2013

Continuous Brightness Estimation (CoBE): Implementation and its Possible Applications

G. Presti

Davide Andrea Mauro
Marshall University, maurod@marshall.edu

Recommended Citation

Presti G., Mauro D.A. "Continuous Brightness Estimation (CoBE): Implementation and its possible applications." 10th International Symposium on Computer Music Multidisciplinary Research (CMMR), Marseille (France), 15-18 October, 2013.

Continuous Brightness Estimation (CoBE): Implementation and its Possible Applications

G. Presti,[1] and D.A. Mauro[1,2]

[1] Laboratorio di Informatica Musicale (LIM), Dipartimento di Informatica (DI), Universitá degli Studi di Milano, Milan, Italy
[2] Institut Mines-Télécom, TÉLÉCOM ParisTech, CNRS-LTCI, Paris, France
davide-andrea.mauro@telecom-paristech.fr

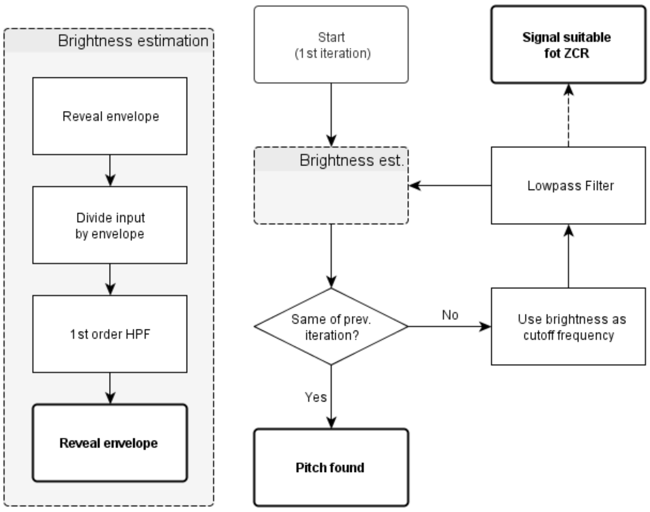
Abstract. This paper presents a time-domain technique based solely on simple operations, like filtering and scaling, to continuously track a sound descriptor correlated to brightness and pitch. In opposition to existing algorithms this approach does not rely on input framing or frequency-domain transforms, ensuring a better temporal resolution and the possibility to model analog-like implementations. In the first part of the document we present the details of our approach to brightness estimation; then we will compare CoBE to the brightness estimation implemented in MIR Toolbox; we introduce and define the concept of "Equivalent Brightness Frequency" (EBF) and finally we show how to exploit CoBE and EBF to obtain a rough monophonic pitch estimator or an efficient pre-processor for zero crossing rate (ZCR) pitch trackers.

Keywords: brightness estimation, real-time, time domain approach, feature extraction, preprocessing, pitch-tracking

1 Introduction

In the context of Music Information Retrieval (MIR) the use of complex descriptors such as "brightness" as well as the fundamental frequency estimation of signals are well-researched problems that can be solved in a number of different ways for a number of different constraints (e.g. mono/poly-phonic signal, offline or in realtime). No algorithm can solve the issues for all the requirements and with any possible input condition so the design of dedicated algorithms is still an open issue.

When one of the requirements is the real-time execution of the algorithm it become very important to avoid any unnecessary computation. See [1] for a review of other algorithm for realtime pitch tracking and refer to a classical paper on time-domain pitch extraction by Gold and Rabiner [2]. A possible solution is to perform only operation in time domain. This is the case of the proposed solution that exploits the behaviour of filters. A workflow of the proposed algorithms is shown in Fig. 1. The adoption of such a methodology turns out to provide both a brightness estimator and a low cost pitch tracker for monophonic sound.



(a) CoBE

(b) Iterative use of CoBE as a pitch tracking helper.

Fig. 1. A flow chart of the proposed techniques.

