FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Universitätsklinikum Erlangen

Glottis Analysis Tools 2020

- User guide -

IMPORTANT NOTE: You cannot have two versions of GAT installed at the same time on one computer.

We have reworked core aspects of the code at the foundation of GAT for the new version. This means:

1. Segmentations created with the previous version can be used in the new version
2. Please **do not mix parameter computations** from previous versions with the new version

While we provide backward compatibility towards GAT 2018 segmentation files from earlier version can not longer be opened with GAT 2020. Furthermore we would strongly advise to avoid mixing the computations of parameters between GAT 2018 and GAT 2020 to ensure consistency. Hence, we recommend to wait with updating your current version until your studies are finished or to recompute the parameters with the new version.

Division for Phoniatrics and Pedaudiology at the ENT Department,
University Hospital Erlangen
Waldstr. 1
91054 Erlangen, Germany

Contents

1 Changelog

This version of GAT features several improvements. The biggest changes are the implementation of the new and fully automatized glottis segmentation as well as various performance improvements. Furthermore we added a function in the segmentation tool to allow for adjustment of the glottal area of single frames and some more features to improve usability in segmentation, analysis and audio analysis tool. Multiple performance improvements are also included, which should in particularly speed up the segmentation process. A detailed list of changes is given below.

1.1 Added or changed Features

- Speed up segmentation process by about 300%
- Changed Shimmer Percent measure for a more stable version (defined analogously to Jitter Percent)
- Added TensorFlow for deep neuronal networks
- Added three different automatized glottis segmentation models
- Added two buttons in segmentation for an user friendly first step in the segmentation process
- Added automatic region of interest prediction
- Added rotation option in region of interest selection form
- Added brightness and contrast regulators in region of interest selection form
- Added Histogram Equalization option in region of interest selection form
- Added sliders for frame selection in region of interest selection form
- Video container is now FFMPEG based
- Set default setting for audio import to "entire"
- Added "Edit Segmentation" option for manual correction of the glottal area in single video frames
- Added caption to "Change Workspace" and "Head to Analysis" Buttons to make them more prominent
- Added two cycle switching buttons in Analysis for switching between audio and other cycles
- Improved cycle detection for GAWs with long closed phases
- Changed depiction of single cycles (now last frame shown of cycle x is first frame shown for cycle x+1)
- Shifted "clear all" button in audio analysis to make it more prominent
- Added "select all" buttons in audio analysis for files and parameter selection
- Added new popup dialogs informing the user
- Added new more user friendly licensing procedure

1.2 Fixed bugs

- Fixed lagging and rare crashes in the segmentation tool
- Fixed performance issues in the segmentation tool
- Fixed several display bugs

2 Glottis Analysis Tools

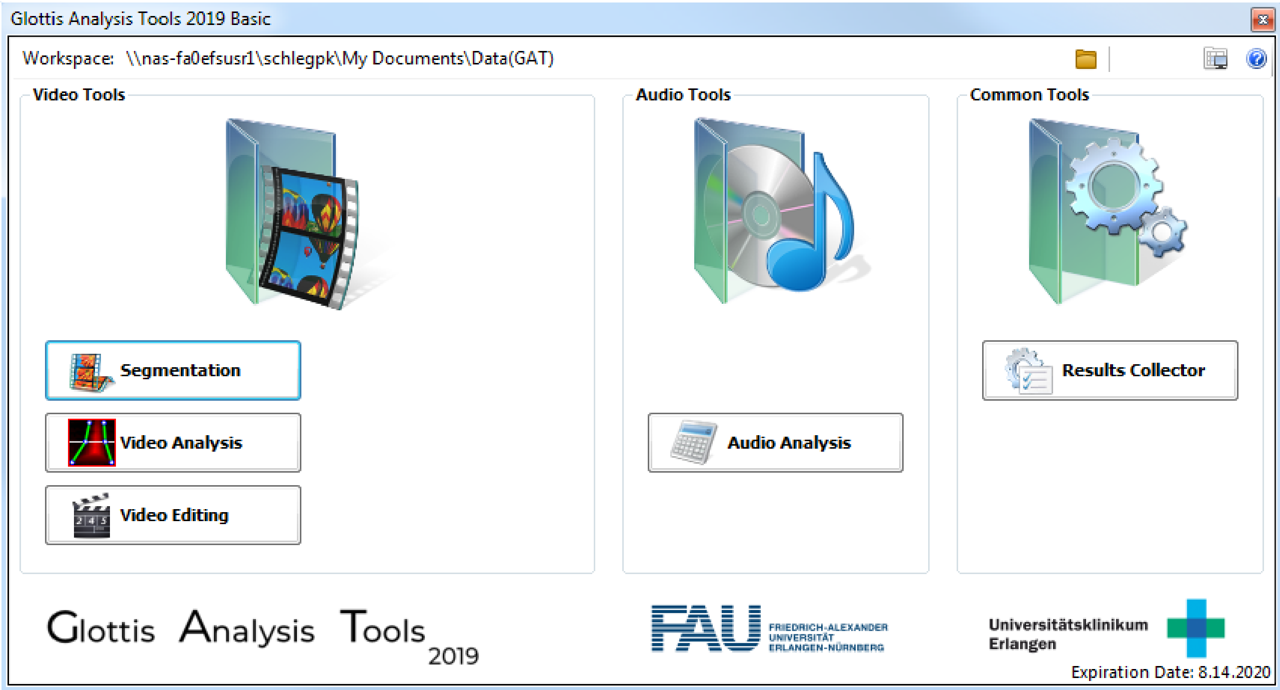


Figure 1 *– Glottis Analysis Tools: Main menu*

The Software *Glottis Analysis Tools* allows the segmentation of high-speed endoscopy (HSE) recordings and analysis of both audio and video recordings. This program provides the following features (fig. 1):

- **Video Processing** (sec. 4): performs the processing of HSE recordings, in particular, the segmentation of HSE recordings to create a representation of some characteristics of the vocal folds' vibratory motions like a phonovibrogram (PVG) or a glottal area waveform (GAW). Furthermore it allows the analysis of these characteristics as well as the analysis of an acoustic signal, i.e. audio recording, corresponding to the HSE recording.
- **Audio Processing** (sec. 5): performs the processing and analysis of an acoustic signals without an accompanying HSE recording.
- **CommonTools** (sec. 6): can be used to perform statistical evaluations on sets of analysis data

2.1 Installation / Licensing

Installation To install the software run the *Glottis Analysis Tools Setup.exe* and follow the instructions.
Note:

- During the installation and at the first start of *Glottis Analyse Tools (GAT)* no internet connection is required.
- On Windows XP/Vista/7/8 some problems with unauthorized access can occur when running the *Glottis Analysis Tools* after the installation. To avoid this, **do not** install the program into the folder *Program Files*!
  Predefined installation folder: *System drive:\Program Files\Glottis Analysis Tools*.
- ***GAT* can only be installed using admin rights.**

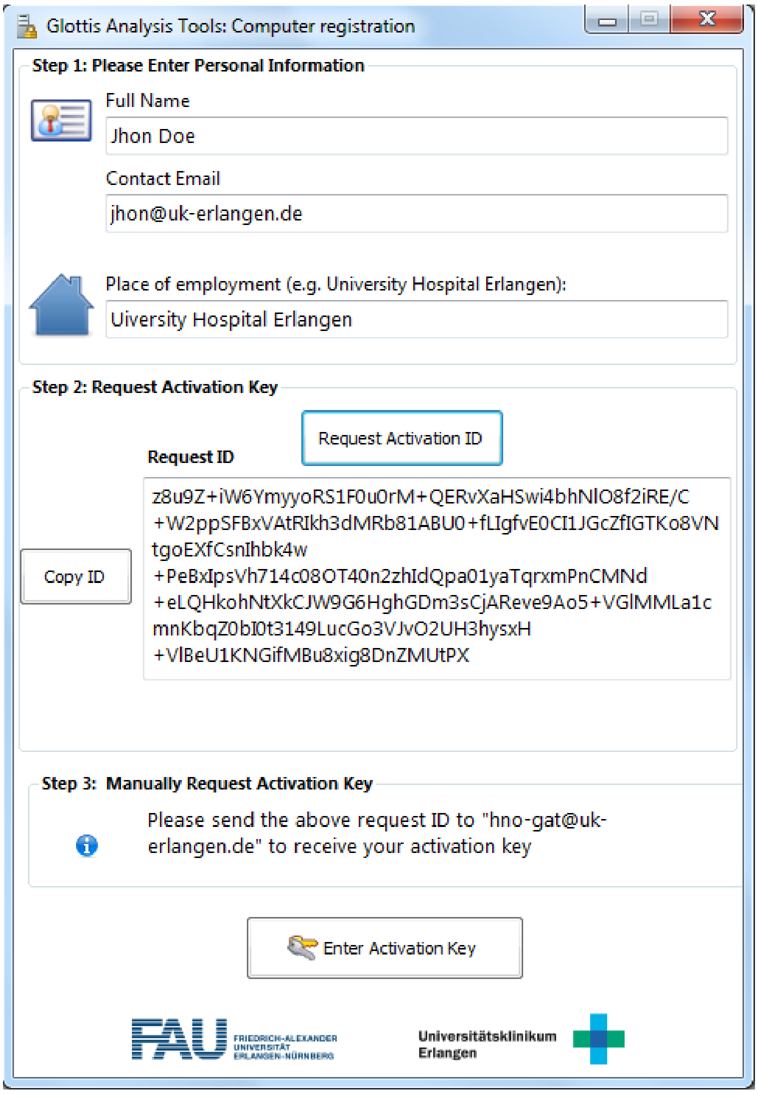


Figure 2 *– Glottis Analysis Tools: Computer registration*

Licensing License validation allows running *Glottis Analysis Tools* without an internet connection at any time. Your license information is written in the *User Info File* (*.UIF) on your computer. This is the license file obtained by registering the software.
Computer Registration: If your computer is not yet registered or the license file (*.UIF) is deleted or expired, an error message *"ACCESS DENIED!"* will be displayed. In this case, the process to register the computer will be started after closing the error message (fig. 2).
The computer registration consists of the following few steps:

1. Entering your Name, E-mail and Workplace (fig. 2).
2. Click "request activation ID" to generate the "Request ID".
3. The software will send the "Request ID" automatically to the admin if you have internet connection.
4. **note:** The Request ID will only be shown if automatically sending it failed.
5. If sending the request failed, Copy the ID and send it to the email below:
   - email: hno-gat@uk-erlangen.de
6. After that you will receive an activation key via email. Click "Enter activation key" to open a window where you can enter your key and activate *GAT*.

