

2010-01-01

On the Use of a Dynamic Hybrid Tempo Detection Model for Beat Tracking

Mikel Gainza

Technological University Dublin, Mikel.Gainza@tudublin.ie

Follow this and additional works at: <https://arrow.tudublin.ie/argcon>



Part of the [Other Engineering Commons](#), and the [Signal Processing Commons](#)

Recommended Citation

Gainza, M. (2010) On the Use of a Dynamic Hybrid Tempo Detection Model for Beat Tracking. *IEEE International Conference on Multimedia & Expo*, Singapore, 2010.

This Conference Paper is brought to you for free and open access by the Audio Research Group at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, vera.kilshaw@tudublin.ie.

Audio Research Group

Articles

Dublin Institute of Technology

Year 2010

On the Use of a Dynamic Hybrid Tempo
Detection Model for Beat Tracking

Mikel Gainza

Dublin Institute of Technology, mikel.gainza@dit.ie

— Use Licence —

Attribution-NonCommercial-ShareAlike 1.0

You are free:

- to copy, distribute, display, and perform the work
- to make derivative works

Under the following conditions:

- Attribution.
You must give the original author credit.
- Non-Commercial.
You may not use this work for commercial purposes.
- Share Alike.
If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For any reuse or distribution, you must make clear to others the license terms of this work. Any of these conditions can be waived if you get permission from the author.

Your fair use and other rights are in no way affected by the above.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License. To view a copy of this license, visit:

- URL (human-readable summary):
<http://creativecommons.org/licenses/by-nc-sa/1.0/>
 - URL (legal code):
<http://creativecommons.org/worldwide/uk/translated-license>
-

ON THE USE OF A DYNAMIC HYBRID TEMPO DETECTION MODEL FOR BEAT TRACKING

Mikel Gainza

Audio Research Group, Dublin Institute of Technology

Email: mikel.gainza@dit.ie

ABSTRACT

In this paper, an approach that estimates the times at which musical beats occur is presented. The system uses a hybrid multi-band decomposition in order to estimate the music tempo. Following this, beat events are tracked by using a dynamic programming approach, which is updated by using short time tempo estimates. The hybrid decomposition is used in order to calculate the tempo by using different onset detection functions in different frequency bands. In addition, a method that estimates which frequency bands provide reliable periodicities is also presented. The accuracy of the model is evaluated by comparing the presented system against existing approaches using a database of 474 songs.

Keywords— tempo detection, beat tracking, onset detection, dynamic programming, audio content analysis.

1. INTRODUCTION

The primary rhythmic units are denoted as “beats”, which can also be understood as the musical events located at the times at which humans tap along their feet whilst listening to music. The speed at which musical beats repeat provides the tempo of a piece of music, which is expressed in beats per minute (bpm). There are numerous applications derived from the automatic detection of the beat positions such as assisted audio editing, DJ mixing applications, audio effect generation, automatic metronome generation and synchronised trigger of different media events such as lights or virtual robots.

There has been a considerable amount of research undertaken in the area of beat tracking. Multi-band approaches have been widely used in the literature; in [1], Scheirer splits the audio signal into 6 frequency bands. Next, a bank of comb filter resonators is used in order to extract the periodicity of the amplitude envelopes of each band. Following this, the resonator outputs are summed across bands in order to estimate the global tempo of the piece of music. Finally, the resonators phase lock is used in order to predict the location of the beats. Klapuri et al. combine the loudness differences of 36 frequency bands [2]. The model uses a comb filter bank to seek periodicities in 3 different metrical levels (tatum, beat and measure) in each of the 4 accent bands. Then, a probabilistic model of the dependencies and temporal relations between the 3 metrical levels is performed. Ellis’s beat tracking model generates an onset detection signal by using a log-magnitude 40 channel

mel-frequency spectrogram [3]. The periodicity detection is performed after summing across channels. Then, a dynamic programming approach is used in order to estimate the beat positions [3]. Alonso et al’s use spectral methods to calculate the periodicity of an onset detector based on a spectral flux representation [4]. Then, as in [1], a pulse train of the estimated tempo is used in order to calculate the position of the first beat. In [5], a method that uses the Viterbi algorithm in order to explain over time the observed tempo hypotheses and its relationships to other metrical levels. In [6], a context dependent beat tracker that uses the complex spectral change onset detector introduced in [7] is presented. The method uses contextual information such as the song’s meter in order to limit the range of beat period hypothesis and future beat location predictions. This method is extended in [8] by using 3 onset detection functions, which uses a confidence measure in order to select the onset detection function that provides the most reliable sequence of onsets.

