

5th Aubio Study Results

Study on MKL **only**

Domenico Stefani

24 February 2021

Abstract

This brief report presents the results of a comparison study on the MKL onset detection method of the Aubio suite. This derives from the intention on improving the onset detection performance on a specific application and acknowledging that the Adaptive whitening technique developed in [1] proved to improve the performance of several OD methods, excepts for MKL, which showed consistently better performances on the regular, non whitened version (even better than the 2 version of the other methods).

1 Details

Refer to the previous study for the introduction and more general details. For this version, the **MKL** methods was analyzed in different settings:

1. Default MKL settings in aubio (with adaptive whitening), unprocessed audio signal.
2. Default MKL settings in aubio (with adaptive whitening), noise gated signal (threshold -50dB, ratio 1:20).
3. MKL **without adaptive whitening**, unprocessed audio signal.
4. MKL **without adaptive whitening**, noise gated signal (threshold -50dB, ratio 1:20).
5. MKL **without adaptive whitening**, highpassed signal (Cutoff: 2000Hz, Mix 0.2 (20% original signal, 80% filtered)).

2 Results

The best f1-score results along with the latency metrics connected to them are presented in tables 1 and 2.

Table 1: The best f1-score avg. values are shown. Different combinations of Buffer size and Method produce different latency values, which are reported in the following tables. Bold values represent the points in the Pareto front defined by the points in the space of 2 objectives: the f1-score (to maximize) and the Inter Quartile Range (to minimize). More info in fig. 1 and table 3.

		Buffer size					
		64	128	256	512	1024	2048
Method	mkl	0.8482	0.8522	0.8718	0.8661	0.8690	0.8706
	mkl (whitening, noisegate)	0.8815	0.8819	0.8831	0.8771	0.8715	/
	mkl (No whitening)	0.9511	0.9702	0.9764	0.9731	0.9732	0.9613
	mkl (No whitening, noisegate)	0.9376	0.9525	0.9580	0.9484	0.9557	0.9324
	mkl (No whitening, highpass)	0.9483	0.9638	0.9658	0.9569	0.9491	0.9421

Table 2: The results of the latency recorded on the examples which f1-score is reported in table 1 are shown here. Each cell contains 3 values: the first and the last are the lower and upper Tukey fences with $k = 1.5$, which are defined starting from the Interquartile range and are commonly used to define outliers of a distribution, while the central value is the sample mean of the latency distribution.

Method		Buffer size					
		64	128	256	512	1024	2048
	mkl	2.3/4.6/6.7	2.9/5.3/7.5	3.8/6.4/8.8	5.4/8.2/11.0	8.5/11.0/13.6	11.4/14.5/18.1
	mkl (W, NG)	2.0/4.0/6.0	2.7/5.0/7.2	3.1/5.2/7.3	4.1/6.1/8.3	5.5/9.3/13.4	/
	mkl (No W)	2.8/4.6/6.3	3.2/5.1/6.8	4.2/6.0/7.7	5.3/7.4/9.2	8.3/10.9/13.3	10.5/13.7/16.5
	mkl (No W, NG)	2.7/4.5/6.3	3.2/5.0/6.8	4.2/5.9/7.6	5.3/7.3/9.2	7.8/10.5/13.5	10.8/13.4/16.0
	mkl (No W, HP)	2.4/4.5/6.4	2.9/4.9/6.9	3.9/5.9/7.8	5.2/7.3/9.2	8.8/11.1/13.5	11.7/14.1/16.6

Pareto front results are shown in table 3 and fig. 1 for the first analysis, and in table 4 and fig. 2 for the second one.

