# COMPARATIVE EVALUATION AND COMBINATION OF AUDIO TEMPO ESTIMATION APPROACHES

JOSE R. ZAPATA AND EMILIA GÓMEZ {joser.zapata,emilia.gomez}@upf.edu
Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain

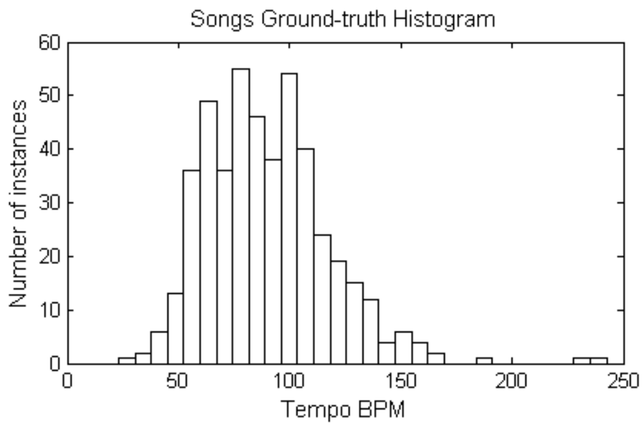
## APPROACHES

| Algorithm | Aubio | BeatIt | Beatroot | Ellis | Essentia | Fixedtempo | IBT | jAudio | MIRTempo | Mpeg7-xm | Qmtempo | Tzanetakis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | Brossier | Bonada and Gouyon | Dixon | Ellis | Aylon and Wack | Cannan | Oliveira et al. | McEnnis and McKay | Lartillot | Rohden | Davies and Plumbley | Tzanetakis |
| **Infrastructure** | Sonic annotator | Windows Binary | Java | Matlab | Essentia | Sonic annotator | Marsyas | Java | Matlab | Matlab | Sonic annotator | Marsyas |
| **Output** | Beats Positions in Time | One Bpm | Beats Positions in Time | slower and faster tempo, relation between them | One Bpm | One Bpm | One Bpm | One Bpm | One Bpm | BPM value each time the estimated tempo changes | BPM value each time the estimated tempo changes | slower and faster tempo, relation between them |
| **Feature list** | Onset detection (Complex spectral difference) | Energy envelope differences for 8 bands | Onset detection (Spectral Flux) | Onsets obtained from the Mel spectrogram | Energy enveloped differences (Novelty curves) for 5 bands | Overall energy rise function. | Onset detection (Spectral Flux) | Energy envelope (256 window) | Onset curve, (10 channel gammatone filterbank and a low pass filter) | Energy envelopes for 6 bands | Onset detection (Spectral Flux) | Onset strength signal (based on Spectral Flux) |
| **Pulse induction** | ACF | ACF | IOI Clustering | ACF | ACF | ACF | ACF | ACF | ACF | ACF | ACF | ACF |

| Algorithm | AlonsoACF | AlonsoSP | DixonACF | DixonI | DixonT | Klapuri | Scheirer | Tzan_H | Tzan_MM | Tzan_MS | Uhle | Combination |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | Alonso | Alonso | Dixon | Dixon | Dixon | Klapuri | Scheirer | Tzanetakis | Tzanetakis | Tzanetakis | Uhle | Heuristic combinatination of the results of Klapuri, Beatit, Qmtempo, Ellis, IBT(c),Essentia and Mpeg7-xm |
| **Infraestructure** | Matlab | Matlab | Matlab | Java | Java | Linux binary | Linux binary | Linux binary | Linux binary | Linux binary | Windows Binary | |
| **Output** | One Bpm | One Bpm | One Bpm | One Bpm | One Bpm | Beats in Time | Beats in Time | One Bpm | One Bpm | One Bpm | One Bpm | |
| **Feature list** | Onsets of Notes | Onsets of Notes | Downsampled and smoothed Energy of 8 frequency bands | Energy based Onset Detector | Energy based Onset Detector | The differentials of the loudness in 36 frequency subbands | Energy envelope differences for 6 bands computed by a filterbank | Energy Envelopes of 5 octave – space frequency bands obtained by Wavelets | Energy Envelopes of 5 octave – space frequency bands obtained by Wavelets | Energy Envelopes of 5 octave – space frequency bands obtained by Wavelets | Differential smoothed Energy envelopes for logarithmically space frequency bands | |
| **Pulse induction** | ACF | Spectral Product | ACF | IOI Clustering | IOI Clustering | Bankcomb Filter | Bankcomb Filter | ACF | ACF | ACF | ACF | |

## MUSIC COLLECTION

| Genre | # Songs |
|---|---|
| Rock | 68 |
| Classical | 70 |
| Electronic | 59 |
| Latin | 44 |
| Samba | 42 |
| Jazz | 12 |
| Afrobeat | 3 |
| Flamenco | 13 |
| Balkan and Greek | 144 |
| Fado | 10 |




## COMBINATION METHOD

```
Main() {
for n =1:number_of_songs
   Result(n) = klapuri(n);
   if  (Result(n) ~= BeatIt(n) & Result(n) ~= Essentia(n))
         [p, Out] =Relation(n);
         if (p ==1)
                Result(n) = Out;
         end
   End

% To avoid the double error
     if (Result(n) == BeatIt(n)*2 & BeatIt(n) = = Ellis(n))
          Result(n) = BeatIt(n);
      end
end
}
```