Other methods estimate the location of the onsets prior to the detection of existing periodicities in the audio signal [9, 10]. In [9], clusters of onsets are used as tempo hypotheses in a multi-agent approach that seeks for the best description of the rhythmic content of the piece of music. This model is modified in [8] by incorporating a complex based onset detector. In [10], Goto uses a multi-agent based model in 3 different rhythmic levels (quarter note, half note and bar note levels) for songs in 4/4. His model uses information of chord changes and onset information in order to select the agent that best represent the observed data. In [8], a detailed comparison between the approaches that participated in MIREX 2006 beat tracking contest is reported, which includes both [2] and [3] approaches, as well as the modifications of [9] and [6] methods.

In this paper, a tempo detector based on the hybrid tempo detection system introduced in [11] is used. The system uses different onset detectors in different frequency ranges in an attempt to exploit the music properties of certain instruments in different frequency regions. The model presented in [11] is modified in order to obtain both a global tempo and short time tempo estimates. The short time tempo estimates are then used in a dynamic programming method based on [3]’s model, in an attempt to better model fast changing tempos. In addition, the content of the frequency bands is analysed in order to choose only bands that provide reliable periodicities. This overcomes one of the limitations of [11]’s model, which uses a fixed frequency decomposition.

The paper is organised as follow; firstly, the proposed beat tracking method is introduced in section 2. Following this, a set of results that validate the presented contributions is shown in section 3. Finally, some conclusions are discussed in section 4.

2. PROPOSED BEAT TRACKING SYSTEM

Figure 1 illustrates the different blocks that comprise the beat tracking system proposed in this paper. Firstly, the band onset and transient detectors ODF_j , where $j \in \{1..3\}$ denotes the band number, are obtained for the whole audio signal. Next, the reliability of the frequency bands is evaluated in an attempt to estimate the bands that provide valuable periodicities. Following this, the $ODFs$ are split into non-overlapping segments of 16sg. Then, the PEriodicity Detection Function ($PeDF$) of each segment i is obtained. Following this, the band $PeDFs$ are combined and weighted in each segment. Next, the segment $PeDFs$ are combined together and both a global tempo and short time tempo estimates are calculated. Finally, the short time tempo estimates are utilised by a dynamic programming approach in order to estimate the beat positions that best describe the observed Onset Detection Function (ODF).

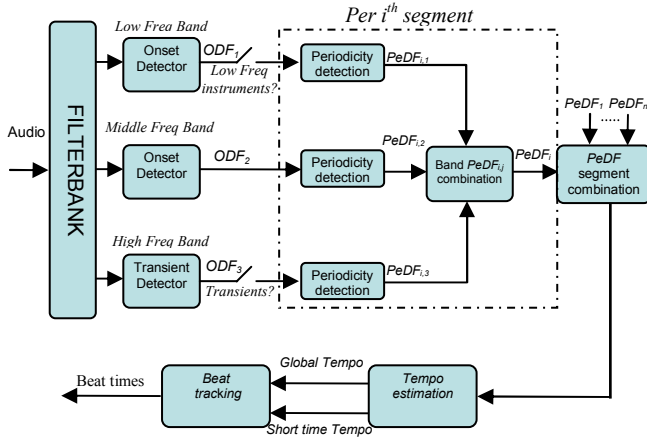


Fig. 1: Hybrid beat tracking approach block diagram

2.1. Global tempo estimation

The different tasks required to calculate the song's global tempo are explained as follows:

Filterbank

The presented beat tracking system splits the audio signal into 3 different frequency bands, which cut-off frequencies are chosen to cover regions of common activity of certain instrument types:

- **Low frequency band (LFB):** frequency range: [0-200Hz]: this band covers the presence of a bass line or percussive instruments such as a snare or a kick drum.
- **Middle Frequency band (MFB):** frequency range: [200-5000Hz]. The chosen band range roughly covers the fundamental frequencies of a wide range of instruments.

