Statistical methods for time signature inference

Pablo Hernán Rodríguez Zivic
prodriguez@dc.uba.ar
University of Buenos Aires - CONICET

Guillermo Cecchi
gcecchi@us.ibm.com
T.J. Watson IBM Research Center.

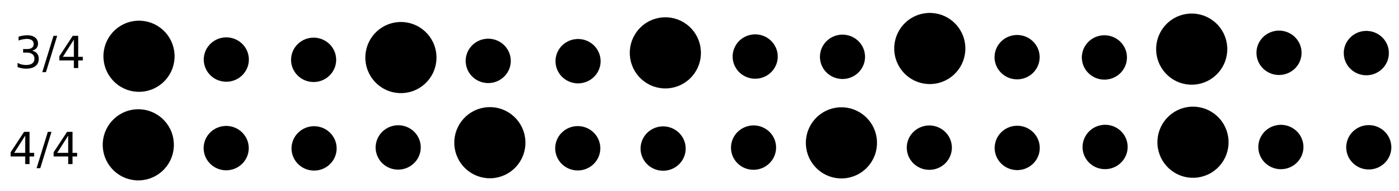
Objective

We want to quantify an elusive perceptual quality of a song: its time signature. We are using a standard classification framework to be able to make an objective evaluation and perceptual interpretation.

Background

What is the time signature of a song?

Before that, we must talk about meter. Meter refers to the **periodic structure** of beats that appears in almost every musical piece. Think of that big rhythmic difference between a Waltz and the Beethoven's 9th symphony.
This periodic structure is not present in the **music surface**, instead the listener must **infer** it from perceptual clues, like the relative volume of events.
There are some **standard metrical structures**. These are notated in a musical score with a fraction, the **time signature**, which its numerator refers to the period of a perceptual salient pulse. For example, a Waltz has a 3/4 time signature andBeethoven's 9th symphony has a 4/4.



A schematized version of two pulses. First pulse has a strong beat every 3 units. Second pulse every 4.