2 Continuous Brightness Estimation (CoBE)

The main idea behind CoBE is to remove from the input data all the amplitude modulations dividing the signal by its envelope, then the demodulated signal is filtered and its new envelope is revealed. This envelope is a measure of how much of the input sound matches the filter characteristic (examples in Fig. 2 and Fig. 3). If the filter is an high-pass filter, this provides an estimation of brightness, which can be defined as the level of the signal over a certain cut-off frequency (see [3] and [4]).

In opposition to other algorithms we use a very gentle 1st order filter to ensure some sensitivity to lower frequencies (see Fig. 4). This characteristic makes the algorithm slightly different from others brightness trackers, but gives us the ability to exploit this behaviour to also track pitch (details in Sections 4 and 5). Moreover, brightness is generally measured with 0 to 1 real values, while CoBE can give values higher than 1, but this happens only for pure tones near Nyquist frequency, which is a negligible circumstance in case of music signals.

CoBE works in spite of amplitude modulations over the input signal, because those modulations are removed in the first step of the algorithm. Nevertheless, as in many brightness tracking algorithms, the response to high frequencies in conjunction with the insensitivity to level changes, implies a sensitivity to background noise, which can be treated with the implementation of an input gate which stops the processing for sounds under a certain threshold (this function has not been implemented in the present work).

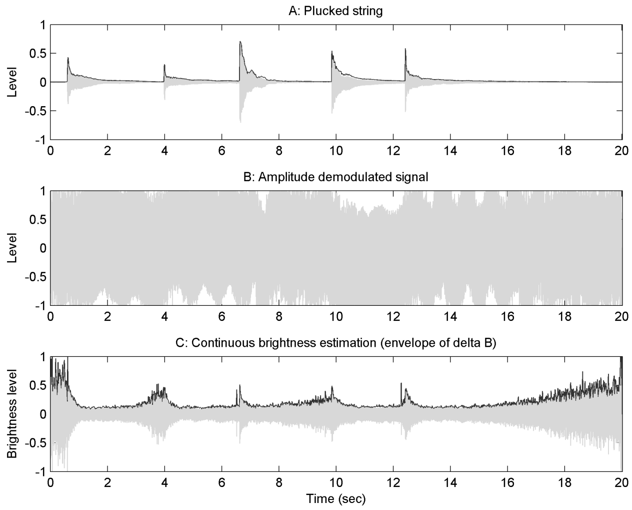


Fig. 2. A: Input signal. We recorded a plucked instrument with some background noise coming from the amplifier; B: Input divided by its amplitude envelope; C: Filtered version of B, showing the percentage of the input matching the filter characteristic. Brightness increases when the notes fades out due to the presence of noise.

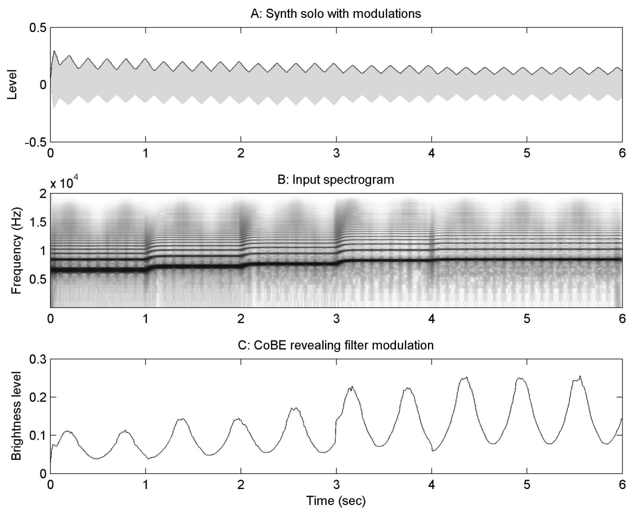


Fig. 3. Due to the HPF, high frequencies have a great influence over brightness estimation, this can help revealing lowpass filter modulations despite of amplitude and pitch changes.

Finally, the method used to track the envelopes inside CoBE may be of any kind. For example a "peak meter" like algorithm with a short decay time can provide good results and can be easily implemented.