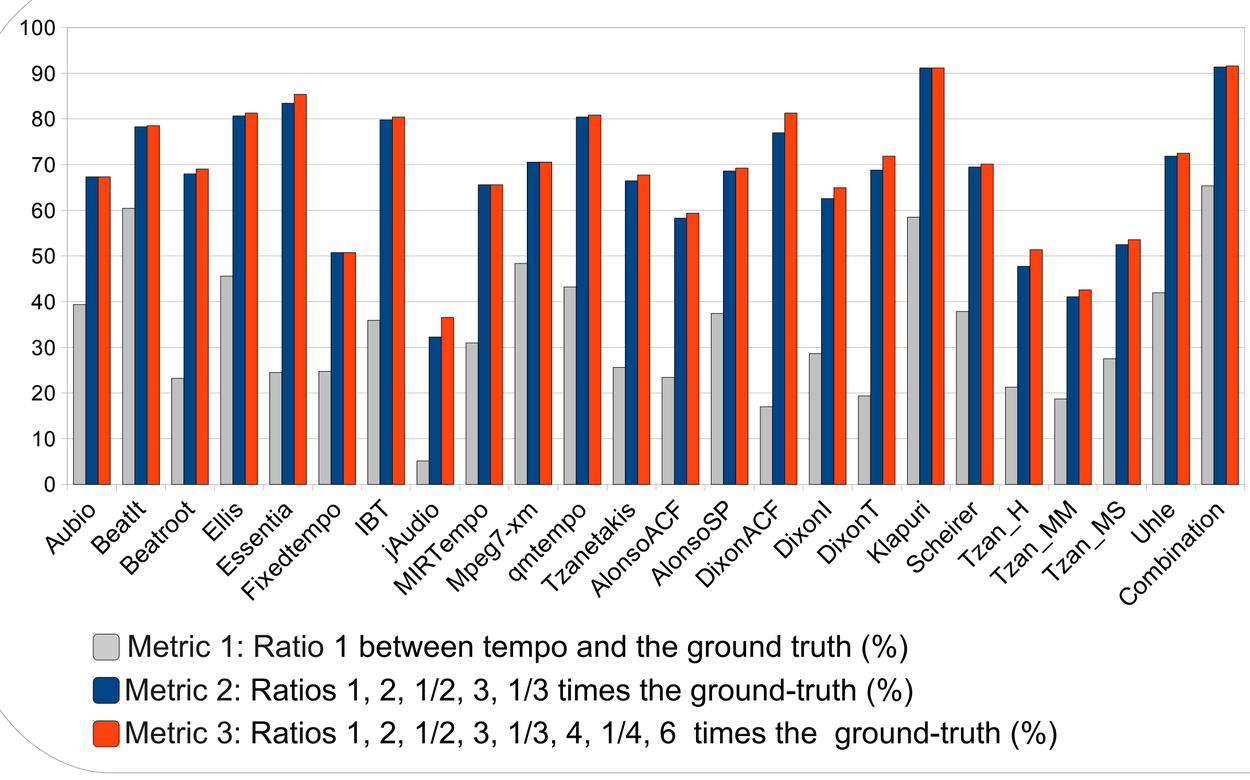
```
function [p, Out] = Relation(n)
{  array = [Beatit(n), Ellis(n), Qmtempo(n),
        IBT(n), Essentia(n), Mpeg7-xm(n)];

    if  three or more values from array are equal and
    if exist the double of these values in the array

        p  = 1;
        Out = the repeat value in the array;

    else
        p = 0; Out = 0;
    end
}
```

## RESULTS




| Algorithm | Pulse Induction | Combinig Bands |
|---|---|---|
| Klapuri | B | After |
| BeatIt | A | After |
| Ellis | A | |
| qmtempo | A | |
| IBT | A | |
| Essentia | A | After |
| Mpeg7-xm | A | After |
| Uhle | A | After |
| Scheirer | B | After |
| AlonsoSP | SP | |
| DixonACF | A | Before |
| Aubio | A | |
| MIRTempo | A | Before |
| Beatroot | IOI | |
| DixonT | IOI | |
| Tzanetakis | A | |
| DixonI | IOI | |
| AlonsoACF | A | |
| Tzan_MS | A | Before |
| Fixedtempo | A | |
| Tzan_H | A | Before |
| Tzan_MM | A | After |
| jAudio | A | |

| General Ratio | % |
|---|---|
| 1/4 | 0,03% |
| 1/3 | 0,09% |
| 1/2 | 3,62% |
| 2/3 | 1,89% |
| 3/4 | 0,62% |
| Good | 32,12% |
| 5/4 | 1,02% |
| 4/3 | 5,36% |
| 3/2 | 1,88% |
| 2 | 29,00% |
| 9/4 | 1,65% |
| 5/2 | 1,14% |
| 8/3 | 1,37% |
| 3 | 1,80% |
| 4 | 1,25% |
| 16/3 | 0,07% |
| 6 | 0,15% |
| Other | 16,94% |

## CONCLUSIONS

- The best estimation had an overall accuracy of 91% but can be improved with a heuristic combination of the best methods
- The best performing methods (Klapuri and BeatIt) are based on the following steps: differential energy values, frequency decomposition, periodicity detection prior to the multi-band integration, and multi-tempo level detection in the pulse-induction block.
- The tested algorithms involving band decomposition and periodicity detection before the multi-band integration achieved better results.

## REFERENCES

F. Gouyon, A. Klapuri, S. Dixon, M. Alonso, G. Tzanetakis, C. Uhle, and P. Cano, “An experimental comparison of audio tempo induction methods,” Audio, Speech, and Language Processing, IEEE Transactions on, vol. 14, 2006, p. 1832–1844.

F. Gouyon and S. Dixon, “A Review of Automatic Rhythm Description Systems,” Computer Music Journal, vol. 29, 2005, pp. 34-54.

**https://sites.google.com/site/tempoandbeattracking/**