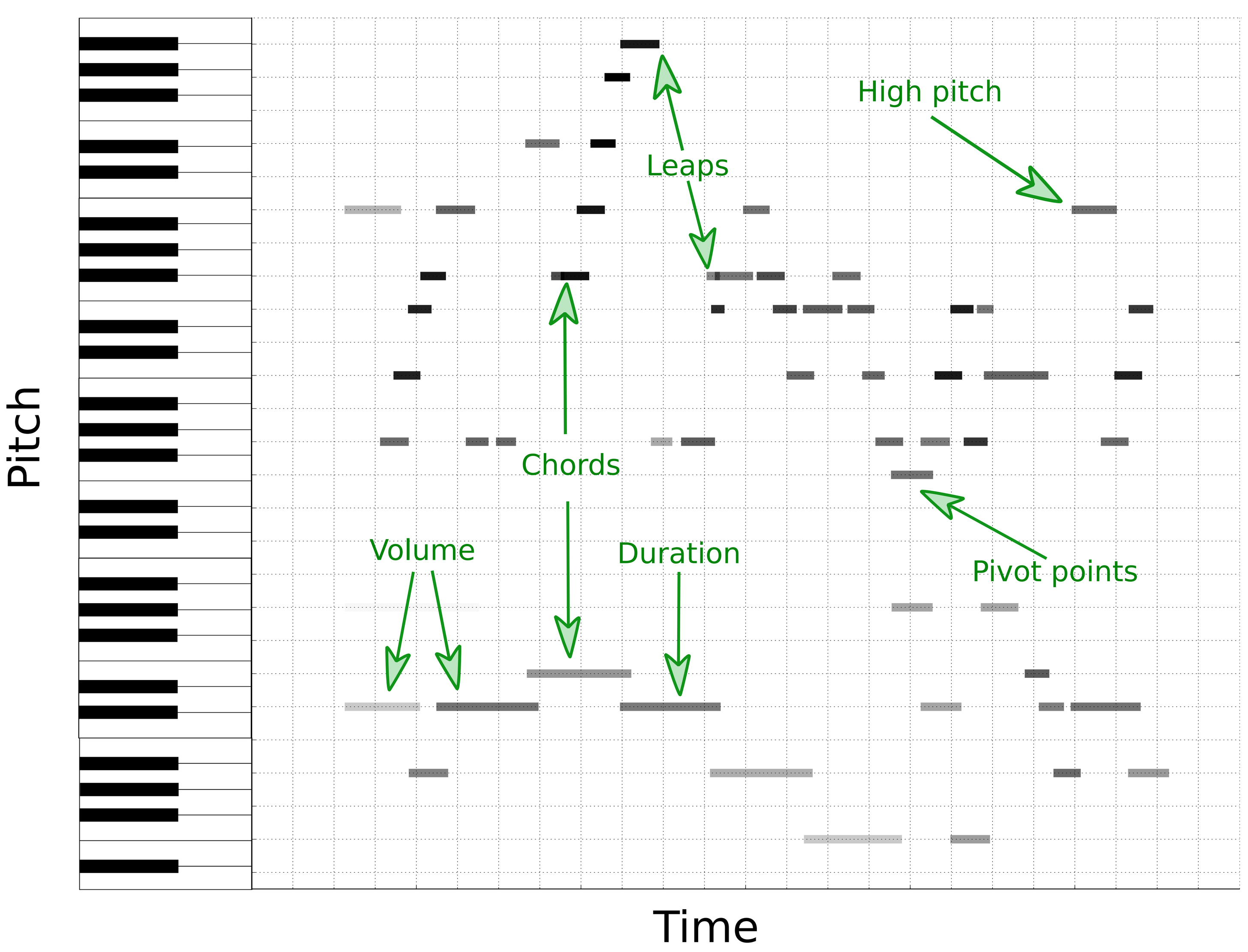
Previous work

Most of the work have been focused on **ad-hoc algorithms** to do such inference (Brown, 1993; Temperley and Sleator 1999; Meudic 2002; Gouyon and Herrera, 2003b). They all use some perceptual features to create an heuristic to compute the time signature of a given song.

There are two main problems with such heuristics:
1) They are **difficult to evaluate objectively**.
2) It's not clear how a perceptual conclusion can be made from such algorithms.

