

# **AudioFeX v1.8**

## **(MUVIS Xt Audio Feature Extraction Tool)**

### **Quick Reference**

by Serkan K.



The *AudioFeX v1.8* is a standalone tool, capable of playing (rendering) and performing automatic segmentation and classification of the audio files (clips). In this version, compressed/uncompressed audio files (AMR, MP3, AAC, PCM/ADPCM) in several file formats (AMR, MP3, AAC, AVI, MP4) are supported. After performing the analysis, it displays the final classification and segmentation result as well as classification and segmentation result based on individual feature (Pause Rate, Fundamental Frequency and Subband Centroid Features). It is also capable of displaying advanced feature information that is related to the analysed file. This application basically serves a test-bed application to develop, test and make performance comparisons of the automated audio classification and segmentation algorithms that will be used in audio indexing and retrieval framework within MUVIS.

### **1 Installation/Uninstall**

Simply run “Setup AudioFeX v1.8.exe” file. Once installed, you can run the same file to uninstall the program.

**Q: Which operating systems *AudioFeX* can be installed on?**

A: Like the other MUVIS applications, it is developed and tested for the following OSs:

- Windows 9x
- Windows NT
- Windows ME
- Windows 2000

- Windows XP

**Q: Are there any pre-requisite programs that should be installed before *AudioFex* installation?**

A: No.

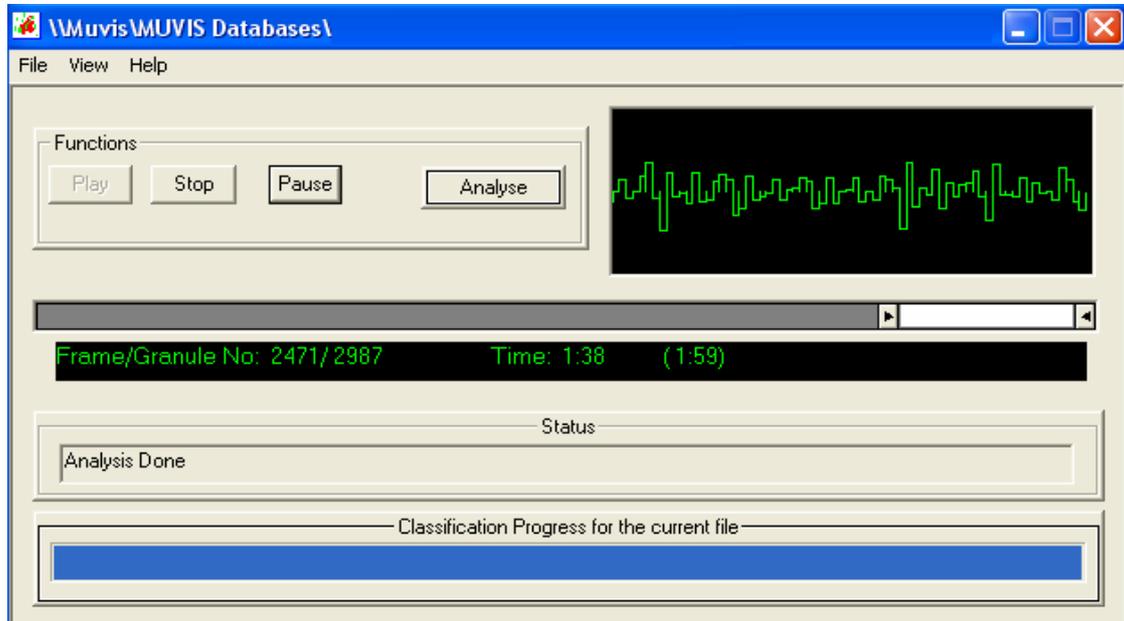


Figure 1. AudioFex 1.8 GUI

## 2 *AudioFex* in Run Time

When you run *AudioFex*, the UI appears on the screen with no current action. User should initiate the actions.

**Q: Which types of files are supported by *AudioFex*?**

A: In *AudioFex* version 1.8, both compressed audio files (MP3, AAC, AVI or MP4) that contain MP3/AAC audio and other audio types (AMR, PCM/ADPCM) are supported. *AudioFex* may support more file formats in future versions.

