Tony: a tool for melody transcription - Bug #897
display doesn't extend to highest extracted frequencies

2014-03-07 11:23 AM - Matthias Mauch

Status:	New		**Start date:**	2014-03-07
Priority:	Normal		**Due date:**	
Assignee:			**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:				

Description

I understand that Tony adapts the maximum frequency shown to the data, but that seems to be limited to around, what, 1500Hz?

I attach a screenshot where pyin has extracted a high frequency at the end of a file, but Tony doesn't display it. It does play (re-synthesise) it.

This can be reproduced using the flute file (also attached).

History

#1 - 2014-03-07 02:11 PM - Chris Cannam

The spectrogram uses SV's standard MelodicRangeSpectrogram config, which goes up to (yes) 1500Hz.

We should adjust this based on the extracted features, but with caution -- earlier versions of Tony clamped the displayed range quite harshly to the extracted feature range, which is much worse. Equally, we don't really want to display up to 2KHz when the entire melody is in the 200-500 range.

#2 - 2014-03-07 02:24 PM - Matthias Mauch

Some considerations:

  - I notice we never really need the bottom two octaves, so if we want to risk that the pitch track overlaps the waveform, then those could just go *a priori*
    - setting the frequency range to pyin's original estimate range +/- one octave would be easy to do (for Chris)
      - is there any better way of doing it?
      - once the pitch track is changed the range could be adapted
  - a problem emerges when we need to change the frequency by more than one octave (or even the candidates are outside +/- one octave
    - one could expand to the range covering all candidates
    - one could expand to the range covering the currently selected candidate
    - one could risk blindly selecting some candidates and update the frequency range after it has been confirmed

#3 - 2014-03-10 04:01 PM - Rachel Bittner

How difficult would it be to allow the user to adjust the displayed frequency range? Maybe as a menu option in "View > Change Frequency Range..." --> which could take user entered frequencies or a set of predefined choices (for ex. min freq = 0 Hz, 100 Hz, 200 Hz, ... 800 Hz, max freq = 300 Hz, 400 Hz, ... 1500 Hz)

In theory then, Tony could choose the initial range as it does now or using one of the ways Matthias proposed, and then the user can manually adjust as it suits them. I don't think this is the ideal solution, but it would be flexible and remove the corner cases where there is no way to see a frequency out of range.

#4 - 2014-03-11 09:18 AM - Chris Cannam

Re. Matthias's point, the melodic-range spectrogram in SV (which is what this essentially is[*]) has a range from 40-1500Hz, although the low end rounds down to 37Hz. But for some reason Tony's Analyser code was setting the low end explicitly to 15Hz (which again would round down). Simply removing that line improves things there, I think, though there is one little problem -- if a recording has 50Hz hum, that now overlaps with the waveform.

Will look into Rachel's request.

[*] I still haven't got around to finishing the constant-Q integration because it turned out to need a significant amount of work in v-axis alignment for arbitrary plugin-generated layers -- that would be really useful work to do regardless of Tony, but I haven't quite found the time to do it yet

#5 - 2014-03-11 09:44 AM - Chris Cannam

I've added (in commit:ae0bf8dc5833) a View -> Set Displayed Frequency Range function.

I hope this works well enough to be going on with for now, but it's far from ideal. It simply adjusts the bin range of the spectrogram layer, which (whether visible or not) is what the other layers are aligned to. But this means the values are always rounded to spectrogram bins, and in particular the way the rounding works (or fails to work) means that the lower bound will be rounded down even if you leave it unchanged!

This really should set hard boundaries for the display range, but that sort of job should go along with something like the constant-Q display alignment, on a day when I have more time to make sure it's actually correct. This way the results should be right, they're just fiddly to set to what you want.

Also, just for further annoyance, the extents are not remembered -- you have to re-enter them each time you load a file.

All the same I hope this is of some help.

Downloads

Screen Shot 2014-03-07 at 11.07.25.png	50.8 KB	2014-03-07	Matthias Mauch
MusicDelta_LatinJazz_Premix_04.wav	11.5 MB	2014-03-07	Matthias Mauch