- **High Frequency band (HFB):** frequency range: [5000- $fs/2$ Hz], where fs corresponds to the sampling rate. The presence of transients resulting from percussive instruments in the recording will be better localised in this band [11].

Onset/Transient detection functions

Different signal property changes manifest at different frequency bands. Consequently, different onset detection functions which track specific signal properties are used in the above introduced frequency bands [11]. First, a Short Time Fourier Transform (STFT) is calculated using a Hanning window and an FFT length N , which is given by:

$$X(n, k) = \sum_{m=0}^{L-1} x(m + nH)w(m)e^{-j(2\pi/N)km} \quad (1)$$

where $w(m)$ is the window that selects an L length block from the input signal $x(m)$, n is the frame number and H is the hop length in samples.

The spectral change onset detector is used in the LFB and MFB [7]. This method tracks both energy changes in the magnitude spectrum and unexpected deviations in the phase spectrum. These measurements are used to calculate the difference between a measured S_k and a predicted complex number \hat{S}_k . Thus, the complex difference for k^{th} bin of a frame n is given by:

$$\Gamma_k(n) = \left\{ \left[\Re(\hat{S}_k) - \Re(S_k) \right]^2 + \left[\Im(\hat{S}_k) - \Im(S_k) \right]^2 \right\}^{1/2} \quad (2)$$

where \Re and \Im are the real and imaginary parts respectively [7].

The onset detection function frame is then generated in both the HFB and LFB by summing across frequency bin spectral complex changes as follows [7]:

$$ODF_{\{1,2\}}(n) = \sum_{k=1}^{N/2} \Gamma_k(n) \quad (3)$$

In the HFB, the transient detector presented in [12] is used instead. This method tracks the existence of broadband signals in each frame n as follows:

$$T(n, k) = 20 \log_{10} \frac{|X(n, k)|}{|X(n-1, k)|} \quad (4)$$

Then, the onset detection function in band HFB, ODF_3 , is calculated by counting the number of bins that reach a threshold $Thresh$, which is set to 3db [12]:

$$ODF_3(n) = \sum_{k=1}^{N/2} \begin{bmatrix} "1" & \text{if } T(n, k) > Thresh \\ "0" & \text{elsewhere} \end{bmatrix} \quad (5)$$

Finally, the $ODFs$ are smoothed in an attempt to remove less significant peaks. In Figure 2, the onset detection functions of bands MFB and HFB are shown for an excerpt of song "Ain't no sunshine" by Kenny Rogers, where it can be seen that the transient signals are accurately localized in the HFB.

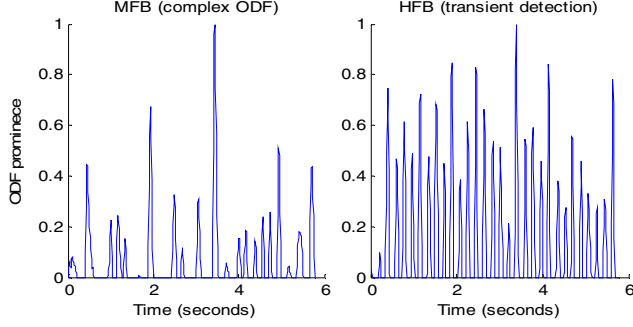


Fig. 2: *ODF* of bands HFB and MFB of an excerpt of song “Ain’t no sunshine” by Kenny Rogers

Dynamic frequency band reliability

The presented system only uses frequency bands which can potentially provide reliable periodicities. Thus, if signal properties that represent certain instrument types in localized frequency regions are not present, the corresponding frequency bands will be disabled. As an example, a song with no presence of low frequency instruments will not produce valuable periodicities in the LFB. In addition, songs with no presence of percussive instruments will not generate transient signals in the HFB and consequently will not provide reliable band periodicities.

• Presence of low frequency instruments in the LFB

The energy content of the audio signal in the frequency ranges [1:200] Hz and [200:500] Hz is respectively obtained, which is denoted as $E1$ and $E2$ respectively. Then, if low frequency instruments are not present in the music recording, it is expected that $E2$ will have a much higher value than $E1$. Thus, by using a conservative threshold $T_{LFB} = 100$, the LFB will not be used if $E2/E1 > T_{LFB}$.