**Q: What are the actions that *AudioFex* can perform?**

A: In basic terms, *AudioFex* can render (play, stop or pause) any and analyse a supported files. It is capable of performing the classification and segmentation of such files, based on the analysis of these files i.e. performing “Audio Features Extraction”. You can load any supported file simply by drag&drop onto the main UI or you can use the <File> menu to load it. Once loaded, it can be played or analysed. Also pressing “<Play>” or “<Analysis>” button on the main UI of the *AudioFex*, may let you load and play or analyse a specific file at the beginning of the run time.

**Q: What happens when I load a supported file?**

A: Once loaded, the file starts playing.

**Q: What actions are possible once I have started playing a supported file?**

A: Once you have played a file, it can be paused or stopped by pressing the <Pause> and <Stop> buttons respectively on the UI. Seeking forward/backward arbitrarily and limiting audio time can also be done by using the seek bar from both ends.

**Q: How can I analyse a supported file?**

A: There are 2 methods of doing this. When no file is loaded in the *AudioFex*, pressing <Play> button will let you directly load&play a file. Similarly, pressing <Analyse> button will let you directly load&analyse a file. Once a file is loaded these button will always play or analyse, the loaded file. Second method is “drag&drop” method. A file can be directly loaded&played (and may be analysed after loading) by this method.

**Q: How can I change analysis parameters?**

A: Once a file is loaded, from the <File> menu, press <Parameters> to get Audio Parameters dialog. In this dialog you can change any parameter you wish to.

**Q: What does it mean by “Generic Analysis” and “Bit-stream Analysis”?**

A: In the Audio Parameters dialog, by pressing the button labelled <Generic Analysis> will take you to counterpart of it, i.e., the <Bitstream Analysis>. Pressing it again will bring you back to <Generic Analysis>. Bitstream Analysis can only be done for compressed files while Generic Analysis common for both compressed and uncompressed files. Bitstream Analysis is made directly on the compressed audio bitstream, while Generic Analysis is done in the (decoded) uncompressed domain. If the file is in a compressed format, Generic Analysis will uncompress the bitstream and then commence the analysis.

**Q: Do I need to optimize the parameters?**

A: *AudioFex* runs on the default “Generic Analysis” audio parameters unless the parameters are changed. They should be changed manually for “Bitstream Analysis”. The parameters that needed to be changed for “Bitstream Analysis” are given below. Once changed manually, the parameters should be set again to default parameters manually. Although, *AudioFex* provides the luxury of changing the parameters, but for best results it is recommended to run the application with default parameters as these are already optimized. For “Bitstream Analysis”, it is recommended to change the following parameters:

- 1) Energy Threshold % = 12
- 2) FF No. of Harmonics = 8
- 3) FF Peak2Mean Ratio = 2.5

Both Generic and Bitstream Analysis parameters are shown in Figure 7.

**Q: Once the analysis is finished, what exactly does the popup window show?**

A: Once audio analysis is complete for a specific file, the final result of classification and segmentation is shown in the “*Classification and Segmentation Window*” with tracing of a line.

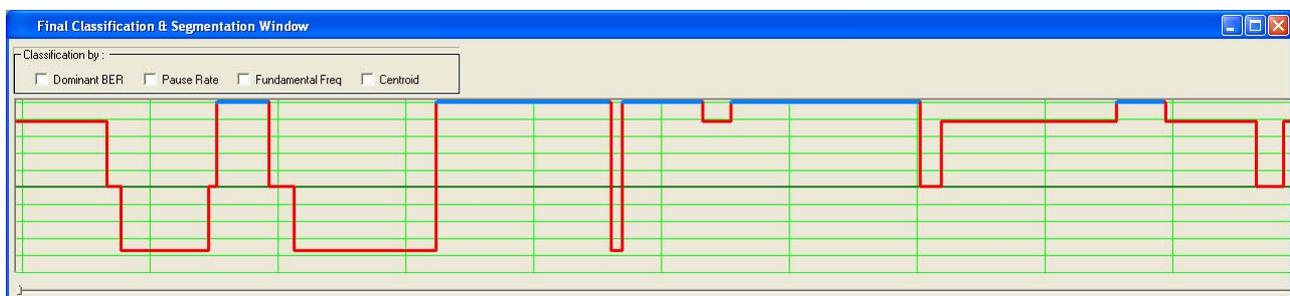


Figure 2. Classification and Segmentation window

**Q: In the “Classification and Segmentation Window”, what is the legend to understand the meaning of the traced lines?**

A: The legend is such that if the red-line lies above the centre position, that specific segment is classified as *speech*, if it is below the centre position; it shows that the segment is classified as

*music*. All segments with blue trace lines are classified into *fuzzy* class type. All the segments that have their display line lying on the centre line are classified as *silent* segments. If a segment is shown with black line means that it has a forced class type (details ahead).