3 CoBE vs. MIR Toolbox Brightness

Typically, wide band or high pitched signals result in peaks of brightness, while low pitched and/or narrow bandwidth signals result in falls of brightness. Classical

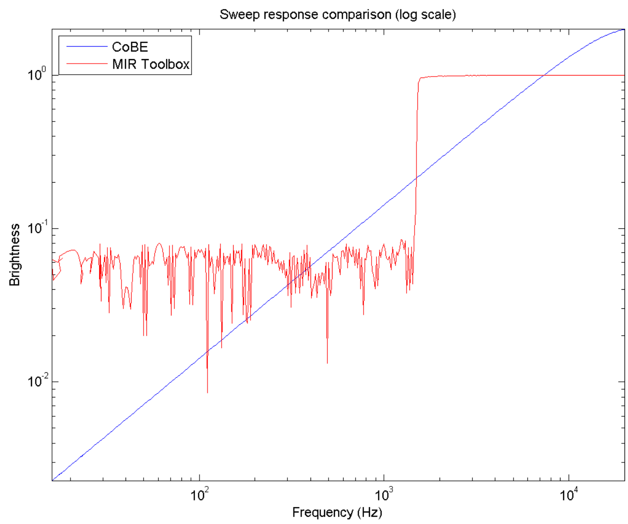


Fig. 4. Sweep response for CoBE and MIR Toolbox. The main difference between classical brightness and our definition of brightness lies in the filter used, which ensures a continuous increasing sensibility over the spectrum instead of a steep distinction between bright and non-bright frequencies.

brightness implementations have an abrupt transition between this two cases (see Fig. 4), due to the fact that the portion of the signal which is not considered "bright" is cut out with a very steep curve. To obtain a smoother behaviour, instead of using high-order filters, CoBE relies on a simple differential filter (1st order Hi-Pass, or derivative filter). As shown in Fig. 5 and Fig. 6, this alters the range of the readings, but without changing the general shape of data. In particular, Fig. 6 shows that a non linear function seems to fit the data: further studies will investigate the compatibility between CoBE and other different approaches to brightness estimation.

On the computational side, a Matlab test implementation run on various musical instruments and songs samples revealed that CoBE, on average, takes 8

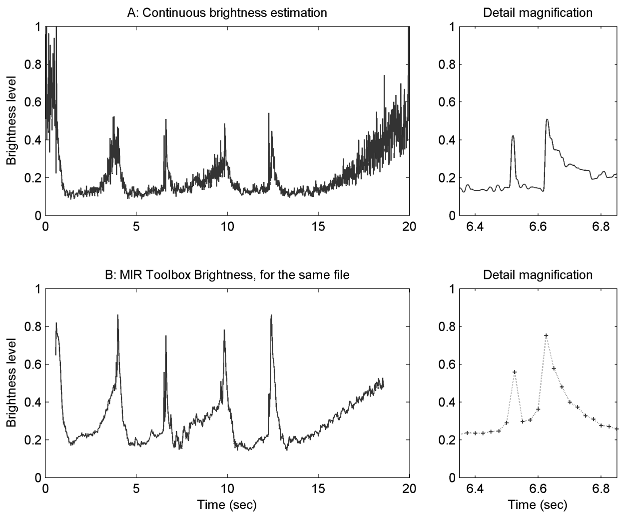


Fig. 5. The same sound of Fig. 2 analyzed with CoBE and with MIR Toolbox Brightness: As shown in the right side of the plot CoBE can obtain finer temporal precision because it does not rely on signal framing.

ms to process one second of audio at CD quality, while “MIR Toolbox brightness” takes about 27 ms. Test results are shown in Fig. 6. The test has been done on a commercial laptop with 6GB of RAM and an Intel i5-3317U @ 1.7GHz processor.

