2nd Aubio Study Results

With Noise Gate and Dynamic Range Compressor

Domenico Stefani

27 January 2021

Abstract

This brief report presents the results of a comparison study on all the methods of the Aubio suite for onset detection. Some choices are strongly affected by the target application for the onset detector.

This is a variant of a previous study, with the difference of having a Noise gate and compressor added

1 Details

Refer to the previous study for the introduction and more general details. The only difference is that now files are processed with a Noise Gate and a compressor. The specifics are:

- Noise Gate:
  - Threshold: *-36dB*
  - Ratio: *9:1*
  - Attack: *0 ms*
  - Release: *0 ms*
- Compressor:
  - Threshold: *-20dB*
  - Ratio: *6:1*
  - Attack: *0 ms*
  - Release: *0 ms*

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	hfc	0.9201	0.9316	**0.9383**	**0.9417**	0.9358	0.9305
	energy	0.9063	0.9135	0.9231	0.9227	0.9333	0.7875
	complex	0.9217	0.9288	0.9298	0.9313	0.9314	0.9225
	phase	0.8739	**0.8955**	0.9071	0.9207	0.9255	0.9150
	specdiff	0.9016	0.9156	**0.9302**	**0.9355**	0.9257	0.9314
	kl	0.9034	0.9149	0.9258	0.9372	0.9268	0.9353
	mkl	0.9143	**0.9293**	**0.9351**	0.9392	0.9329	0.9332
	specflux	0.8927	**0.9040**	**0.9065**	**0.9250**	0.9225	0.9234

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.

		Buffer size					
		64	128	256	512	1024	2048
Method	hfc	1.7/4.1/6.2	2.0/4.2/6.1	2.6/4.8/6.8	3.6/6.0/8.1	4.6/7.2/9.4	4.9/9.9/14.9
	energy	1.0/3.4/5.5	1.4/3.6/5.6	3.4/5.6/7.5	5.1/7.4/9.5	8.5/10.8/13.1	16.1/18.2/20.8
	complex	1.8/4.0/5.9	2.0/4.2/6.0	2.3/4.5/6.5	2.7/4.8/6.8	3.3/5.7/8.0	3.5/5.9/8.1
	phase	0.4/2.6/4.4	2.0/4.0/5.6	2.3/4.3/6.1	2.8/5.0/6.9	3.0/5.2/7.1	3.4/6.1/8.4
	specdiff	2.0/4.0/5.8	2.6/4.6/6.4	3.4/5.5/7.4	4.4/6.6/8.5	5.9/8.3/10.5	6.4/12.2/18.0
	kl	0.5/2.9/5.0	0.7/3.2/5.4	1.7/3.9/6.0	3.4/6.1/8.7	4.2/6.7/9.2	7.9/10.8/14.1
	mkl	1.2/3.5/5.5	1.9/4.0/5.8	2.3/4.5/6.4	3.5/5.9/8.2	4.6/7.3/9.9	5.6/9.3/12.8
	specflux	0.6/2.6/4.5	1.6/3.5/5.2	1.8/3.8/5.5	2.4/4.4/6.2	2.5/4.5/6.3	3.5/5.9/8.0

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.

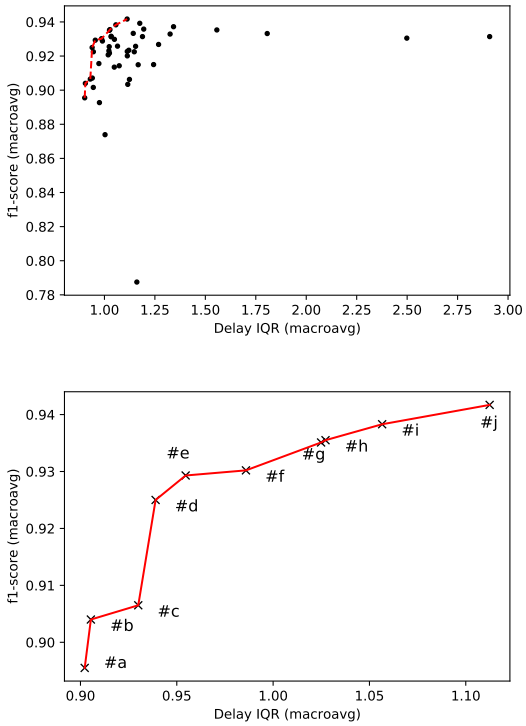

Figure 1: Pareto front computed for f1-score and the Interquartile Range of the latency distribution. The upper plot shows all the solution while the lower plot represents only the points in the front. The labels refer to the detailed information that can be found in table 3.

