**Instructor Feedback for the Signal Processing Cup 2017**

**Team ID:** 24746
**Team Name:** Real Time Beat Tracking: IIT Bombay Entry
**Team Members:**
Yash Sanghvi, Undergraduate, Indian Institute of Technology, Bombay - sanghviyash95@gmail.com
Mehul Shah, Undergraduate, Indian Institute of Technology, Bombay - mehulshah2095@gmail.com
Krishna Reddy, Undergraduate, Indian Institute of Technology, Bombay - 130020118@iitb.ac.in
Rajbabu Velmurugan, Supervisor, Indian Institute of Technology, Bombay - rajbabu@ee.iitb.ac.in
Ashish Goyal, Postgraduate, Indian Institute of Technology, Bombay - goyal2608@gmail.com
Arunabh Ghosh, Undergraduate, IIT Bombay - arunabhghosh98@gmail.com

**Phase 1**
***Evaluation Procedure***
To subjectively evaluate the quality of the annotations (for the two submitted pieces and one challenge piece), four judges loaded each piece into the Sonic Visualiser software and overlaid the annotations as a time-instants layer. Each judge then listened back to the audio with the annotations synthesized as short percussive clicks on top in order to assess their quality. Having done so, each judge then assigned a score per piece as follows:

- **2** – *the annotations were tapped at a metrical level consistent with the opinion of the judge, in-phase (i.e. on the perceived "on-beat"), and with sufficiently accurate temporal localization.*
- **1** – *the annotations were tapped on the "off-beat", or at some related metrical level to one considered best by the judge, or the annotations had minor localization issues.*
- **0** – *the annotations could not be understood as related to the beat by the judges, or had severe temporal localization issues, and were thus considered unacceptable.*

For each of the three pieces, the mean of the four ratings was recorded.

After the subjective ratings of the annotations had been collected, and prior to the release of the data teams, all annotation sequences were verified again, and those in need of modification were updated. This was achieved by the manual correction of the annotations, again using the Sonic Visualiser software. In addition to the subjective annotation score described above, an additional **1** point per piece was given if no manual correction was required.

Thus, for any team the maximum possible score would be **9**, if all judges scored each excerpt **2** points, and no manual correction was required. This accumulated score, out of **9**, was multiplied by 100/9 to give a final total **(A)** out of **100**.

**Phase 2**
***Evaluation Procedure***
*Beat tracking evaluation process:* The evaluation of the beat tracking algorithms followed the well-established continuity-based approach for which a MATLAB reference implementation (`beatEvaluator2.m`) was provided in the Resources of the SP Cup on Piazza. Around each annotation a tolerance window was defined as +/-17.5% of the inter-annotation-interval. Beat times were considered accurate if they fell within this tolerance window, and the inter-beat-interval was within +/- 17.5% of the inter-annotation-interval. In addition to calculating accuracy at the annotated metrical level, alternative scores were also calculated using a set of excerpt-specific alternative metrical levels. Per excerpt, the maximum across this range of scores was taken as the accuracy. Each beat tracking algorithm was evaluated on the Closed (25 excerpts) and Challenge piece (33 excerpts) datasets, i.e. those for which the ground annotations were not made available to the teams.

Due to a request by several teams based on their difficulties in obtaining a precise synchronization between the audio playback and microphone capture on their embedded devices, each evaluation score was recalculated with the predicted beat times shifted back up to 300ms (in 10ms increments). Since this offset was not consistent, the score corresponding to the "best" offset per excerpt was recorded and the overall mean score was taken from these values. In order to enable a fair comparison across all submissions, the same offset correction was applied to all submitted beat predictions.

The final score **(B)** was measured directly out of 100 and thus required no additional adjustment.

*Video judging criteria:* A team of seven judges watched each video and recorded scores according to the criteria:

- Novelty of the creative application (**0-3**)
- Clarity of the relationship to beats in music (**0-3**)
- Overall Quality of Video (**0-3**)

An additional **1** point was available for following the instructions and including the hyperlink(s) to the videos in the PDF files.

The final score for the video **(V)**, out of **10**, was multiplied by 10 to give a total out of **100**.

**Overall Score**
As specified in the overview document of the SP Cup, the final score per team would be based on the weighted contribution of each component as follows:

**Total =** Annotation Score * (1/6) + Beat Tracking Score * (1/2) + Video Score * (1/3)

**Instructor Feedback**

**Phase 1:**
The audio material was well chosen and interesting. While the annotations scored well under the subjective listening tests by the judges, ultimately each sequence required some modification and thus the scoring is not as high as it could have been with greater temporal precision of the annotations.

**Phase 2:**
The written paper, while rather short (just 2 pages), was clearly presented and a nice summary of the approach. The information about the creative application was rather limited, but in the end the application itself was quite self-explanatory. That said, it is not clear from the description of the method how the measure boundaries are inferred in order to have the lights *count* 1,2,3,4 in time with the music. Nevertheless, the idea of using higher level metrical information did make for a more interesting use of lights than merely flashing a single LED on and off in time with the beat. Concerning the beat tracking algorithm, it was among the better systems submitted, yet like many algorithms it struggled in the more challenging examples with greater tempo variation.

To summarize: Great audio examples, paper could have used more detail, but the beat tracking method shows definite promise. To that end, the team may wish to consider doing further work on their system and then submitting it a music / audio related conference. There are many such conferences happening in 2017 with deadlines coming soon: http://mtg.upf.edu/research/workshops

**Awarded Scores**
**(A)** Annotation Score = 50.00
**(B)** Beat Tracking Score = 55.13 (combined for Closed and Challenge datasets)
**(V)** Video Score = 45.00

**Total = A** * (1/6) + **B** * (1/2) + **V** * (1/3)
**Total = 50.90 / 100**
**Ranking Group: middle third**

We are choosing not to reveal the precise ranking, but instead split the overall positions into thirds across all the 21 teams who completed Phase 2 including the top 3 teams. Thus, **top third** implies (ranking 1-7), **middle third** (ranking 8-14), and **lowest third** (ranking 15-21).