4 Equivalent Brightness Frequency (EBF)

As in MIRToolbox, we can interpret CoBE brightness as the percentage of signal made of “bright” frequencies, but in case of sinusoidal inputs CoBE brightness is strictly related to the pitch of the input, and so it can acquire another meaning. This because the delta (i.e. differential) of a sine wave consists in quite the same signal, delayed and downscaled proportionally to its frequency, by a factor which is simply the transfer function of the filter used. In other words: CoBE brightness value can be used to calculate the frequency of a sinusoidal input. Looking at Fig. 7: “A” is an amplitude modulated gliding sinusoidal wave; “B” is the demodulated and filtered signal: the transfer function of the differential filter used in CoBE is:

$$|F(f)| = 2\sin(\pi\frac{f}{f_s}) \quad (1)$$

“C” is the envelope of “B” processed with the inverse transfer function of the filter:

$$f = (\frac{f_s}{\pi})\arcsin(\frac{A}{2}) \quad (2)$$

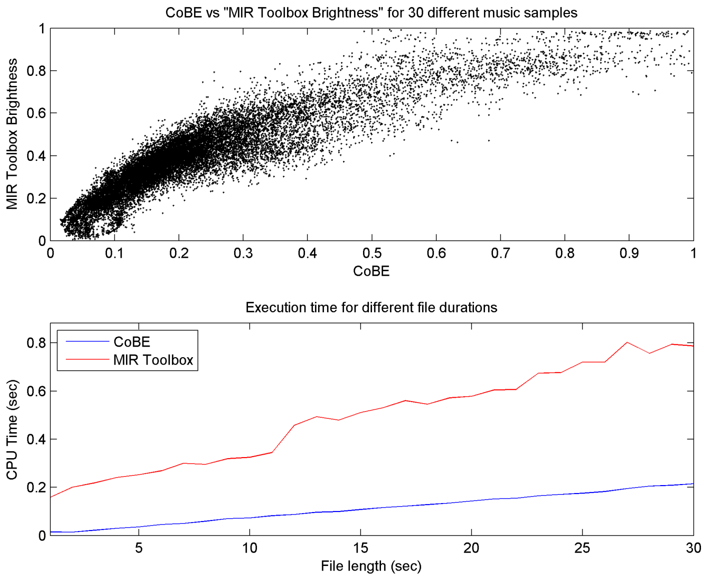


Fig. 6. (upper) Plotting the results of 30 different music samples, analyzed with CoBE and with MIR Toolbox Brightness. (lower) Execution time for different samples length. The computational cost for CoBE seems to be of a lower order of complexity.

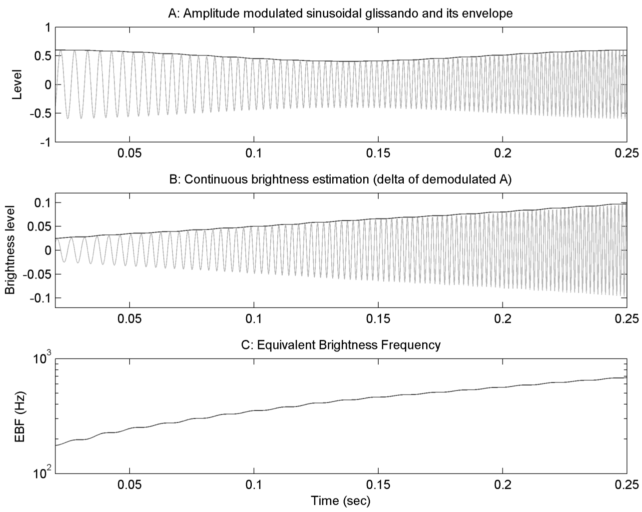


Fig. 7. The brightness value "B" of the sinusoidal input "A" can be turned into the frequency value "C" because of the known transfer function used in "B" and the independency of CoBE from the amplitude modulations.

Now that we saw that CoBE brightness can be a simple function of the pitch of sinusoidal inputs, we can define "Equivalent Brightness Frequency" (EBF) as the frequency that a sinusoidal signal must have to match the brightness of a given sound.