If your license expires in less than 14 days, you will be informed. For license renewal please contact to above mentioned email.

2.2 General

***Figure 3** – Glottis Analysis Tools: Info form*

Each included tool contains the button *"Help"* ( ) located in top right corner of the window. Clicking it opens the following menu:

- **Info** (fig.3): provides information about the software as well as the computer-license.
- **User guide**: opens the software documentation.

Furthermore each included tool contains different context menus, which can be opened by clicking the right mouse button.

2.2.1 Workspace

To store the generated data, e.g. analysis results or segmentation settings, you should define the **workspace** (Default: *MyDocuments**Data(GlotAnTools)*) folder. All output data of all projects will be saved there. Click on the button *"Set workspace..."* (( ) which is located on center top edge of most menus) to change it.
Note: You should only change the main folder! All required subfolders will be automatically created as illustrated in *"Architecture"* (fig. 4).
The paths for the *Video source folder* and *Audio source folder* (fig. 4) are the default folders used to find Audio and Video files for analysis . Click *"Ok"* to apply the changes or *"Cancel"* to abort and close the current window.

2.2.2 Analysis: Cycle Detection

The *Glottis Analysis Tools* uses the following settings for the detection of cycles (fig 5).

C1. **Video sampling rate** (Hz): allows the change of the sampling rate if necessary.

C2. **Detected cycles import**: in order to prospectively import detected cycles, they should be exported after detection (sec. 4.2).

- **Fundamental frequency (f_0)**

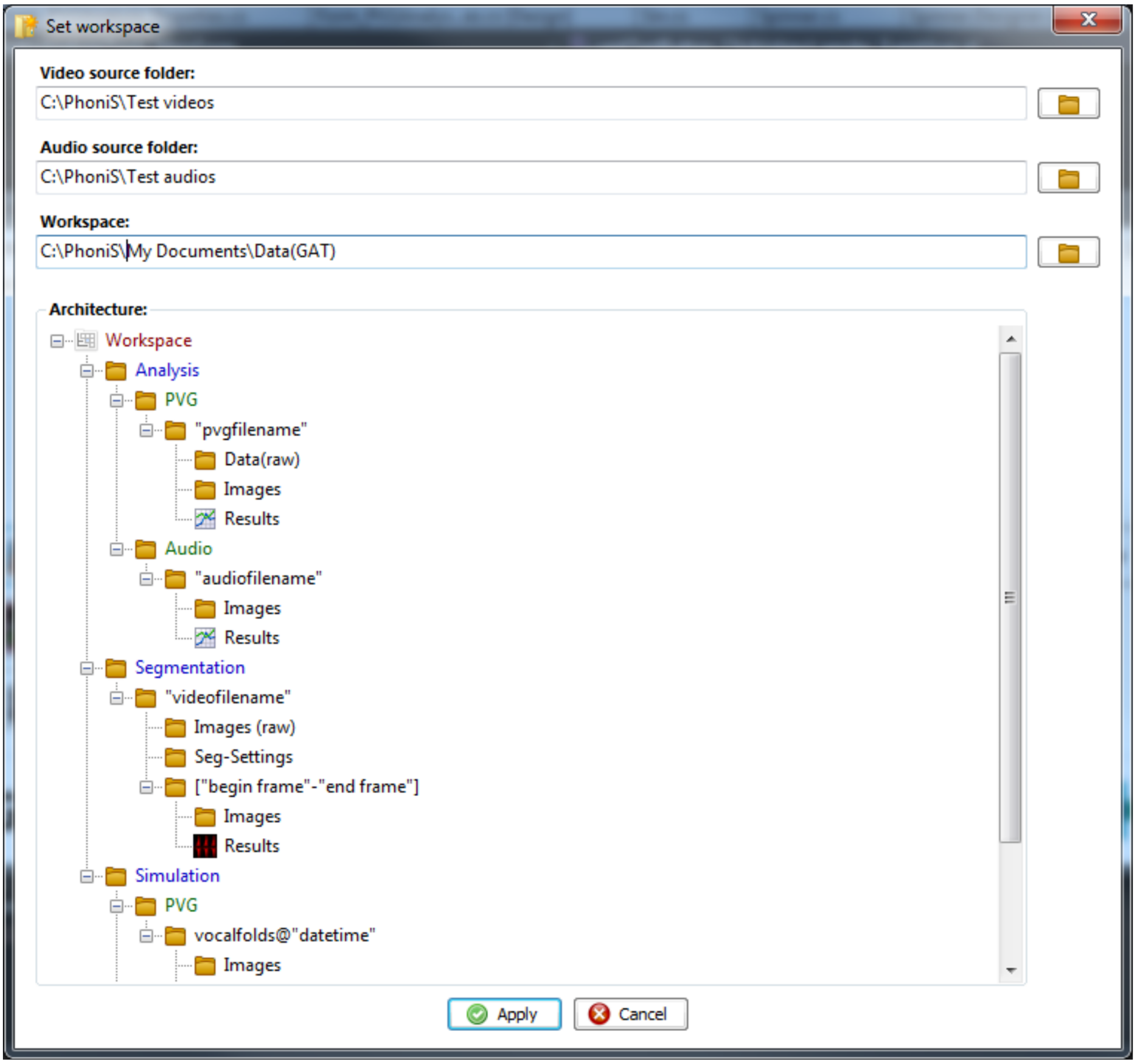


***Figure 4** – Glottis Analysis Tools: Change workspace form*

C3. **Tolerance range (Hz)**: defines an allowed range for the fundamental frequency. For a correct detection of cycles it is important that the actual fundamental frequency of your signal is within this range. You can adjust the tolerance range to fit the approximate fundamental frequency of your signal better (e.g. choose higher values, if you have very short cycles or choose lower values, if you have broader cycles) to ensure a fitting cycle detection.

C4. **Magnitude tolerance limit (%)**: defines a threshold for the Fourier coefficients representing the frequencies, which are in the allowed range of the fundamental frequency (f_0). This threshold is defined as a percentage of the maximum occurring Fourier coefficient magnitude.

C5. **Magnitude tolerance limit of harmonics (%)**: all the approximately integer multiples of the fundamental frequency will be accepted as harmonics, if their Fourier coefficients are not below this tolerance limit. This threshold is defined as a percentage of the magnitude of the Fourier coefficient representing the fundamental frequency f_0.

- **Cycles**

C6. **Detection method**: specifies the algorithm that will be used for the determination of the fundamental frequency f_0 and thereby cycles calculation (sec. 3.1.1)

C7. **Correlation method**: specifices the method used for the accuracy estimation in the fundamental frequency determination and cycles calculation (sec. 3.1.1). The frequency with maximum accuracy will be accepted as f_0.

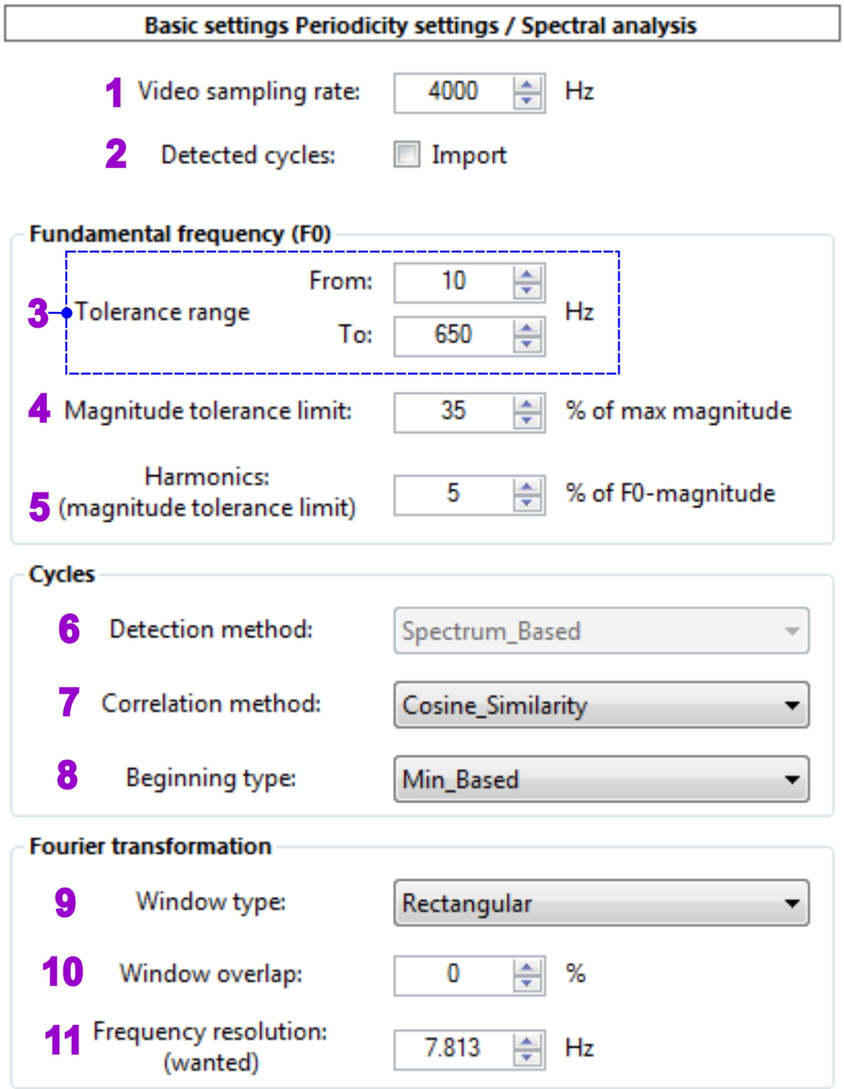


Figure 5 *– Glottis Analysis Tools: Analysis settings*

   i. *Cosine similarity*
   ii. *Euclidean distance*
   iii. *Pearson distance*
   iv. *Standard deviation*

C8. **Beginning type**: defines the initial value of the cycles.
   i. *Min-based* - each cycle begins at its minimum value.
   ii. *Max-based* - each cycle begins at its maximum value.