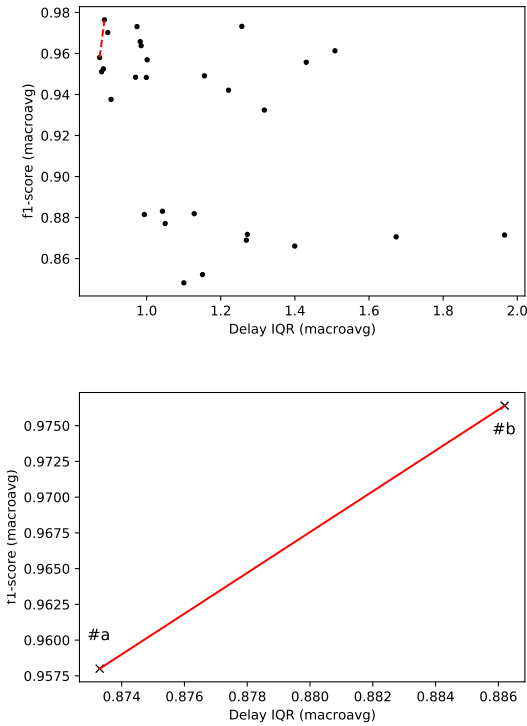


Figure 1: Pareto front computed for f1-score and the Interquartile Range of the latency distribution. The upper plot shows all the solution while the lower plot represents only the points in the front. The labels refer to the detailed information that can be found in table 3.

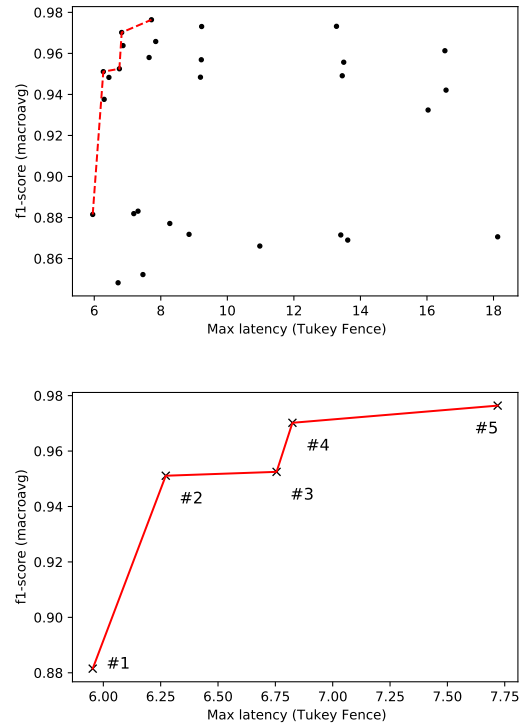


Figure 2: Pareto front computed for f1-score and upper Tukey fence. The upper plot shows all the solution while the lower plot represents only the points in the front. The labels refer to the detailed information that can be found in table 4.

3 Results on dynamics

References

- [1] Dan Stowell and Mark Plumbley. Adaptive whitening for improved real-time audio onset detection. In *Proceedings of the 2007 International Computer Music Conference, ICMC 2007*, pages 312–319, 2007.

Table 3: Pareto front solution with *f1*-score (macro average over all techniques) as the first objective and Interquartile Range of latency as the second.

#	Method	F1-score	Low Tukey fence (ms)	Delay mean (ms)	High Tukey fence (ms)	Onsets inside fences (%)
a	mkl (No whitening, noisegate)	0.9580	4.1551	5.9223	7.6484	95.25
b	mkl (No whitening)	0.9764	4.1744	6.0215	7.7192	95.70

Table 4: Pareto front solution with *f1*-score (macro average over all techniques) as the first objective and maximum latency as the second, in the form of upper Tukey fence.

#	Method	F1-score	Low Tukey fence (ms)	Delay mean (ms)	High Tukey fence (ms)	Onsets inside fences (%)
1	mkl (whitening, noisegate)	0.8815	1.9788	4.0052	5.9537	93.92
2	mkl (No whitening)	0.9511	2.7591	4.5746	6.2722	96.90
3	mkl (No whitening, noisegate)	0.9525	3.2225	5.0274	6.7553	95.86
4	mkl (No whitening)	0.9702	3.2442	5.0959	6.8244	96.74
5	mkl (No whitening)	0.9764	4.1744	6.0215	7.7192	95.70

Table 5: Results on dynamics

ID	Piano Accuracy	Piano Precision	Piano Recall	Mezzoforte Accuracy	Mezzoforte Precision	Mezzoforte Recall	Forte Accuracy	Forte Precision	Forte Recall	Piano F1	Mezzoforte F1	Forte F1
a	0.8699	0.9719	0.8923	0.8559	0.9484	0.8978	0.9635	0.9677	0.9956	0.9304	0.9224	0.9815
b	0.9341	0.9693	0.9625	0.8678	0.9554	0.9044	0.9868	0.9890	0.9978	0.9659	0.9292	0.9934