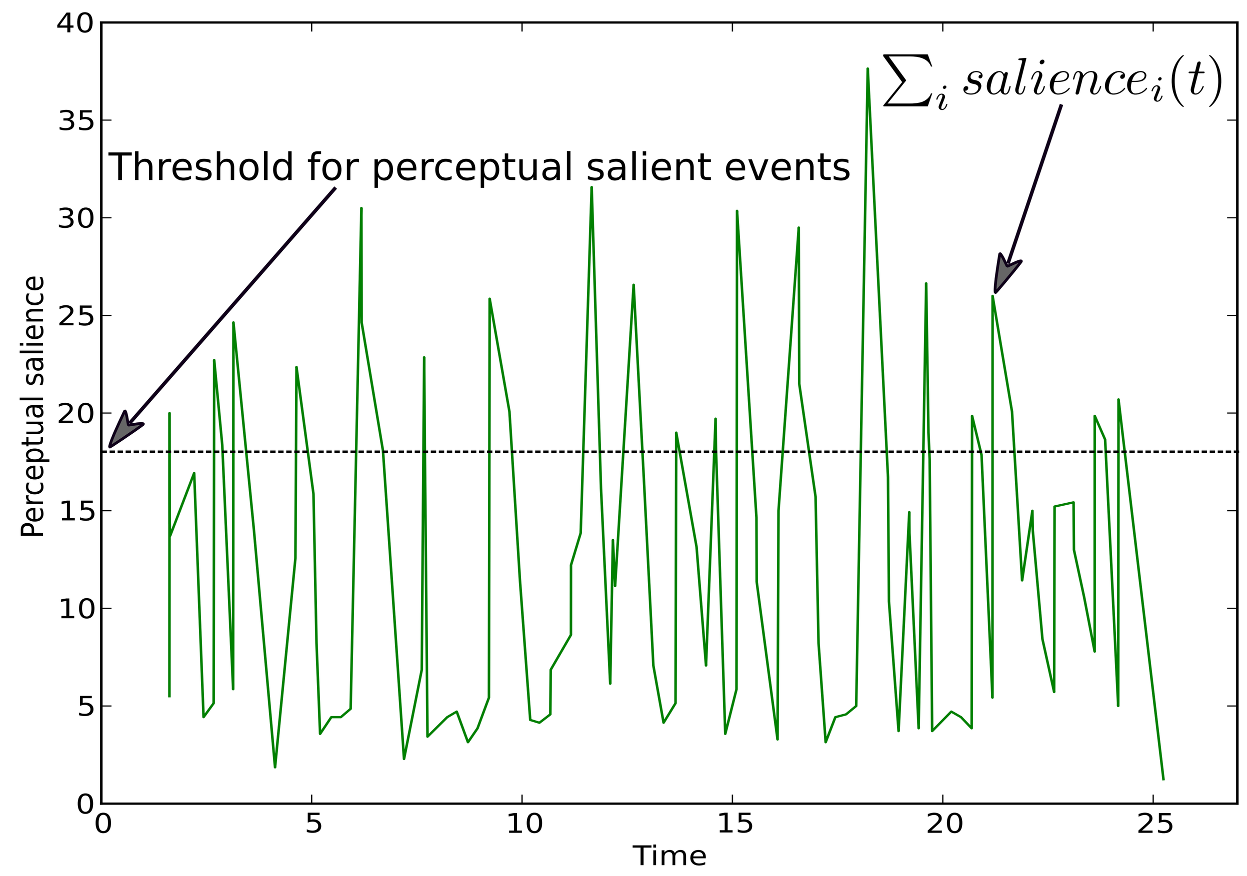
Perceptual features

There are several features that have a strong correlation with the time signature of a given song. Examples of these are that **long events**, with **high pitch**, etc. For a survey, see Hannon et al. (2004). This figure shows in a piano-roll notation the features used to score the perceptual salience of an event.
The horizontal axis represents time, and vertical axis represents pitch.