- **Fourier transformation** Note: Most parameters have fixed window lengths as described in their source papers and may provide significantly differing results if the windows are changed. The only parameters which are influenced by these settings are:
   i. *Harmonics Intensity* (eq. 20)
   ii. *Spectral Flatness* (eq. 24)
   iii. *Cepstral Peak Prominence (CPP)* (sec. 3.3- CPP)

C9. **Window type**: type of the window used for the computation of the Fourier transform of the given parameters above.

C10. **Window overlap**: defined as a percentage of the window length for the Fourier transform of the given parameters above.

C11. **Frequency resolution**: this value defines the **desired** bandwidth of each Fourier coefficient (Freq-step) in Hz and influences the length of the window for the Fourier transform

$$\text{Freq-step} = \frac{\text{Sampling Rate}}{\mathrm{Length}(\text{Fourier window})}. \tag{1}$$

Note: The length of the window for the Fourier transform depends on the sampling rate and the desired frequency resolution. The number of Fourier windows used depends on the length of the window and the given signal. If the signal is not long enough the actual frequency resolution will be coarser than the value which is set here. Hence we use the term **"desired"**.

2.2.3 Analysis: Parameter Analysis Settings

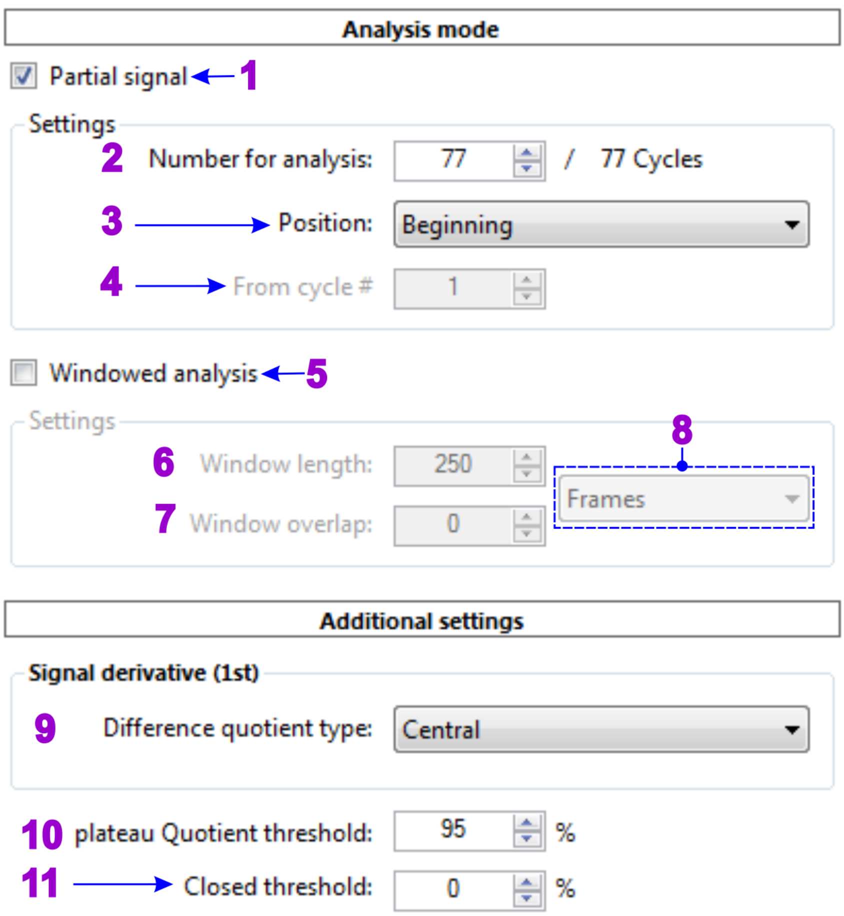


Figure 6 *– Glottis Analysis Tools: Analysis mode. (a) Analysis Tools (PVG), (b) Analysis Tools (Audio)*

After Cycle Detection the following Settings for the calculation of parameters (fig. 6) are available:

- **Partial signal analysis**
  Note: **these settings only influence the cycle based parameters.** Not cycle based parameters are calculated with the entire signal independent of here selected cycles.

  C1. **Partial signal checkbox**: allows the selection of only a certain coherent set of cycles for analysis

  C2. **Number for analysis**: select the number of cycles you want to analyze here. Default is all detected cycles.

  C3. **Position**: select the starting position of your subset of cycles, possible are

    i. *Beginning*

ii. *Middle*

iii. *End*

iv. *Manual*

C4. **From cycle#**: Only enabled if "Manual" was chosen from the "Position" drop down menu. Allows to define a definite starting cycle from which onwards the given number of cycles for analysis will be selected for analysis.

- **Windowed signal analysis**

C5. **Windowed analysis checkbox**: allows the subdivision of the signal in windows if selected

C6. **Window length**: the length of the windows which should be used for analysis.

C7. **Window overlap**: the overlap of each window with its previous window.

C8. **Unit selection**: the different selectable units for window length definition are

i. *Frames*

ii. *Cycles*

Note: The unit frames refers to the video frames. If audio data is loaded, the audio signal is windowed appropriate to the selected video windows. Analogously this applies if the chosen Unit is cycles. (i.e. both windows have the same duration in time). To ensure the congruency of the audio- and videodata during windowing, the chosen window is extended by 6 frames.

- **Additional settings**

C9. **Difference quotient type**: the two different selectable difference quotient types are defined as follows

i. *Central*: $\frac{d}{dt}f(t_0) = \frac{f(t_0 + \Delta t) - f(t_0 - \Delta t)}{2 \cdot \Delta t}$

ii. *Forward*: $\frac{d}{dt}f(t_0) = \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$

whereby Δt is the time step.

Note: To get a derivative vector of the same length as the signal, the last sample of this vector is generated by using the backward difference of the last two samples in the signal vector (= forward difference of last two datapoints). Analogously the first sample of the central difference is the forward difference of the first two datapoints and the last sample is the backward difference of the last two datapoints in the signal vector again.

C10. **Plateau quotient threshold**: the threshold value of the parameter "plateau quotient". (eq. 38)

C11. **Closed threshold**: the threshold value of the zero line in percent of the maximal GAW height. Everything below or equal this threshold will be considered as closed phase. (E.g. with the default value 0 only phases during which the value of the GAW is exactly zero pixels will be ascribed as closed phase. If the closed threshold is set to 5% regions of the GAW which have an value of equal or less than 5% of its global maximum are considered as closed phase)

Note: the starting and ending points of the closed phase are the datapoints which are closest to the given threshold (and therefore can be marginally greater than the given threshold).

- **Parameter selection**

C12. **Parameter domain selection**: Allows the selection of parameter sets sorted by their umbrella term.

C13. selects all available domains

C14. deselects all available domains

C15. **Details** for each selected parameter set a detailed description is displayed here

C16. starts or cancels the cycle detection or parameter calculation with the given settings.

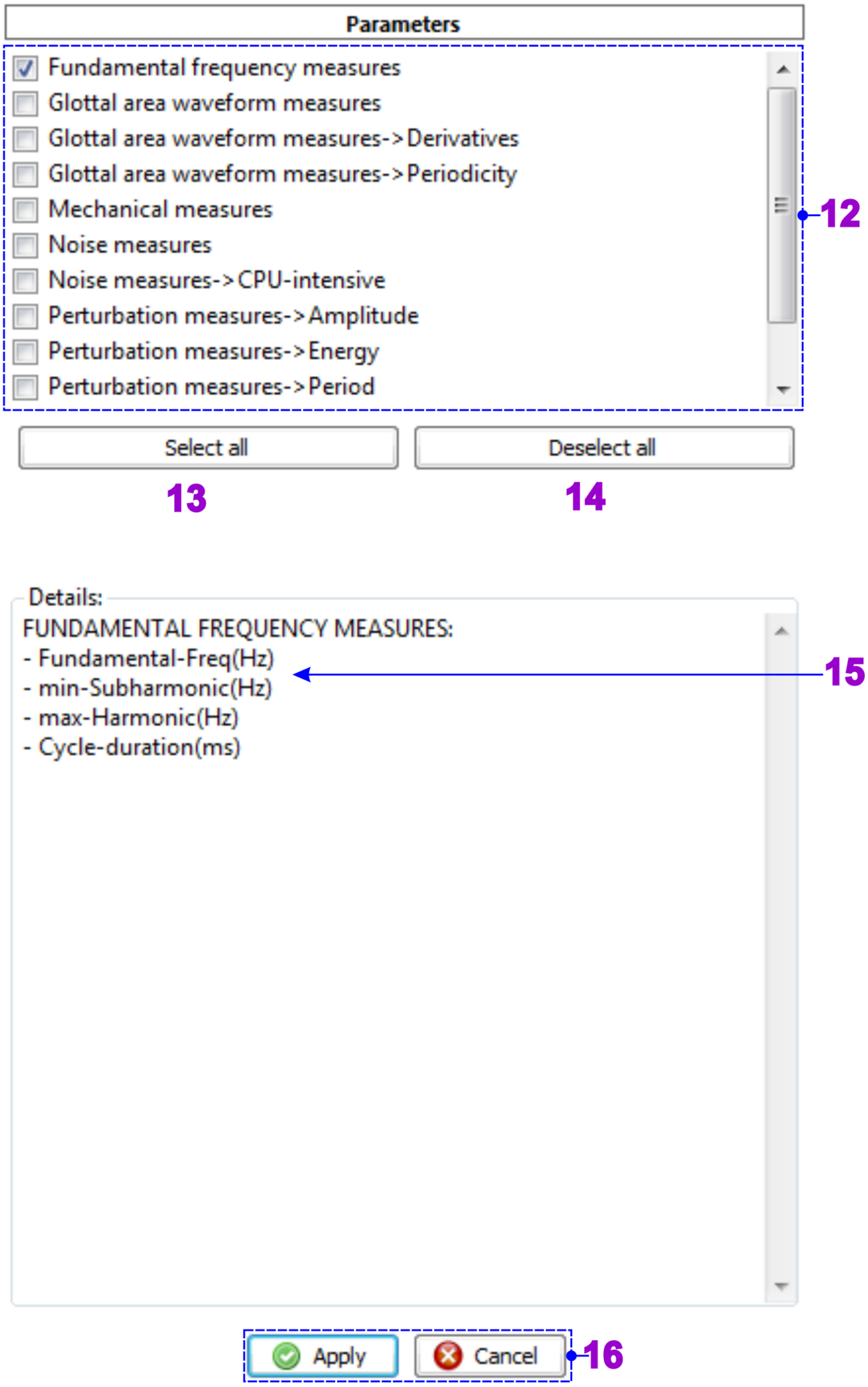


Figure 7 *– Glottis Analysis Tools: Parameters*

Analysis: Results Saving The result of the analysis can be saved in the following formats:

1. XML-File (*.xml)
2. CSV-File (*.csv)
3. ARFF-File (*.arff)

Before the results will be saved, you will be asked to select the desired file format (fig. 8). It is possible to select multiple formats.
In order to save not only summarized values (**mean, standard deviation, minimum, maximum**) but also the parameter values for each considered cycle check the *"Detailed list of parameters"* checkbox.