**Q: In the classification and segmentation window, why are some lines drawn with “blue” colour?**

A: Segments that are above the centre line and are displayed in a special “blue colour” are classified as fuzzy segments (mixed classes within this segment).

**Q: On the classification and segmentation window, what are the checkboxes for?**

A: Besides the final classification and segmentation result that you get after each successful analysis, on the same window, you can also see the classification and segmentation results based on individual features like Dominant Band Energy Ratio (blue line), Fundamental Frequency (green line), Subband Centroid (*yellow line*)\_and Pause Rate (modified version) (*white line*) by checking the respective checkboxes from left to right. Multiple checkboxes can be selected at a time.

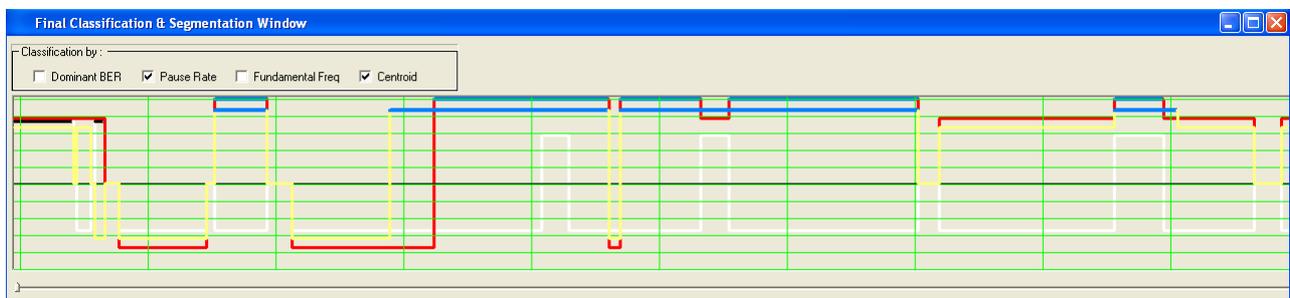


Figure 3. Displaying segment feature classification and segmentation

**Q: In the classification and segmentation window (for classification based on a specific feature), why are some lines drawn with “black colour”?**

A: Segments whose trace lines are above the centre line and are displayed in a “black color”, have a **forced-speech** classification with respect to that specific feature. Similarly, segments with black color lines below the centre line have a **forced-music** classification. An example is shown in Figure 4 below.

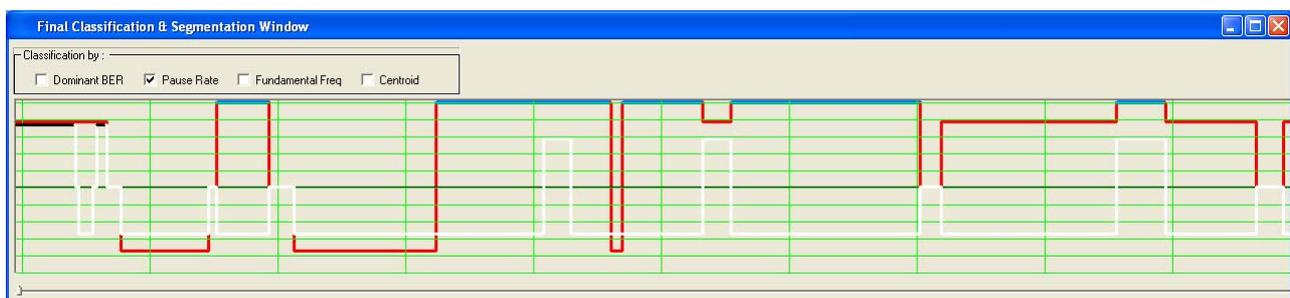


Figure 4. Displaying forced classification

**Q: Do I have to re-analyse a file if I accidentally close the final classification and segmentation window and want to see it again?**

A: No. You can re-display the classification and segmentation window by going to the <View> menu and press the button for the specific classification and segmentation window you want to view.