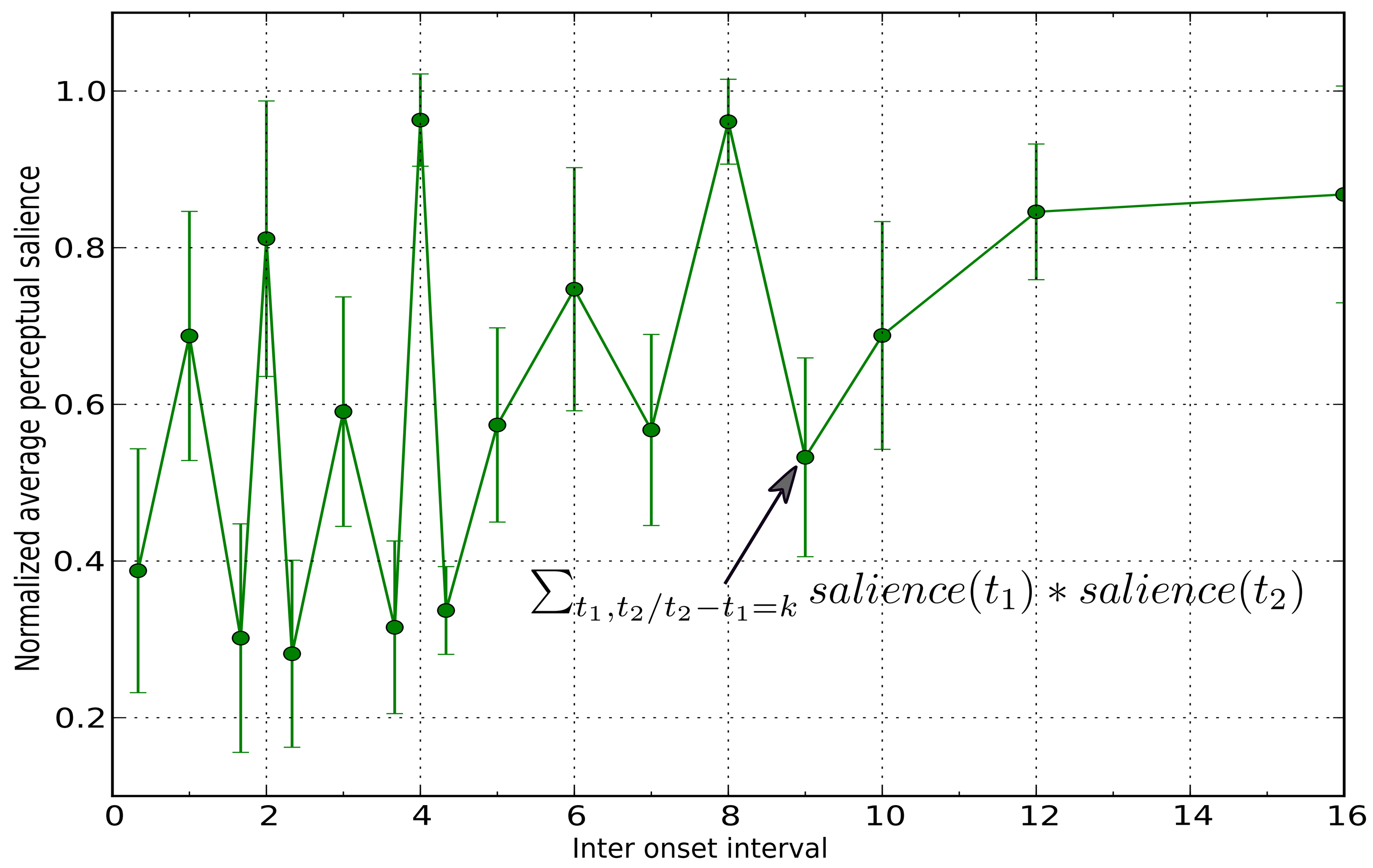
Additive interaction

Using the fact that there is an interaction which seems additive between the salience features (Ellis and Jones, 2009), it seems reasonable to sum them for a given moment. Afterwards, only the most prominent events will be taken into account. This follows from the fact that only the most prominent events are the ones that will be used for meter inference (Lerdhal and Jackendoff, 83).



Final feature vector

Since what we are looking for is a regular pattern, its essence is a period. Each period will have a weight which is the sum of the weights of all events it spans. This sum is only calculated for perceptual salient events.



This graph was computed for 280 songs with a 4/4 time signature, normalizing for the maximum value of the vector. Error bars show standard error.

Results and conclusion

We tested with a Naïve Bayes classifier and a Support Vector Machine. Results are shown in the next table:

	F-Measure
Naïve Bayes	0.73
Support Vector Machine	**0.92**

The fact that we got a score above chance with Naïve Bayes let us conclude that the vectorial representation is adecuate for the problem. Further more, averaged feature vectors, like the above figure, reveal the musically well known internal structure time signatures.
Besides, such an F-Measure with the SVM classifier, indicates that most of the data is explained with the features we used, so we can go further on exploring how the brain process musical rhythm.

References

Brown, J. (1993). Determination of the meter of musical scores by autocorrelation.
Ellis and Jones (2009). The role of accent salience and joint accent structure in meter perception.
Gouyon, F. and Herrera, P. (2003a). A beat induction method for musical audio signals.
Hannon et al. (2004). The role of melodic and temporal cues in percieving musical meter.
Lerdhal and Jackendoff (1983). A generative theory of tonal music.
Meudic, B. (2002). Automatic meter extraction from midi files.
Temperley and Sleator (1999). Moleling Meter and Harmony: A Preference Rule Approach.