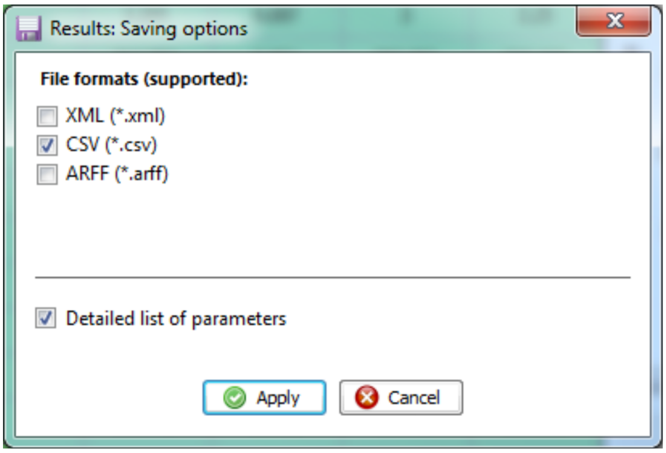


Figure 8 *– Glottis Analysis Tools: Results saving*

Note: The *Results Collector* (sec. 6.1) supports the XML- or CSV- formats and only allows collecting the summarized values!

3 Data Analysis

3.1 Fundamental frequency

These parameters can be calculated for the following input data:

- Glottal area waveform (GAW)
- Audio signal
- Glottal trajectories

The following parameters are calculated:

1. **Fundamental frequency**
   Fundamental-Freq $(\text{Hz})_i$ - oscillation frequency of the i^{th} cycle in Hz (eq. 2).
2. **Cycle Duration**
   Cycle-duration $(ms)_i$ - duration of the i^{th} cycle in ms.
3. **Maximum Harmonic**
   max-Harmonic (Hz) - maximum occurring harmonic frequency (a multiple of the fundamental frequency) in Hz.
4. **Minimum Subharmonic**
   min-Subharmonic (Hz) - minimum occurring subharmonic frequency (the fundamental frequency is a multiple of this frequency) in Hz.

3.1.1 Fundamental frequency: Determination

The fundamental frequency f_0^i of the i^{th} cycle is calculated as follows:

$$f_0^i\ (\text{Hz}) = \frac{\text{Sampling rate (Hz)}}{T_i}, \tag{2}$$

where T_i is duration of the i^{th} cycle in samples.

The cycles, and thereby the global fundamental frequency (f_0) will be detected using the following method:

> **Spectrum-based detection**: The Fourier spectrum of the given signal or signal window is calculated and the local maximum plus possible side peaks in the given *Tolerance range* (sec. 2.2.2 - C1) are detected. For every peak greater than the *Magnitude tolerance limit* (sec. 2.2.2 - C2) possible cycles are calculated. Subsequently the correlation between these cycles based on a cosine similarity like approach is calculated. The fundamental frequency which results in the cycles with the highest correlation is chosen as the final fundamental frequency f_0
>
> Note: the paramter Fundamental-Freq $(\text{Hz})_i$ is calculated based on the cycels which are determined by using the global f_0. Therefore the mean of all f_0^i and the directly calculated f_0 can differ marginally. Since the maximum and minimum harmonics are calculated by using directly f_0 they also tend to differ slightly from mean f_0^i if no significant harmonics or subharmonics in the signal exist.

3.2 Perturbation

The perturbation parameters quantify the average variability of the measured values.
These parameters can be calculated for the following signals:

- Glottal area waveform (GAW)
- Audio signal
- Glottal trajectories

Perturbation parameters will be calculated for the following signal features:

- Amplitude (sec. 3.2.1)
- Period: duration of the oscillation cycles (sec. 3.2.2)
- Energy (sec. 3.2.3)

each perturbation parameter is *cycle based*.

3.2.1 Amplitude

Further parameters:

- $A(i)$ - dynamic range $(\max - \min)$ of the i^{th} cycle,
- N - the number of analyzed cycles (equivalent to the number of elements in A).

The following *cycle based* amplitude-related parameters are calculated:

1. **mean Shimmer** [1, 2]

$$\text{Mean-Shim } (dB) = \frac{20}{N-1} \sum_{i=0}^{N-2} \left| \log_{10} \left[\frac{A(i)}{A(i+1)} \right] \right|. \tag{3}$$

2. **Shimmer (%)** [3] (**new version since GAT 2019**)

$$\text{Shim } (\%) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A(i) - A(i-1)|}{\frac{1}{N} \sum_{i=0}^{N-1} A(i)} \cdot 100. \tag{4}$$

3. **APQ (Amplitude Perturbation Quotient)** [4, 5, 6]

$$\text{APQ } (\%) = \frac{1}{N-k} \sum_{i=\frac{k-1}{2}}^{N-\frac{k-1}{2}-1} \left| 1 - \frac{k \cdot A(i)}{\sum_{j=-\frac{k-1}{2}}^{\frac{k-1}{2}} A(i+j)} \right| \cdot 100, \tag{5}$$

where k represents the number of cycles considered for computation of the quotients:

- $k = 3$: APQ-3 (%)
- $k = 5$: APQ-5 (%)
- $k = 11$: APQ-11 (%)

4. **APF (Amplitude Perturbation Factor)** [4, 5, 6]

$$\text{APF } (\%) = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| \frac{A(i) - A(i-1)}{A(i)} \right| \cdot 100. \tag{6}$$

5. **AVI (Amplitude Variability Index)** [2]

$$\text{AVI} = \log_{10} \left(1000 \cdot \frac{\frac{1}{N} \sum_{i=1}^{N} \left[A(i) - \bar{A} \right]^2}{\bar{A}^2} \right), \tag{7}$$

where $\bar{A}$ represents the mean amplitude for over all analyzed cycles.

3.2.2 Period

Further parameters:

- $p(i)$ - duration of the i^{th} cycle in ms,
- N - the number of analyzed cycles (equivalent to the number of elements in p).

The following *cycle based* period-related parameters are calculated:

1. **mean Jitter** [1]

$$\text{Mean-Jitter } (ms) = \frac{\sum_{i=1}^{N-1} |p(i) - p(i-1)|}{N-1}. \tag{8}$$

2. **Jitter (%)** [1]

$$\text{Jitter (\%)} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1} |p(i) - p(i-1)|}{\frac{1}{N}\sum_{i=0}^{N-1} p(i)} \cdot 100. \tag{9}$$

3. **Jitter Ratio** [2]

$$\text{Jitter Ratio} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1} |p(i) - p(i-1)|}{\frac{1}{N}\sum_{i=0}^{N-1} p(i)} \cdot 1000. \tag{10}$$

4. **Jitter Factor** [2]

$$\text{Jitter-Factor} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1} |f_i - f_{i-1}|}{\frac{1}{N}\sum_{i=0}^{N-1} f_i} \cdot 100, \tag{11}$$

where $f_i = \frac{1}{p(i)}$ is the frequency of the i^{th} cycle in Hz.

5. **PPQ (Period Perturbation Quotient)** [4, 5, 6]

$$\text{PPQ (\%)} = \frac{1}{N-k} \sum_{i=\frac{k-1}{2}}^{N-\frac{k-1}{2}-1} \left| 1 - \frac{k \cdot p(i)}{\sum_{j=-\frac{k-1}{2}}^{\frac{k-1}{2}} p(i+j)} \right| \cdot 100, \tag{12}$$

where k represents the number of cycles considered for the computation of the quotients:

- $k = 3$: PPQ-3 (%)
- $k = 5$: PPQ-5 (%)
- $k = 11$: PPQ-11 (%)

6. **PPF (Period Perturbation Factor)** [4, 5, 6]

$$\text{PPF (\%)} = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| \frac{p(i) - p(i-1)}{p(i)} \right| \cdot 100. \tag{13}$$

7. **RAP (Relative Average Perturbation)**

V1. [1]

$$\text{RAP-v1} = \frac{\sum_{i=1}^{N-2} \left| \frac{p(i-1)+p(i)+p(i+1)}{3} - p(i) \right|}{\sum_{i=0}^{N-1} p(i)}, \tag{14}$$

V2. [7, 2]

$$\text{RAP-v2} = \frac{\frac{1}{N-2} \sum_{i=1}^{N-2} \left| \frac{p(i-1)+p(i)+p(i+1)}{3} - p(i) \right|}{\frac{1}{N} \sum_{i=0}^{N-1} p(i)}. \tag{15}$$

8. **PVI (Period Variability Index)** [2]

$$\text{PVI} = 1000 \cdot \frac{\frac{1}{N} \sum_{i=1}^{N} [p(i) - \bar{p}]^2}{\bar{p}^2}, \tag{16}$$

where $\bar{p}$ represents the mean cycles duration.