**Q: Can I get any advance feature information related to an analysed file?**

A: Yes. Once an audio analysis is performed, feature information can be displayed in separate windows. In the <View> menu on the main UI of *AudioFex*, selecting <Fundamental Frequency feature> will display the FF feature for all the segments in the “Fundamental Frequency feature

dialog”. Similarly, selecting the <Subband Centroid feature> will show the SC feature of the segments of the file in Subband Centroid feature dialog, while pressing <Pause Rate feature> will display the (modified) PR feature value for each segment within the analysed file in the “Pause Rate feature dialog”.

**Q: What does the number represent in the black box, placed in the upper left corner of each feature window?**

A: This is the maximum value that the feature has attained within any segment.

**Q: What does the button <Showing Feature> in the feature-windows mean?**

A: It means that the window currently shows the result based on a segment feature. Pressing this button, it will show the feature curve, which will convert the window into a more detailed feature-window. Pressing this button on the Fundamental Frequency feature-window will display the detected Fundamental Frequency values of each granule. Pressing this on Subband Centroid window will display the SC sliding-windowed standard deviation curve, while pressing this on Pause Rate window will display the segmentation information (breakpoints applied) for that file. Figure 5 gives an example of this case. Pressing it again will show the feature again.

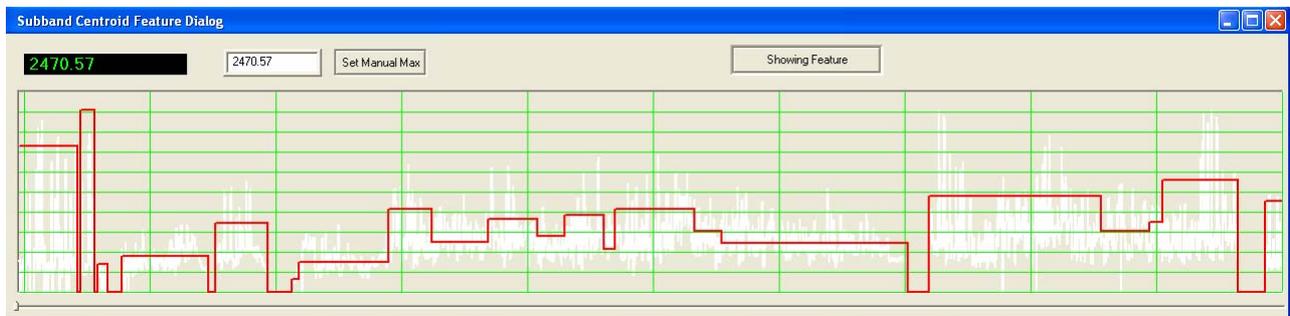


Figure 5. Showing segment feature and curve

**Q: Why are some lines short and some long, in the segmentation (breakpoints) window?**

A: For *AudioFex*, segmentation is based on either silence occurrence or by intra-segmentation (segmentation of sub-segments within a segment that contain no silent part in between). The first segmentation based on silence is shown by short lines, while the intra-segmentation breakpoints are shown by the longer lines. An example of such segmentation window is Figure 6.