• Presence of transients in the HFB

In order to investigate the presence of percussive instruments in the HFB, transient signals are first detected in this band by using (4) and (5). This is achieved by counting the number of transients, in the ODF_3 that reach a threshold $TranThresh$, which is set to correspond to 30% of the length of the HFB in bins.

$$numTran = \sum_{n=1}^{frames} \begin{cases} "1" & \text{if } TDF(n) > TranThresh \\ "0" & \text{elsewhere} \end{cases} \quad (6)$$

where $TranThresh = 0.3 * N_{HFB}$, and where N_{HFB} and $frames$ correspond to the length of the HFB in bins and the number of frames in the analysis respectively.

It should be noted that transient signals separated by less than 50 ms are assumed to belong to the same transient event. Thus, only the most prominent transient will be kept. Finally, if the number of detected transients $numTran$ is lower than a conservative threshold $minTran$, it will be concluded that percussive instruments were not present in

the HFB. In this case, the HFB will be disabled in the analysis. Threshold $minTran$ is estimated as follows:

$$minTran = \frac{ls}{0.5} \quad (7)$$

where ls is the duration of the audio signal in seconds and 0.5 corresponds to the duration of a beat played at a common 120 bpm tempo.

Segment Periodicity detection

The hybrid multi-band tempo detection model presented in [11] estimates the global tempo of the piece of music. In order to better accommodate the presented beat tracking model to tempo deviations, [11]’s approach is modified. The *ODFs* are split into non-overlapping segments of 16sg. This roughly corresponds to the duration of a common 8 bars section played at a tempo of approximately 120bpm. In each segment, the periodicities of the band *ODFs* in the lag range $D=[minLag..maxLag]$ are tracked, where $minLag$ and $maxLag$ correspond to the beat period (in frames) of a tempo equal to 250 bpm and 40 bpm respectively.

$$r_{i,j}(D) = \sum_{n=1}^{Lo-D} ODF_{i,j}(n)ODF_{i,j}(n+D) \quad (8)$$

where i and j denote the segment and band number respectively, and Lo denote the segment length.

Then, the first D frames of segment i are correlated with the last D frames of the previous segment $i-1$. This ensures the whole range of periodicities are tracked in every frame:

$$r'_{i,j}(D) = \sum_{n=1}^D ODF_{i,j}(n)ODF_{i-1,j}(L_1 - D + n) \quad (9)$$

where L_1 denotes the previous segment length.

Then, the resulting function $r'_{i,j}$ is added to the first D frames of $r_{i,j}$ as follows

$$r_{i,j}(D) = \sum_{n=1}^D r_{i,j}(D) + r'_{i,j}(D) \quad (10)$$

Next, the j^{th} band $PeDF$ of segment i is calculated by using a comb filter type of function within $r_{i,j}$, where more weight to low multiples of D is given:

$$PeDF_{i,j}(D) = \sum_{l=1}^4 \frac{\max_{i,j}(D * l + R)}{l} \dots \text{where } R = \{1..l..l-1\} \quad (11)$$

where $\max(x)$ denotes the maximum value of region x

Then, the i^{th} segment $PeDF$ is obtained by combining the 3 maximised band $PeDF_{i,j}$ as follows:

$$PeDF_i(D) = \sum_{j=1}^3 \frac{PeDF_{i,j}(D)}{\max(PeDF_{i,j})} \quad (12)$$

Next, the i^{th} segment $PeDFs$ are combined into a single $PeDF$:

$$PeDF(D) = \sum_{i=1}^{ns} \frac{PeDF_i(D)}{\max(PeDF_i)} \quad (13)$$

where ns corresponds to the number of segments used in the analysis.

Maximising the *PeDF* ensures that each segment *PeDF* is equally weighted. Thus, loud and quiet passages will contribute equally to the global tempo calculation. The resulting segment *PeDF* is weighted by using the method proposed in [11], which uses a Rayleigh function in order to weight only specific periodicity ranges. The method attempts to avoid weighting periodicities non-multiples of the most prominent periodicity of the *PeDF*.