3.2.3 Energy

Further parameters:

- $E(i)$ - signal energy within the i^{th} cycle,
- N - the number of analyzed cycles (equivalent to the number of elements in E).

The following *cycle based* energy-related parameters are calculated:

1. **EPQ (Energy Perturbation Quotient)** [4, 5, 6]

$$\text{EPQ (\%)} = \frac{1}{N-k} \sum_{i=\frac{k-1}{2}}^{N-\frac{k-1}{2}-1} \left| 1 - \frac{k \cdot E(i)}{\sum_{j=-\frac{k-1}{2}}^{\frac{k-1}{2}} E(i+j)} \right| \cdot 100, \tag{17}$$

where k represents the number of cycles considered for computation of quotients:

- $k = 3$: EPQ 3 (%)
- $k = 5$: EPQ-5 (%)
- $k = 11$: EPQ-11 (%)

2. **EPF (Energy Perturbation Factor)** [4, 5, 6]

$$\text{EPF (\%)} = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| \frac{E(i) - E(i-1)}{E(i)} \right| \cdot 100. \tag{18}$$

3.3 Noise

These measures can be calculated for the following signals:

- Glottal area waveform (GAW)
- Audio signal
- Glottal trajectories

Further parameters:

- $F(k)$ - k^{th} coefficient of the Fourier transform of the signal ($F(0)$ is the DC component),
- $C(k)$ - k^{th} Cepstrum coefficient [8]:

$$C(\omega) = 10 \cdot \log_{10}\left(\left|\mathcal{F}\{10 \cdot \log_{10}(|F(\omega)|^2)\}\right|^2\right), \tag{19}$$

- ω_0 - index of the Fourier coefficient that represents the fundamental frequency (f_0),
- H_{max} - maximum order of the harmonics of f_0,
- ω_{min} - index of the Fourier coefficient that represents the minimum occurring subharmonic for f_0.
- $f_{sampling}$ - sampling rate of the data.

The following parameters are calculated:

1. **Harmonics Intensity [litGAT:Hiraoka1984]**

   The parameter Harmonics Intensity is *window based*. The used window size can be influenced by the Fourier transformation settings (sec. 2.2.2 - C9 - C11)

   $$\text{Harmonics-Intensity (\%)} = 100 \cdot \frac{\sum\limits_{n=2}^{H_{max}} |F(n \cdot \omega_0)|}{\sum\limits_{\omega \geq 1} |F(\omega)|}. \tag{20}$$

2. **Harmonics-to-Noise Ratio (HNR)** [9]

   The parameter HNR is *cycle based*.

   $$\text{HNR } (dB) = 10 \cdot \log_{10}\left(\frac{H}{N}\right), \tag{21}$$

   with

   $$H = n \int_0^T f_A^2(t)dt,$$

   $$N = \sum_{i=1}^{n} \int_0^{T_i} |f_A(t) - f_i(t)|^2 \, dt,$$

   where

   - n - number of considered cycles,
   - T_i - duration of the i^{th} cycle in ms,

- $f_i(t)$ - i^{th} cycle of the signal ($0 \leq t \leq T_i$),
- $T = \max\limits_{1 \leq i \leq n} (T_i)$,
- $f_A(t)$ - averaged cycle of the signal

$$f_A(t) = \frac{1}{n} \cdot \sum_{i=1}^{n} f_i(t), \qquad 0 \leq t \leq T. \tag{22}$$

3. **Normalized Noise Energy (NNE)** [10]

The parameter NNE is *window based*. A fix windowing algorithm is used:

- The signal is separated in 40 ms frames with 20 ms overlap.
- For each frame the median duration of all cycles with their starting position within this frame is calculated.
- From the original signal starting at the starting position of the frame a slice of data is cutted.
- This slice is as long as seven times the median duration of the cycles in the current frame.
- The slice is hamming windowed and afterwards filled with zeros to a length of 102,4 ms.
- The fourier transform of these zeropadded windows is calculated.

For each of these Fourier windows, let $s(\tau)$ be the periodic component, and $w(\tau)$ the noise component of the discretized signal $f(\tau)$, respectively, then

$$\Downarrow$$

$$f(\tau) = s(\tau) + w(\tau), \qquad \tau = 1, \ldots, M.$$

Discrete Fourier transform $\Downarrow$

$$F(\omega) = S(\omega) + W(\omega), \qquad \omega = 0, \ldots, N-1.$$

This leads to:

$$\text{NNE } (dB) = 10 \cdot \log_{10} \left(\frac{\sum\limits_{\omega=\omega_{min}}^{\omega_{max}} \left|\hat{W}(\omega)\right|^2}{\sum\limits_{\omega=\omega_{min}}^{\omega_{max}} |F(\omega)|^2} \right), \tag{23}$$

where

- $\omega_{max} = \min(\lceil \frac{\min(0.5 \cdot f_{sampling}, 5000)}{f_{sampling}} \cdot N \rceil - 1, \frac{N}{2} - 1)$,
- $\omega_{min} = \lceil \frac{50}{f_{sampling}} \cdot N \rceil - 1$,

with

$$\left|\hat{W}(\omega)\right|^2 = \begin{cases} |F(\omega)|^2\,, & \omega \in D_i \\ \frac{1}{2}\left(\sum\limits_{r\in D_i}\frac{|F(r)|^2}{N_i} + \sum\limits_{r\in D_{i+1}}\frac{|F(r)|^2}{N_{i+1}}\right), & \omega \in P_i \end{cases},$$

where

- $\left|\hat{W}(\omega)\right|^2 \approx |W(\omega)|^2$ (Fourier transform of the noise component $w(\tau)$),
- $D_i = \left\{r : (i-1)\cdot\omega_0 + \frac{2N}{M} \le r \le i\cdot\omega_0 - \frac{2N}{M}\right\}$,
- $P_i = \left\{r : i\cdot\omega_0 - \frac{2N}{M} \le r \le i\cdot\omega_0 + \frac{2N}{M}\right\}$,
- $N_i = |D_i|$ (Cardinality), $\quad 1 \le i \le H_{max}$.

The mean of all windows is the returned NNE parameter. If more than two successive D_i regions cant be calculated because of too close peaks, the respective window is excluded from analysis.

4. **Spectral Flatness** [5]

   The parameter Spectral Flatness is *window based*. The used window size can be influenced by the Fourier transformation settings (sec. 2.2.2 - C9 - C11)

$$\text{Spectral-Flatness }(SFM) = \frac{20}{N}\cdot\left(\sum_{i=1}^{N/2}\log_{10}|F(\omega)|^2\right) - 10\cdot\log_{10}\left(\frac{2}{N}\cdot\sum_{i=1}^{N/2}|F(\omega)|^2\right). \qquad (24)$$

5. **Glottal-to-Noise Excitation Ratio (GNE)** [4, 11, 5]

   The parameter GNE is *window based*. The following algorithm is used:

   i) Changing the sampling rate to 50 kHz by sinosoidal interpolation
   ii) High pass filtering of the considered signal using:
      - Filter order: 10
      - cutoff frequency: 10 Hz
      - phase shift: zero
      - Method: Butterworth
      - Window: entire signal
   iii) Changing the sampling rate to 10 kHz
   iv) Discarding the first 10 ms of data (because of filter effects)
   v) Inverse filtering of the considered signal (calculation of the linear-prediction error signal) using:
      - Predictor order: 13
      - Method: auto correlation
      - Prediction window: Hann-window (length 30 ms, overlap 20 ms)
   vi) Discarding the first 13 samples of data (because of LP-filter effects)

vii) Subdividing the linear-prediction error signal in windows (length 500 ms, overlap 250 ms)

viii) Calculating the Fourier transformation of the windows.

ix) Calculating the Hilbert envelopes $\mathcal{H}(f_c)$ (the absolute value of the complex analytic signal) of frequency bands with fixed bandwidth (B_f) and different center frequencies (f_c) for each window. The center frequencies are divided in a group of left (f_{cl}) and right (f_{cr}) center frequencies:

- F_s - sampling rate of the signal $= 10000$
- $B_f = 3000Hz$
- f_c - Step $= 100Hz$
- f_{cl} - Range (Hz): $\frac{B_f}{2} \leq f_{cl} \leq 2000Hz$
- f_{cr} - Range (Hz): $3000Hz \leq f_{cr} \leq 5000Hz - \frac{B_f}{2}$

  Note: every slice of the fft window is multiplied with a hann window before back transformation.

x) Calculating $\rho_{k,j} = \max\left\{\text{CrossCorrelation}\left[\mathcal{H}(f_{cl}^k), \mathcal{H}(f_{cr}^j)\right]\right\}, \quad k, j \in \{1,2,3,4,5,6\}$

xi) $GNE_i = \max_{k,j}(\rho_{k,j})$ Whereas i is the ith 500 ms window
Note: As other fixed window size Parameters GNE will not be calculated for too short signals. The minimum length for which one GNE window is calculated is 531.3 ms.

6. **Waveform Matching Coefficient (WMC)** [5]

   The parameter WMC is *window based*. A fix windowing algorithm is used:

   - The mean Cycle duration of all detected cycles is calculated.
   - The signal is divided in subvectors of this length.
   - These subvectors are used for the following calculation.

   Note: This preprocessing of the signal is necessary since the calculation of cosine similarity requires data vectors of equal length.

$$\text{max-WMC} = \max_{1 \leq i \leq n-1} \frac{\langle f_i(t), f_{i-1}(t) \rangle}{\|f_i(t)\|_2 \cdot \|f_{i-1}(t)\|_2}, \tag{25}$$

   where

   - n - number of considered subvectors,
   - D - duration of each subvector in frames (average of all cycle lengths),
   - $f_k(t)$ - k^{th} subvector of the signal ($0 \leq t \leq D$)

7. **Mean Waveform Matching Coefficient (MWMC)** [5]

   The parameter MWMC is *window based*. A fix windowing algorithm is used:

   - The mean Cycle duration of all detected cycles is calculated.
   - The signal is divided in subvectors of this length.
   - These subvectors are used for the following calculation.

