonators very similar to those used by Scheirer.



This is followed by a probabilistic model where the period lengths of the tactus, tatum, and measure pulses are jointly estimated and temporal continuity of the estimates is modelled. At each time instant, the periods of the pulses are estimated ?rst and act as inputs to the phase model. The probabilistic models encode prior musical knowledge and lead to a more reliable and temporally stable meter tracking. Both causal and noncausal algorithms are presented.

Results and Performance Comparison:

Musical pieces from a range of genres like Classical, Electronic, Dance, Hip hop, Rap, Jazz, Blues, Rock, Pop, Soul, R&B, Funk etc were used in the experiment. The result was compared with 2 state-of-the-art reference systems by Scheirer and Dixon (these 2 systems were very slightly modified to fit best in this case).

The performance rates used are *Correct, Accept d/h, Period Correct*. Among these it seems *Accept d/h* (which means – a pulse estimate is accepted if its phase is correct and period matches either .5, 1 or 2 the annotated reference) gives a single best number to characterize the performance of the system.

Perf1: Tactus estimation perfomance (%) of dif ferent methods

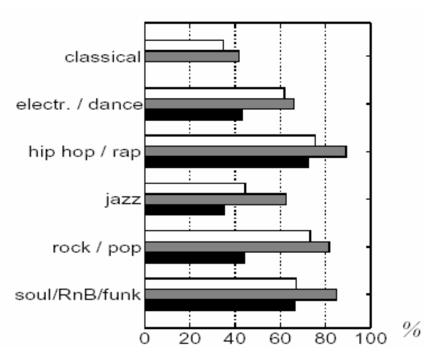
Method	Continuity required			Individual estimates		
	correct	accept d/h	period correct	correct	accept d/h	period correct
Causal	57	**68**	74	63	78	76
Noncausal	59	**73**	74	64	80	75
Scheirer [17]	27	**31**	30	48	69	57
Dixon [11]	7	**26**	10	15	53	25
O + Dixon	12	**39**	15	22	63	30

Perf2: Meter estimatio n performance for the proposed method

Method	Pulse	Continuity required			Individual estimates		
		correct	accept d/h	period correct	correct	accept d/h	period correct
Causal	tatum	44	**57**	62	51	**72**	65
	tactus	57	**68**	74	63	78	76
	measure	42	**48**	**78**	43	51	81
Non-causal	tatum	45	**63**	62	52	**74**	65
	tactus	59	**73**	74	64	80	75
	measure	46	**54**	**79**	47	55	81

As the ?rst observation, it was noticed that the reference methods did not maintain the temporal continuity of acceptable estimates. For this reason, the performance rates are also given as percentages of individual acceptable estimates. Dixon's method has dif?culties in choosing the correct metrical level for tactus, but performs well according to the "accept d/h" criterion when equipped with the new onset detector. The proposed method outperforms the previous systems in both accuracy and temporal stability.

Perf3: Performance for different Genres



Perf4: Performance for different system configurations

Method	Require continuity, accept d/h			Individual estimates, accept d/h		
	tatum	tactus	measure	tatum	tactus	measure
0. Baseline	**62**	**73**	**54**	74	80	55
1. No joint estim.	**58**	**68**	**49**	71	75	50
2. No temporal proc.	**45**	**54**	**31**	72	77	50
3. Neither of the two	**41**	**50**	**25**	70	72	44

Perf3 shows shows the "accept d/h" (continuity required) performance rates for different musical genres. For classical music, the proposed method is only moderately successful, although e.g. the tactus rate still outperforms the performance of the reference methods for the whole material. However, this may suggest that pitch analysis would be needed to analyze the meter in classical music. In jazz music, the complexity of musical rhythms is higher on the average and the task thus harder.

Perf4 shows performance for different configuration of the proposed system. In the ?rst test, the dependencies between the different pulse levels were broken by using a non-informative (?at) distribution. This slightly degrades the performance in all cases. In the second test, the dependencies between temporally successive estimates were broken by using a non-informative distribution for the transition probabilities between successive period and phase estimates, and respectively. This degrades the temporal stability of the estimates considerably and hence collapses the performance rates which use the longest continuous correct segment for evaluation. In the third case, the both types of dependencies were broken. The system still performs moderately, indicating that the initial time-frequency analysis method and the comb-?lter resonators provide a rather high level of robustness as such.

Conclusion:

The critical elements of a meter estimation system appear to be the initial time-frequency analysis part which measures musical accentuation as a function of time and the (often implicit) internal model which represents primitive musical knowledge. The former is needed to provide robustness for diverse instrumentations in e.g. classical, rock, or electronic music. The latter is needed to achieve temporally stable meter tracking and to ?ll in parts where the meter is only faintly implied by the musical surface. A challenge in the latter part is to develop a model which is generic for various genres, for example for jazz and classical music. The proposed model describes suf?ciently low-level musical knowledge to generalize over different genres.