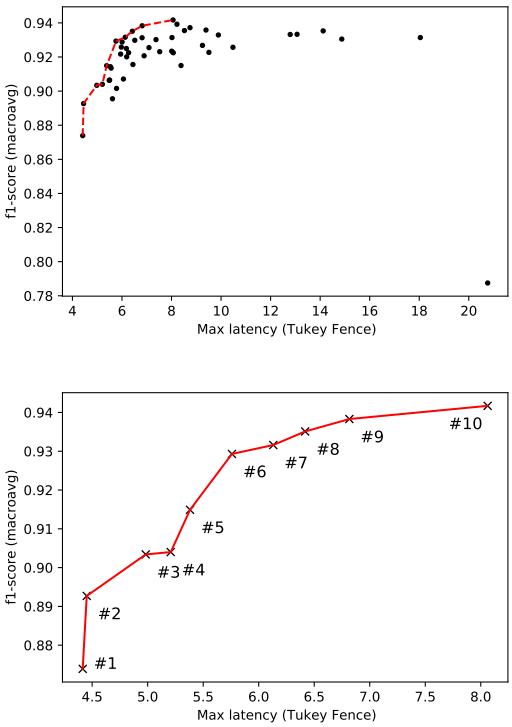

Figure 2: Pareto front computed for f1-score and upper Tukey fence. The upper plot shows all the solution while the lower plot represents only the points in the front. The labels refer to the detailed information that can be found in table 4.

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	phase	0.8955	2.0074	3.9506	5.6163	94.63
b	specflux	0.9040	1.5863	3.5320	5.2078	94.64
c	specflux	0.9065	1.7809	3.7514	5.5008	95.12
d	specflux	0.9250	2.4208	4.4328	6.1767	94.29
e	mkl	0.9293	1.9406	3.9823	5.7591	94.31
f	specdiff	0.9302	3.4297	5.5351	7.3733	94.31
g	mkl	0.9351	2.3188	4.4937	6.4183	93.28
h	specdiff	0.9355	4.4135	6.5849	8.5224	94.65
i	hfc	0.9383	2.5903	4.8357	6.8165	95.01
j	hfc	0.9417	3.6125	5.9758	8.0620	95.13

3 Results on dynamics

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	phase	0.8739	0.4029	2.6010	4.4150	94.68
2	specflux	0.8927	0.5505	2.6176	4.4513	96.66
3	kl	0.9034	0.5202	2.9324	4.9827	94.86
4	specflux	0.9040	1.5863	3.5320	5.2078	94.64
5	kl	0.9149	0.7126	3.2222	5.3815	94.73
6	mkl	0.9293	1.9406	3.9823	5.7591	94.31
7	hfc	0.9316	1.9984	4.2158	6.1313	94.70
8	mkl	0.9351	2.3188	4.4937	6.4183	93.28
9	hfc	0.9383	2.5903	4.8357	6.8165	95.01
10	hfc	0.9417	3.6125	5.9758	8.0620	95.13

Table 5: Results on dynamics

ID	P Accuracy	P Precision	P Recall	MF Accuracy	MF Precision	MF Recall	F Accuracy	F Precision	F Recall	Piano F1	Mezzoforte F1	Forte F1
a	0.7309	0.8846	0.8080	0.7966	0.9337	0.8444	0.8154	0.9054	0.8914	0.8446	0.8868	0.8983
b	0.7038	0.8233	0.8290	0.7737	0.8949	0.8511	0.8625	0.9345	0.9180	0.8261	0.8725	0.9262
c	0.7260	0.8513	0.8314	0.7822	0.9218	0.8378	0.8457	0.9479	0.8869	0.8412	0.8778	0.9164
d	0.7722	0.8862	0.8571	0.8100	0.9305	0.8622	0.8601	0.9227	0.9268	0.8714	0.8950	0.9247
e	0.7563	0.8802	0.8431	0.8330	0.9448	0.8756	0.8910	0.9608	0.9246	0.8613	0.9089	0.9424
f	0.7966	0.8931	0.8806	0.8210	0.9172	0.8867	0.8905	0.9289	0.9557	0.8868	0.9017	0.9421
g	0.7676	0.8955	0.8431	0.8358	0.9538	0.8711	0.9050	0.9722	0.9290	0.8685	0.9106	0.9501
h	0.8197	0.9074	0.8946	0.8442	0.9413	0.8911	0.9011	0.9469	0.9490	0.9010	0.9155	0.9479
i	0.7759	0.8886	0.8595	0.8260	0.9359	0.8756	0.9227	0.9663	0.9534	0.8738	0.9047	0.9598
j	0.8129	0.9087	0.8852	0.8205	0.9313	0.8733	0.9417	0.9732	0.9667	0.8968	0.9014	0.9699