   Note: This preprocessing of the signal is necessary since the calculation of cosine similarity requires data vectors of equal length.

$$\text{mean-WMC} = \underset{1 \leq i \leq n-1}{\text{mean}} \frac{\langle f_i(t), f_{i-1}(t) \rangle}{\|f_i(t)\|_2 \cdot \|f_{i-1}(t)\|_2}, \tag{26}$$

   where

- n - number of considered subvector,
- D - duration of each subvector in frames (average of all cycle lengths),
- $f_k(t)$ - k^{th} subvector of the signal ($0 \leq t \leq D$).

8. **Signal-to-Noise Ratio (SNR)**

V1. [12, 13]

The parameter SNR-V1 is *window based*. The following algorithm is used:

i) Separating the signal in 160 ms gaussian windows with 20 ms signal shift (140 ms overlap).

ii) Zeropadding the signal windows to a length of 320 ms.

iii) Calculating the Fourier transformation of each zeropadded window i: $(F(\omega)_i)$.

iv) Further calculations are only done with the left half of this spectrum.

v) Upscaling $F(\omega)_i : \mathrm{Length}\left(\hat{F}_i\right) = 4 \cdot \mathrm{Length}\left(F_i\right)$

vi) Low pass filtering $\hat{F}_i$ using:
- Filter order: 4
- cutoff frequency: $0.25 \cdot f_{sampling}\, Hz$
- phase shift: zero
- Method: Butterworth

vii) Calculating the positions of f_0 (ω_0) and all possible harmonic positions (ω_i) without considering a threshold.

viii) Calculating an approximation for the noise in between the harmonics for each window i:
- define the "valleys" between the harmonics (D_j):

$$f_{step} = \frac{1}{0.16s \cdot 8}$$

$$D_{j+1} = \left\{ r : \omega_j + \frac{12Hz}{f_{step}} \leq r \leq \omega_{j+1} - \frac{12Hz}{f_{step}} \right\}$$

- special case region before the harmonics:

$$D_0 = \left\{ r : 0 + \frac{12Hz}{f_{step}} \leq r \leq \omega_0 - \frac{12Hz}{f_{step}} \right\}$$

- special case region after the last harmonic:

$$D_{jmax+1} = \left\{ r : \omega_{jmax} + \frac{12Hz}{f_{step}} \leq r \leq \omega_{jmax} + \omega_0 - \frac{12Hz}{f_{step}} \right\}$$

- Averaging $\hat{W}_j = \frac{1}{N_j} \left(\sum_{r \in D_j} F(r) \right)$

  where N_j is the number of samples in region D_j

ix) Calculating the half-value widths of the harmonics

x) All harmonics with half-value widths outside the range of 12 Hz to 24 Hz are excluded.

xi) All harmonics with peak values less than 12 dB above the mean noise of both surrounding "valleys" ($\hat{W}_j$ and $\hat{W}_{j+1}$) are excluded.

xii) Choosing the half-value (B_0) width of the greatest harmonic (usually f_0).

xiii) Approximating the pure harmonic spectrum ($\hat{H}_i$), using the harmonics positions,their peak value and the half-value width as follows:

- Each remaining harmonic is represented by a Gaussian shaped function
- Each Gaussian function has the same maximum and maximum position than their corresponding harmonic
- Each Gaussian function has the half-value with B_0.
- a vector of zeros as long as the first half of the Fourier window is generated
- To this vector the Gaussian functions are added

xiv) Calculating the harmonic short-time energy of window i: $E_h^i = \sum_{\omega=0}^{\text{Length}(\hat{H})} \left|\hat{H}(\omega)_i\right|^2$

xv) Calculating the total short-time energy of window i: $E_t^i = \sum_{\omega=0}^{\text{Length}(\hat{F})} \left|\hat{F}(\omega)_i\right|^2$

xvi) $SNR-v2_i(dB) = \min\left(10 \cdot \log_{10}\left(\frac{E_t^i}{E_t^i - E_h^i}\right), 36dB\right)$

Note: In the original source this algorithm is repeated with a narrower B_0, if $E_h^i > E_t^i$ until $E_h^i < E_t^i$. In such cases our implementation returns the value NaN for the corresponding window.

V2. [14]

The parameter SNR-V2 is *window based*. The following algorithm is used:

i) Short-term inverse filtering of the signal $f(t)$ (calculation of the linear-prediction error signal $s(t)$) using:
   - Predictor order: 14
   - Method: auto correlation
   - Prediction window: Hamming-window (length 20 ms, no overlap.)

ii) Discarding the first 14 samples of $s(t)$ (because of LP-filter effects)

iii) Long-term inverse filtering of $s(t)$ (calculation of the linear-prediction error signal $n(t)$) using:
   - Prediction order: 3,
   - Method: auto correlation
   - Window for minimizing prediction error: 2.5 ms,

iv) Optimizing long term filter over 1.25 - 17.5 ms range before current prediction start-sample $s(t_{windowstart})$.

v) Predicting the following 2.5 ms of signal beginning with $s(t_{windowstart})$ with the long term prediction coefficients.

vi) Discarding the first 17.5 ms of $n(t)$ (because no long term prediction was made here)

vii) Discarding the first 3 samples of $n(t)$ (because of LP-filter effects)

viii) $\hat{f(t)} = f(t + 3 + 14 + \lceil f_{sampling} * 0.0175 \rceil)$ ($\left|\hat{f}\right| = |f| - 3 - 14 - \lceil f_{sampling} * 0.0175 \rceil$ (cardinality)).

ix) SNR-v2 $(dB) = 20 \cdot \log_{10}\left[\frac{\sqrt{\sum_{t=1}^{N}\left|\hat{f}(t)\right|^2}}{\sqrt{\sum_{t=1}^{N}|n(t)|^2}} - 1\right]$.

9. **Cepstral Peak Prominence (CPP)**

I. [15, 8]

The parameter CPP-I is *window based*. The used window size can be influenced by the Fourier transformation settings (sec. 2.2.2 - C9 - C11). The following algorithm is used:

i) Calculating the Cepstrum for the given Fourier window (eq. 19).

ii) Detecting the first Rahmonic.

iii) Calculating the regression line ($R(k)$) in the Range of 1 ms to half the length of the Cepstrum.

iv) Subtracting the maximum of the first Rahmonic from the corresponding value on the regression line: $CPP_i = C(r_0)^i - R(r_0)^i$

II. [15, 8]

The parameter CPP-II is *window based*. The used window size can be influenced by the Fourier transformation settings (sec. 2.2.2 - C9 - C11). The following algorithm is used:

i) Scaling the Fourier windows to Db

ii) Averaging all given rescaled Fourier windows to an average window ($\bar{F}_{log}$) if the have the same length.

iii) Calculating the Cepstrum: $C(\omega) = 10 \cdot \log_{10}\left(\left|\mathcal{F}\{\bar{F}(\omega)_{log}\}\right|^2\right)$.

iv) Detecting the first Rahmonic.

v) Calculating the regression line ($R(k)$) in the Range of 1 ms to half the length of the Cepstrum.

vi) Subtracting the maximum of the first Rahmonic from the corresponding value on the regression line: $CPP = C(r_0) - R(r_0)$

3.4 Mechanical

These parameters can be calculated for the following signals:

- Glottal area waveform (GAW)
- Glottal trajectories

Further parameters:

- T_i - duration of the i^{th} cycle in ms,
- A_i - dynamic range $(\max - \min)$ of the i^{th} cycle,
- L_i - mean glottis length (distance between Anterior- and Posterior- points) for the i^{th} cycle,
- $s(t)$ - absolute magnitude of the 1^{st} derivative of the considered signal for i^{th} cycle ($t \subset T_i$).

The following *cycle based* mechanical parameters are calculated:

1. **Stiffness** [16]

$$\text{Stiffness} = \frac{\max_{t \subset T_i}(s(t))}{A_i}. \tag{27}$$

2. **Peak Closing Velocity** [17]

$$\text{Peak-Clossing-Velocity} = 2\pi \cdot \frac{A_i}{2 \cdot T_i}. \tag{28}$$

3. **Peak Acceleration** [17]

$$\text{Peak-Acceleration} = 4\pi^2 \cdot \frac{A_i}{2 \cdot T_i^2}. \tag{29}$$

4. **Amplitude-to-Length Ratio**

$$\text{Amplitude-Length-Ratio} = \frac{A_i}{L_i}. \tag{30}$$

3.5 Glottal area waveform

These parameters can be calculated for the following signals:

- Glottal area waveform (GAW)
- Glottal trajectories

Further parameters:

- T_i - duration of i^{th} cycle in ms,
- t^i_{closed} - *closed phase duration*, defined when the glottis is closed (glottal area ≤ 0),
- t^i_{open} - *open phase duration*, defined when the glottis is open (glottal area > 0),
- $[C \to O]_i$ - opening part of open phase (closed-to-open) for i^{th} cycle,
- $[O \to C]_i$ - closing part of open phase (open-to-closed) for i^{th} cycle,
- GA_i - glottal area for i^{th} cycle
- A_i - Glottal area dynamic range $[\max(GA_i) - \min(GA_i)]$ for i^{th} cycle,

The Phases are dependent of the zero line, which is 0 by default. However, it can be adjusted using the settings (sec. 2.2.3 - C11).
Note: To avoid discretization gaps between the different cycles (i.e. if one cycle ends at datapoint x and the next one starts at datapoint x+1 the between this datapoints would be undefined), the phases for each cycle are calculated from its first datapoint up to the last plus one datapoint. The following *cycle based* quotients are calculated (fig. 9):

1. **Open Quotient** [2]

$$\text{Open-Quotient } (O_q) = \frac{t^i_{open}}{T_i}. \tag{31}$$

2. **Closing Quotient** [18]

$$\text{Closing-Quotient } (Cl_q) = \frac{[O \to C]_i}{T_i}. \tag{32}$$

3. **Speed Quotient** [2]

$$\text{Speed-Quotient } (S_q) = \frac{[C \to O]_i}{[O \to C]_i} \tag{33}$$

4. **Speed Index** [2]

$$\text{Speed-Index } (SI) = \frac{[C \to O]_i - [O \to C]_i}{t^i_{open}} = \frac{[C \to O]_i - [O \to C]_i}{[C \to O]_i + [O \to C]_i} = \frac{(S_q - 1)}{(S_q + 1)}. \tag{34}$$