Finally, the most prominent periodicity τ after applying the weighting is used in order to calculate the global tempo:

$$tempo = \frac{fs \times 60}{\tau \times H} \quad (14)$$

where H and fs corresponds to the analysis hop size and the sampling rate respectively.

As an example, Figure 3 depicts the *PeDF* of song “25 or 6 to 4” by “Chicago”, which global beat period is estimated at $\tau=35$. For a $fs=44100$ Hz, this corresponds to a tempo equal to 147 bpm.

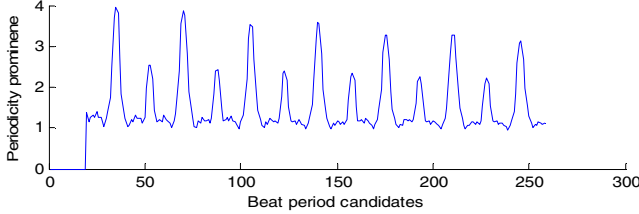


Fig. 3: *PeDF* of song “25 or 6 to 4” by “Chicago”

2.2. Short time tempo estimation

The global beat period τ and the segment *PeDF*s are used in order to obtain the segment tempo t_i . This is performed by extracting the most prominent periodicity τ_i in each *PeDF* _{i} within a region R centered at the global beat duration lag τ . The region width defines the deviation allowed from the global tempo in each segment, which is set to 16% of the estimated τ . Thus, the region R will be defined by the lags within the region $R = \{\tau - 0.16\tau \dots \tau + 0.16\tau\}$.

Finally, the song’s tempo at each individual frame n is obtained by linearly interpolating the tempo of consecutive segments, which will accommodate gradual tempo deviations. Thus, if frames n_1 and n_2 correspond to the centre location of segments i and $i+1$ respectively, the song’s tempo within that range of frames will be given by:

$$te(n = n_1 + k) = te_i + \frac{te_{i+1} - te_i}{L} \times k \quad (13)$$

where $L = n_2 - n_1$, $k = \{0 \dots L-1\}$ and te_i is the tempo of the i^{th} segment.

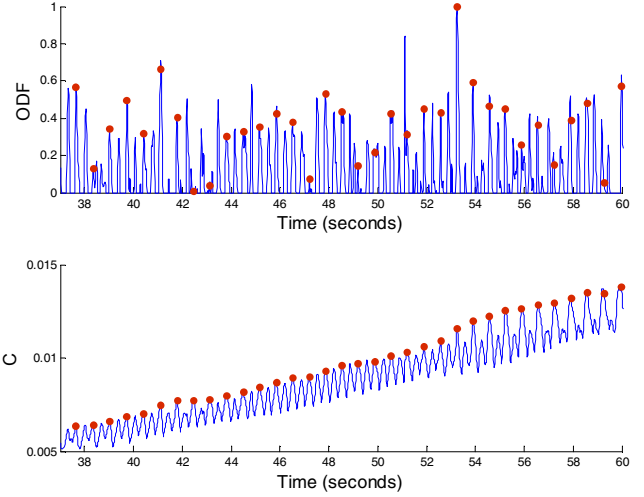


Fig. 4: Example of beat tracking of Handel’s sonata no.3, where the top and bottom plots correspond to the *ODF* and *C* functions respectively. In both plots, the dotted markers correspond to the beat times.

2.3 Beat tracking

The transient detector used in the HFB of the hybrid tempo detector accurately localise transient signals. However, since the method is not dependant on the energy content of the frame, it can generate equally prominent peaks in the *ODF* at beat events as well as at the beat subdivisions (e.g.: half way between beat times at the “ands”). Thus, only the complex based onset detector will be used for beat tracking. Then, dynamic programming techniques are used in order to track a sequence of beats from the complex based *ODF*. In [3], Ellis presents an efficient dynamic programming based beat tracking formulation, which recursively evaluates all possible sequence of beats that end at each frame n :

$$C(n) = (1 - \alpha)ODF(n) + \alpha \max\{F(p)C(p)\} \quad (14)$$

where $p = \{n - 2\tau \dots n - \tau/2\}$ and F is an Guassian function

The parameter α provides a balance between enforcing tempo continuity and allowing a limited amount of tempo deviation. In the presented beat tracker, region p in (13) will dynamically change according to the variable tempo estimation provided by (12). Thus, p in (13) will be given by:

$$p = \{n - 2\tau_n \dots n - \tau_n / 2\} \quad (15)$$

where τ_n is obtained using (11) for $tempo = te(n)$

This modification allows tracking fast tempo changes without affecting the dynamic programming formulation. Next, as in [3], the best predecessor beat is calculated for each frame n :

$$P(n) = \arg \max\{\alpha F(p) + C(p)\} \quad (16)$$

The largest value of C at the end of the time frame, estimated at frame $n=b_N$, will correspond to the last beat of the song. Then, the position of previous beats are obtained

from P by backtracking iteratively [3]. Thus, $b_N - 1 = P(b_N)$, $b_N - 2 = P(b_N - 1)$ and so on. This results in a sequence of extracted beats, which is denoted as $b = \{b_1 \dots b_N\}$.

As an example, Fig.4 shows the result of tracking the beat times within an excerpt of Handel’s sonata no.3. This is a complex example, due to the only presence of pitched instruments and the existence of tempo deviations in the piece of music. The top and bottom plots depict the onset detection function ODF and the dynamic programming function C respectively. In addition, the beat estimates are highlighted in the plots by showing markers at the corresponding beat times. From the top plot, it can be seen that the beat structure of the piece of music is not visually apparent in the ODF . Also, the tempo of the piece of music changes gradually, which is accurately tracked by the dynamic programming formulation.

3. RESULTS

In order to evaluate the proposed beat tracking method, the same manually annotated database and metrics used in [2] are utilised. A summary of the distribution of the 474 songs of the database according to the music genre is shown in Table 1.

Table 1: Details of the database of song excerpts used in [2]

Jazz and Blues	Rock and Pop	Electronic And dance	Hip Hop and Rap	Classical	Soul, RnB and Funk	World and Folk
94	124	66	37	84	54	15

As in [2], the period of an estimated beat is considered as “correct” if it does not deviate more than 17.5% of an annotated beat period. The phase of an estimated beat is considered as correct if it does not deviate from the annotated beats more than 17.5% of the annotated beat period. The following metrics are then used to evaluate the suggested beat tracking system [2]:

- **Correct:** if both period and phase are correctly estimated.
- **Accept d/h :** if the phase is correct and the period is either correct, half or double the annotated beat.

As suggested in [13], results using only the longest continuous segment are also obtained in the evaluation, which is denoted as “*Continuity required*”.

The following state of the art algorithms are included in the comparison:

- **Klapuri’s** model, which results using the same database are published in [2].
- **Davies** submission to MIREX 2006 beat tracking contest [6,8], which code was made available to perform the presented evaluation.
- **Ellis’** beat tracking model [3], which code is publicly available¹.

Both Ellis and the proposed approach use dynamic programming techniques. Thus, parameter α in (13) is set by varying the parameter in the $\{0.5:0.999\}$ range using 10% of the database. The best results were achieved using an α equal to 0.999 and 0.992 for Ellis’ method and the presented approach respectively. The choice of α was found not critical in the presented approach, since the accuracy for every $\alpha > 0.8$ was greater than 80%². However, the accuracy of Ellis’ approach clearly improved as α was increased.

The results of the evaluation are shown in Table 2 for the evaluated systems. The results indicated that the presented approach slightly improves over Klapuri’s method if the “Individual Estimates” metric is used. On the other hand, by using the “Continuity required” metrics Klapuri’s model slightly improves over the proposed method. However, this metric might be seem quite severe, since an incorrect beat estimate at the middle of the song leads to just 50% of accuracy. Thus, a single incorrect estimate will affect the beat detection accuracy differently, which depends on the position of the song in which the incorrect estimate occurs.

The presented approach performs better than both Davies and Ellis methods, which have more problems with beat continuity. Since both Ellis and the presented approach use dynamic programming to track the beats, the accuracy of Ellis’ method might improve if a different onset detector is used.