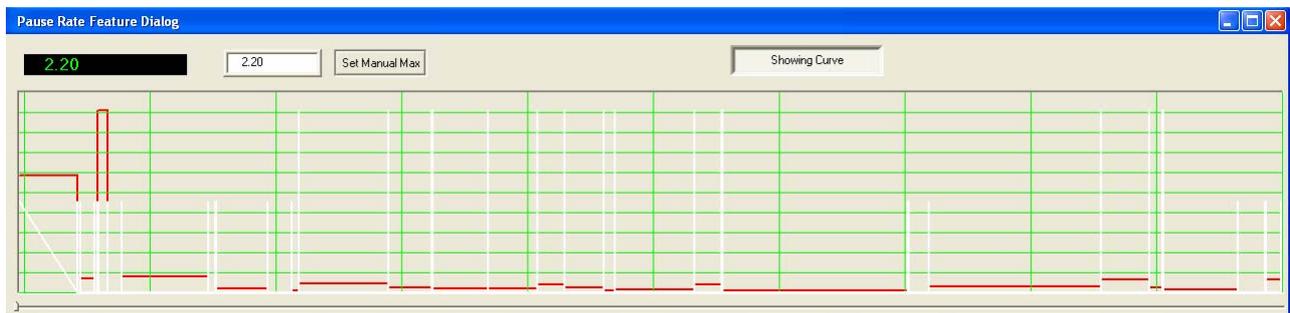


Figure 6. Displaying different breakpoints applied

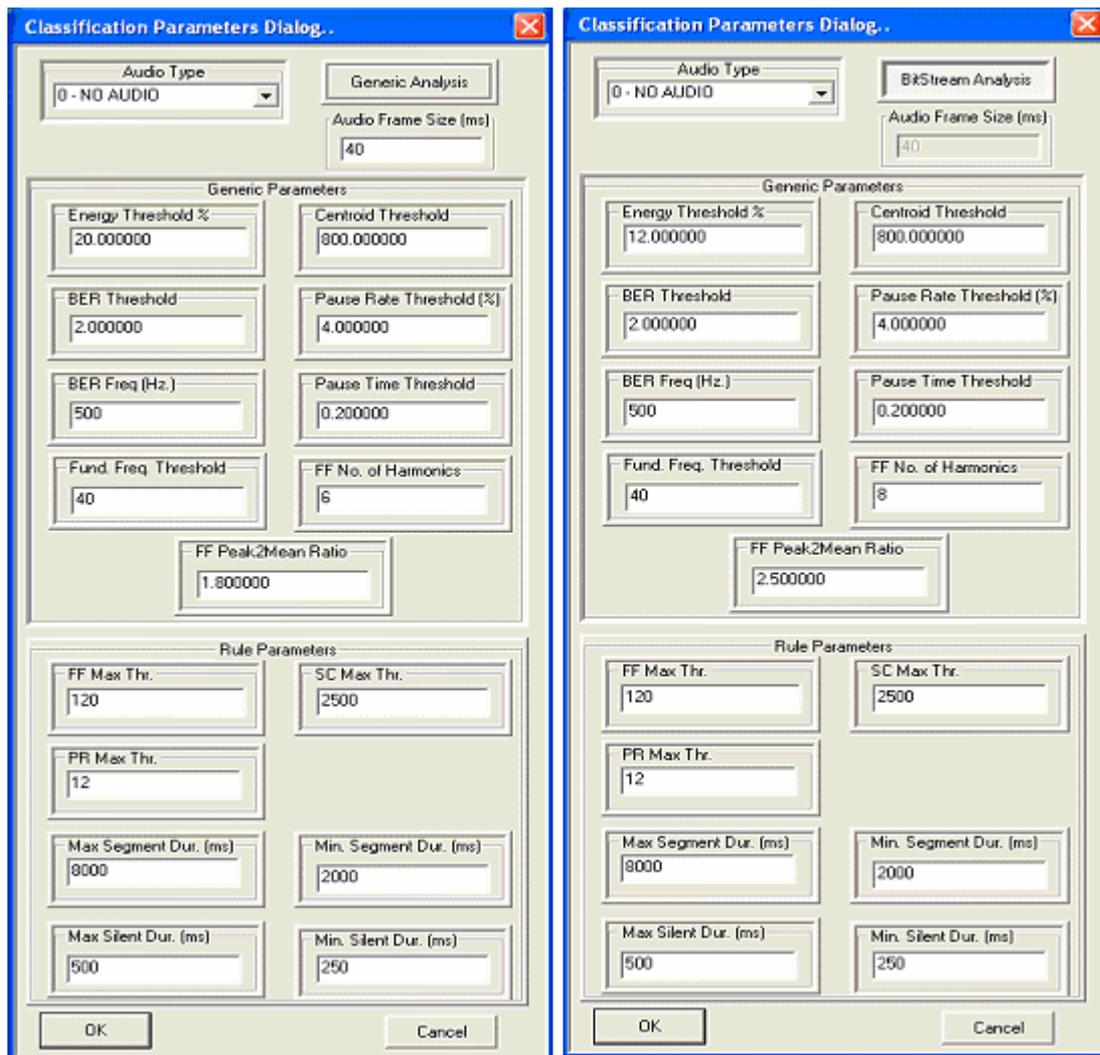


Figure 7. Generic Analysis (default) and Bitstream Analysis (manual) parameters