5 Multipass (or iterative) CoBE

In case of non sinusoidal inputs, EBF generally points far away over the fundamental frequency. If we use EBF as cut-off frequency for a 1st order lowpass filter and then feed the filtered signal again into CoBE, the new EBF will be lower than before. Repeating this operation several times will make EBF converge in the neighbourhood of the lowest frequency of the input (generally the fundamental), this because the higher frequencies will extinct faster, leaving the lowest alone as a pure tone: further filtering will not change the balance of the signal content (but only the amplitude, which is irrelevant to CoBE), resulting in a steady brightness and a steady EBF. The procedure is presented in Fig. 1 and a sample shown in Fig. 8. Running more iterations or choosing abrupt low-pass filters can cause estimation overshooting errors, especially in presence of subharmonics. This iterative setup of CoBE behaves like a rough pitch tracker

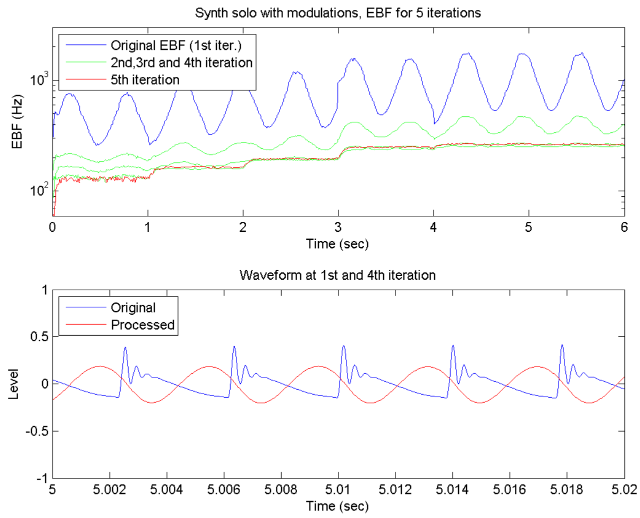


Fig. 8. (upper) EBF over 5 iterations of Multipass CoBE converging to the lowest frequency of the signal (the sample is the same shown in Fig. 3). (lower) Input waveform and waveform after 4 iterations of Multipass CoBE. The processed signal is suitable for ZCR pitch tracking algorithms.

for monophonic signals, but a more detailed study needs to be done to investigate its limitations. If we compare the waveform before the first iteration and after the 4th one (Fig. 8), we can see that the processed signal resembles a sinusoidal wave more than the input, but sharing with it the same fundamental frequency. This behaviour is very useful to prepare the signal for other pitch tracking algorithms, like ZCR.

6 Conclusion

We presented a fast, continuous, and analog-like algorithm to track brightness that can also be exploited to approximate pitch of monophonic audio signals and we briefly compared it with MIR Toolbox Brightness. This tool can be useful in all those situations where a fast and continuous brightness estimation is needed, like in some MIR applications (e.g. see [5] and [6],) or as a control signal for audio effects.

Acknowledgments. This work has been partially funded by the Enhanced Music Interactive Platform for Internet User (EMIPIU) project and by the European Commission under contract "FP7-287723 REVERIE".

References

1. P. De La Cuadra, A. Master, and C. Sapp. Efficient pitch detection techniques for interactive music. In Proceedings of the 2001 International Computer Music Conference, pages 403-406, 2001.
2. B. Gold and L. Rabiner. Parallel processing techniques for estimating pitch periods of speech in the time domain. The Journal of the Acoustical Society of America, 46(2B):442-448, 1969.
3. O. Lartillot, P. Toiviainen, and T. Eerola. A MATLAB toolbox for music information retrieval. Data analysis, machine learning and applications, pages 261-268, 2008.
4. E. Schubert, J. Wolfe, and A. Tarnopolsky. Spectral centroid and timbre in complex, multiple instrumental textures. In Proceedings of the international conference on music perception and cognition, North Western University, Illinois, pages 112-116, 2004.
5. P. Laukka, P. Juslin, and R. Bresin. A dimensional approach to vocal expression of emotion. Cognition & Emotion, 19(5):633-653, 2005.
6. P. Juslin. Cue utilization in communication of emotion in music performance: Relating performance to perception. Journal of Experimental Psychology: Human perception and performance, 26(6):1797, 2000.