The results confirm the performance difference reported in [2] between Klapuri’s and other methods, which included Scheirer method [1] and the first version of Dixon’s “beatroot” [9]. However, the modified versions of “beatroot” and Davies [6, 8] perform better than Klapuri in MIREX 2006 beat tracking contest using different metrics [8]. This performance difference might be explained by the average song duration of the music databases utilised, where in [2] and [8] evaluations correspond to approx. 60s and 30s respectively. Thus, the ability of an algorithm to track continuous segments with instrumentation changes becomes more apparent for longer song durations.

Table 2: Results of beat tracking evaluation

Method	Continuity required		Individual Estimates	
	Correct	Accept d/h	Correct	Accept d/h
Proposed Hybrid method	57	71	64	81
Klapuri Causal [2]	57	68	63	78
Klapuri Non Causal [2]	59	73	64	80
Davies [6,8]	39	47	55	70
Ellis [3]	8	30	20	59

4. CONCLUSIONS

In this paper, a method that tracks musical beats by using a hybrid tempo decomposition has been introduced. The method dynamically updates a dynamic programming technique with short time tempo estimates in an attempt to

¹ Available at <http://labrosa.ee.columbia.edu/projects/coversongs/>

² Evaluating individual beat estimates using “*Accept d/h* ” metric

better model gradual tempo deviations. In addition, the presented approach dynamically chooses the frequency bands that provide the most reliable periodicities. This can be important in genres such as traditional Irish music, which can be played without any percussion or low frequency instruments. The method compares favorably against existing state of the art methods, which validates the presented approach. The proposed method clearly improves over existing methods that employ similar frameworks and it slightly improves Klapuri's probabilistic model using the arguably most representative metric.

The use of different strategies in order to dynamically choose the best hybrid multi-band decomposition should be envisaged as future work. In addition, investigating a method to track abrupt tempo changes also warrants future work.

5. ACKNOWLEDGEMENTS

The author would like to express his gratitude to Anssi Klapuri for allowing access to the database of signals used in [2], and also to Matthew Davies for making his beat tracking code available which made possible the comparison tests.

6. REFERENCES

- [1] E. Scheirer, "Tempo and Beat Analysis of Acoustic Musical Signals," *Journal of the ASA*, vol. 103, pp. 588-601, 1998.
- [2] A. Klapuri et al. "Analysis of the meter of acoustic musical signals," *IEEE Trans. on Speech and Audio Processing*, vol. 4(1), 2004.
- [3] D. P. W. Ellis, "Beat Tracking by Dynamic Programming," *Journal of New Music Research, Special Issue on Beat and Tempo Extraction*, vol. 36 pp. 51-60, 2007.
- [4] M. Alonso et al. "Tempo and beat estimation of musical signals," Proc. *5th ISMIR conference*. 2004.
- [5] G. Peeters, "Time variable tempo detection and beat marking," Proc. *International Computer Music Conference, ICMC*, Barcelona, 2005, p. 0.6.
- [6] M. Davies and M. D. Plumbley, "Context-dependent beat tracking of musical audio," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 15(3), pp. 1009-1020, 2007.
- [7] C. Duxbury, J. P. Bello, M. Davies, and M. Sandler, "Complex Domain Onset Detection For Musical Signals," Proc. *6th Int. Conf. on Digital Audio Effects*, London, 2003.
- [8] M. F. McKinney et al. "Evaluation of audio beat tracking and music tempo extraction algorithms," *Journal of New Music Research*, vol. 36, pp. 1-16, 2007.
- [9] S. Dixon, "Automatic Extraction of Tempo and Beat From Expressive Performances," *Journal of New Music Research*, vol. 30, pp. 39-58, 2001.
- [10] M. Goto, "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds," *Journal of New Music Research*, vol. 30, pp. 159-171, 2001.
- [11] M. Gainza and E. Coyle, "Tempo Detection using a hybrid multi-band approach," accepted for publication in *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 2009.
- [12] D. Barry et al. "Drum Source Separation using Percussive Feature Detection and Spectral Modulation," Proc. *Irish Signals and Systems Conference, ISSC*, Dublin, 2005.
- [13] M. Goto, "Issues in Evaluating Beat Tracking Systems," Proc. *IJCAI-97 Workshop on Issues in AI and Music*, 1997.