5. **Rate Quotient** [2]

$$\text{Rate-Quotient } (R_q) = \frac{(t^i_{closed} + [C \to O]_i)}{[O \to C]_i}. \tag{35}$$

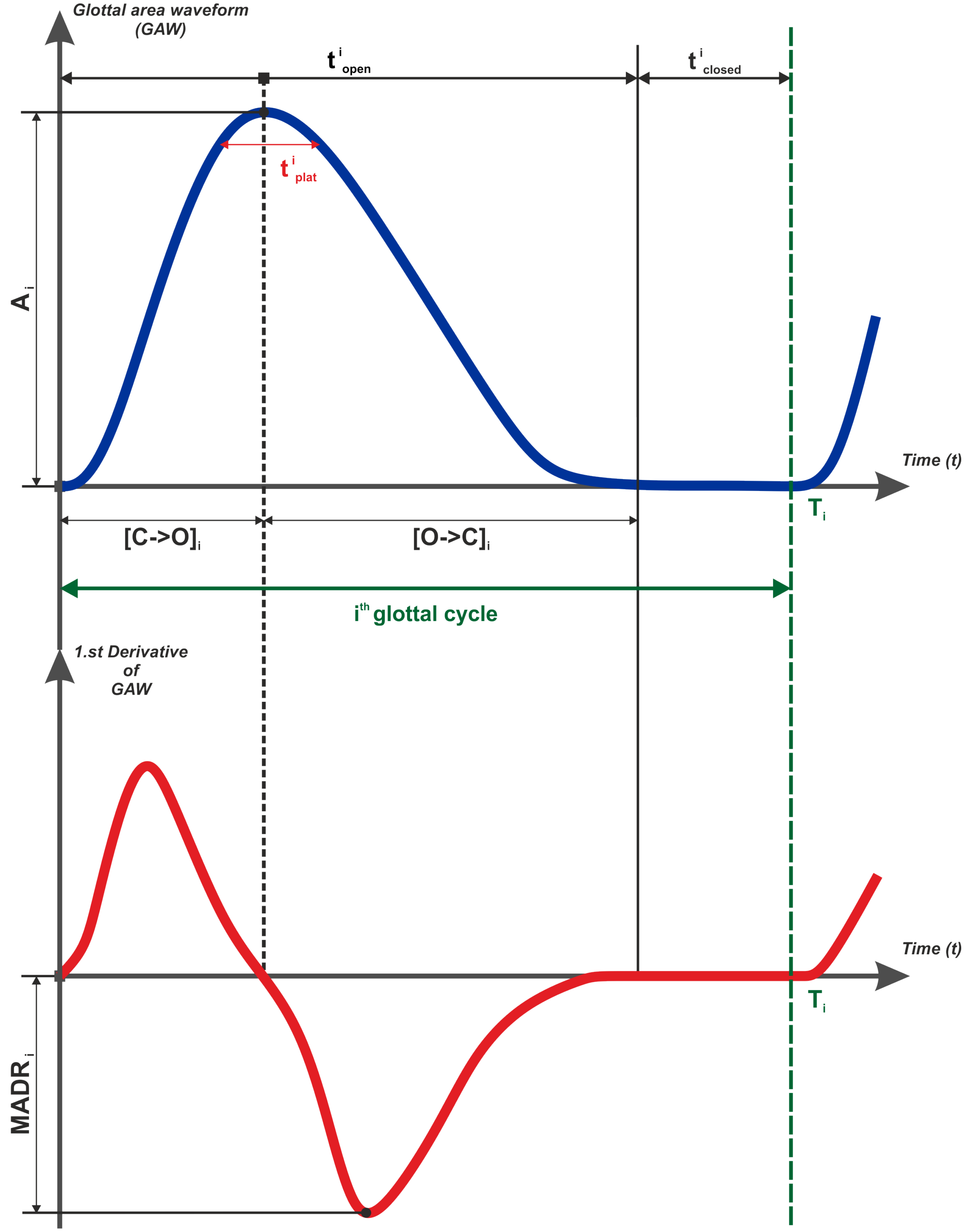


***Figure 9** – Visualization of Open Quotient, Closing Quotient, Speed Quotient, Speed Index, Rate Quotient, Asymmetry Quotient, Plateau Quotient, and GAW Derivatives parameters (sec. 3.5.2).*

6. **Asymmetry Quotient** [19]

$$\text{Asymmetry-Quotient } (A_q) = \frac{S_q}{1 + S_q}. \tag{36}$$

7. **Glottis Gap Index** [20]

$$\text{Glottis-Gap-Index } (GGI) = \frac{\min(GA_i)}{\max(GA_i)}, \tag{37}$$

8. **Plateau Quotient** [21]

$$\text{Plateau-Quotient } (P_q) = \frac{t^i_{plat}}{t^i_{open}}, \tag{38}$$

   where t^i_{plat} is the plateau phase duration of the i^{th} cycle,
   i.e. the phase when the glottal area is greater than 95% of its maximum.
   this threshold value of 95% can be changed using (sec. 2.2.3 - C10)

9. **Glottal Area Index** [22]

$$\text{Glottal-Area-Index } (AC/OQ) = \frac{A_i}{\max(GA_i) \cdot O_q}, \tag{39}$$

3.5.1 Glottal area waveform: Periodicity

These parameters originally proposed by Qiu et al. [23] for the digital kymography were adapted to the signals generated with high speed imaging (fig. 10). Both parameters are *cycle based*.

1. **Amplitude periodicity** [23]

$$\text{Amplitude-Periodicity} = \frac{\min(A_i, A_{i+1})}{\max(A_i, A_{i+1})}. \tag{40}$$

2. **Time periodicity** [23]

$$\text{Time-Periodicity} = \frac{\min(T_i, T_{i+1})}{\max(T_i, T_{i+1})}. \tag{41}$$

3.5.2 Glottal area waveform: Derivatives

These parameters originally used for the glottal air flow [18, 24] were adapted to the signals generated from high speed imaging (fig. 9). Both parameters are *cycle based*.

1. **Maximum Area Declination Rate** [18, 24, 25, 26]

   Maximum-Area-Declination-Rate $(MADR)$ is defined as the absolute maximum amplitude of the negative peak of the 1^{st} derivative of the considered signal (here: glottal area waveform or glottal trajectories).

2. **Amplitude Quotient** [24, 26]

$$\text{Amplitude-Quotient } (AQ) = \frac{A_i}{MADR_i}. \tag{42}$$

3.6 Symmetry

These parameters can be calculated for the following signals:

- Glottal area waveform (GAW)

- Glottal trajectories

Symmetry parameters are separated into following the subsets:

- Unilateral symmetry parameters (sec. 3.6.1)
- Lateral symmetry parameters: not suitable for between-group comparison (sec. 3.6.2)

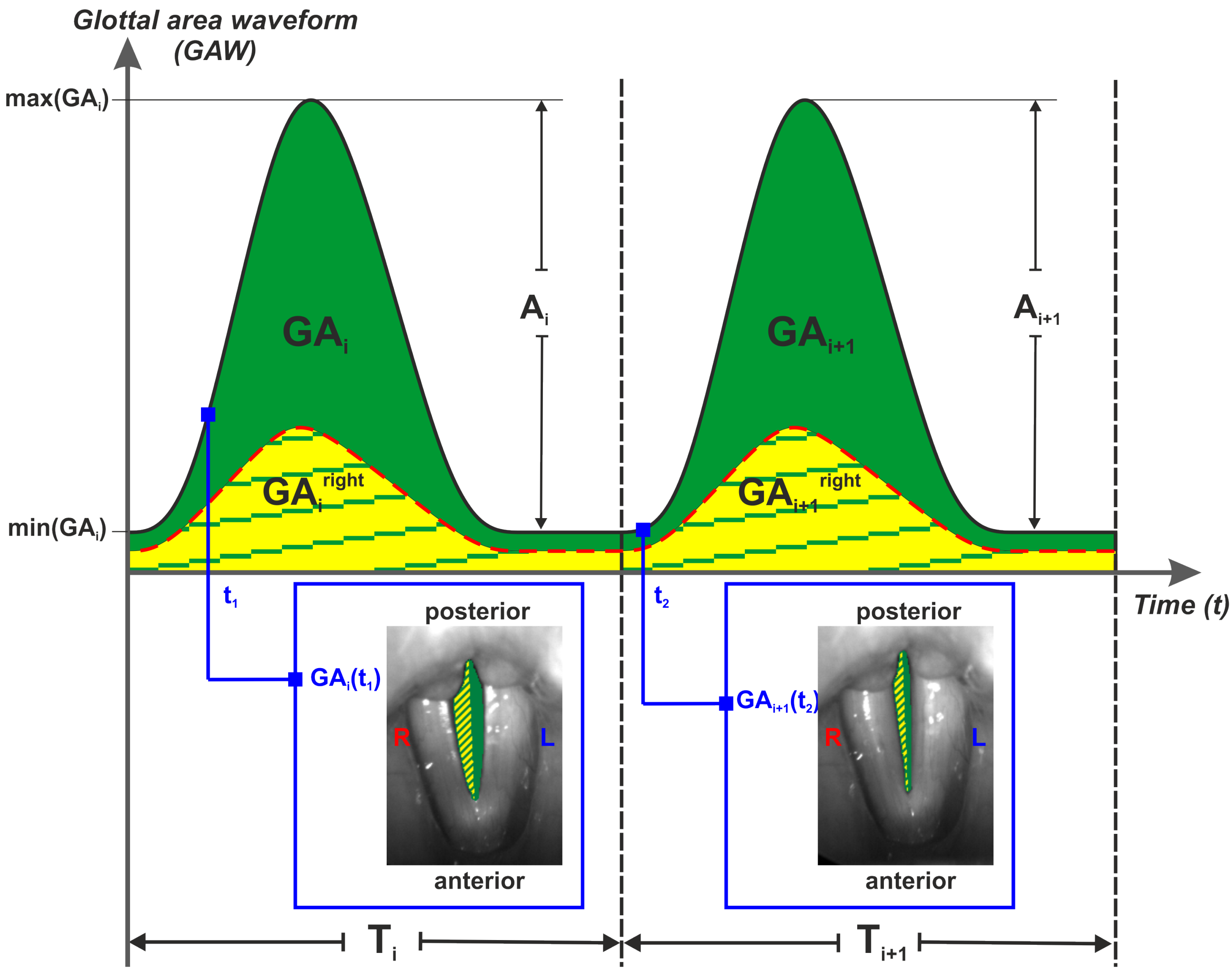


Figure 10 *– Visualization of the Glottis Gap Index, GAW Periodicity parameters (sec. 3.5.1) and Symmetry parameters (sec. 3.6).*

Furthermore:

- T_i - duration of the i^{th} cycle in ms,
- GA_i - glottal area waveform of the i^{th} cycle,
- A_i - glottal area dynamic range $[\max(GA_i) - \min(GA_i)]$ of the i^{th} cycle,
- L - left side, R - right side.

3.6.1 Symmetry: Unilateral

The following unilateral *cycle based* symmetry parameters are calculated (fig. 10):

1. **Phase Asymmetry Index** [23, 20]

$$\text{Phase-Asymmetry-Index} = \frac{\left|t_i^L(\max) - t_i^R(\max)\right|}{T_i}, \tag{43}$$

where $t_i^{side}(\max)$ is the time at which the glottal area of the left/right side in the i^{th} cycle is maximal.

2. **Spatial Symmetry Index** (for GAW only!) [20]

$$\text{Spatial-Symmetry-Index} = \frac{\left| \sum\limits_{t \subset T_i} GA_i^L(t) - \sum\limits_{t \subset T_i} GA_i^R(t) \right|}{\sum\limits_{t \subset T_i} GA_i(t)}. \tag{44}$$

3. **Dynamic Range Symmetry Index**

$$\text{DynamicRange-Symmetry-Index} = \frac{\min(A_i^L, A_i^R)}{\max(A_i^L, A_i^R)}. \tag{45}$$

4. **Amplitude Symmetry Index**

$$\text{Amplitude-Symmetry-Index} = \frac{\min\left(\max[GA_i^L], \max[GA_i^R]\right)}{\max\left(\max[GA_i^L], \max[GA_i^R]\right)}. \tag{46}$$

5. **Waveform Symmetry Index** [27]

$$\text{Waveform-Symmetry-Index} = 0.5 \cdot \left[1 + \frac{\langle GA_i^L, GA_i^R \rangle}{\|GA_i^L\|_2 \cdot \|GA_i^R\|_2}\right]. \tag{47}$$

3.6.2 Symmetry: Lateral

In the *Glottis Analysis Tools* the following lateral *cycle based* symmetry parameters are calculated (fig. 10):

1. **Phase Asymmetry** [23]

$$\text{Phase-Asymmetry*} = \frac{t_i^L(\max) - t_i^R(\max)}{T_i}, \tag{48}$$

where $t_i^{side}(\max)$ is the time at which the glottal area of the left/right side in the i^{th} cycle is maximal.

2. **Spatial Symmetry** (for GAW only!)

$$\text{Spatial-Symmetry*} = \frac{\sum\limits_{t \subset T_i} GA_i^L(t) - \sum\limits_{t \subset T_i} GA_i^R(t)}{\sum\limits_{t \subset T_i} GA_i(t)}. \tag{49}$$

3. **Dynamic Range Symmetry**

$$\text{DynamicRange-Symmetry*} = \frac{A_i^L}{A_i^R}. \tag{50}$$

4. **Amplitude Symmetry**

$$\text{Amplitude-Symmetry*} = \frac{\max(GA_i^L)}{\max(GA_i^R)}. \tag{51}$$

3.7 Phonovibrogram

These parameters can only be calculated for the Phonovibrogram (PVG).
The following*cycle based* PVG parameters are calculated:

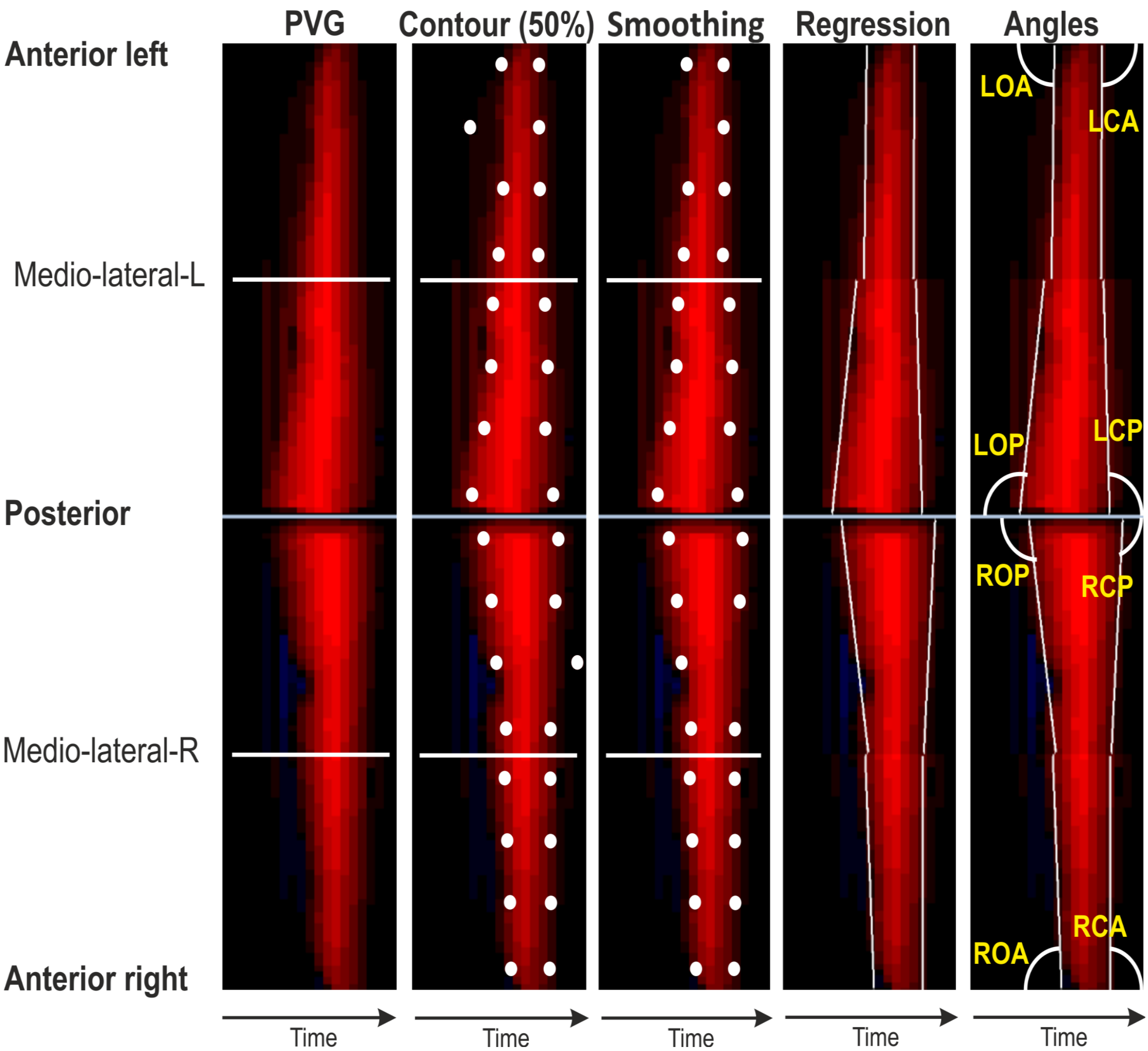


***Figure 11** – Visualization of the PVG Contour Angles.*

1. **Contour Angles** (fig. 11) [28, 29, 30]
   Contour-Angles (deg) are calculated in both anterior and posterior parts during opening as well as closing of the vocal folds for the left and right side of the PVG, respectively. Figure 11 roughly illustrates the calculation algorithm of the contour angles.
   Therein, $CA_i^{side,Item}$ denotes the Contour-Angles of the i^{th} cycle,
   where $side$ represents the corresponding side of the PVG:

   - L - Left side,
   - R - Right side,

   and $Item$ indicates the position of the related Contour-Angle:

   - $Item = OA$: Opening - Anterior,

- $Item = OP$: Opening - Posterior,
- $Item = CA$: Closing - Anterior,
- $Item = CP$: Closing - Posterior.

2. **Contour Angles Symmetry Index** [31]

$$\text{ContourAngles-Symmetry-Index}_{Item} = \frac{\min(CA_i^{L,Item}, CA_i^{R,Item})}{\max(CA_i^{L,Item}, CA_i^{R,Item})}. \tag{52}$$

3. **Contour Angles Symmetry** (Not suitable for between-group comparison!) [31]

$$\text{ContourAngles-Symmetry*}_{Item} = \frac{CA_i^{L,Item}}{CA_i^{R,Item}}. \tag{53}$$

4 Video Processing

Video Processing provides the following features:

- **Segmentation Tools** (sec. 4.1): perform the segmentation of the HSE recordings to create a representation of some characteristics of the vocal folds' vibratory motions like a phonovibrogram (PVG) and a glottal area waveform (GAW).
- **Analysis Tools (PVG)** (sec. 4.2): perform the analysis of the obtained characteristics from the HSE recording as well as the analysis of an acoustic signal corresponding to the HSE.

4.1 Segmentation Tools

These tools allow the segmentation of the glottis contour from HSE recordings, the extraction of some characteristics of the vocal folds' vibratory motions like a glottal area waveform (GAW) and the computation of a phonovibrogram (PVG). Furthermore, the *Segmentation Tools* facilitate the synchronization between the HSE recording and the related acoustic signal. The program supports the following video and multipage-Image formats:

- **Videofile** (*.avi),
- **Basic Loadable picture file** (*.bld).
- **Vision Research Phantom Cine Camera Video Format** (*.cine)
- **Videofile** (*.mp4),

4.1.1 Main menu

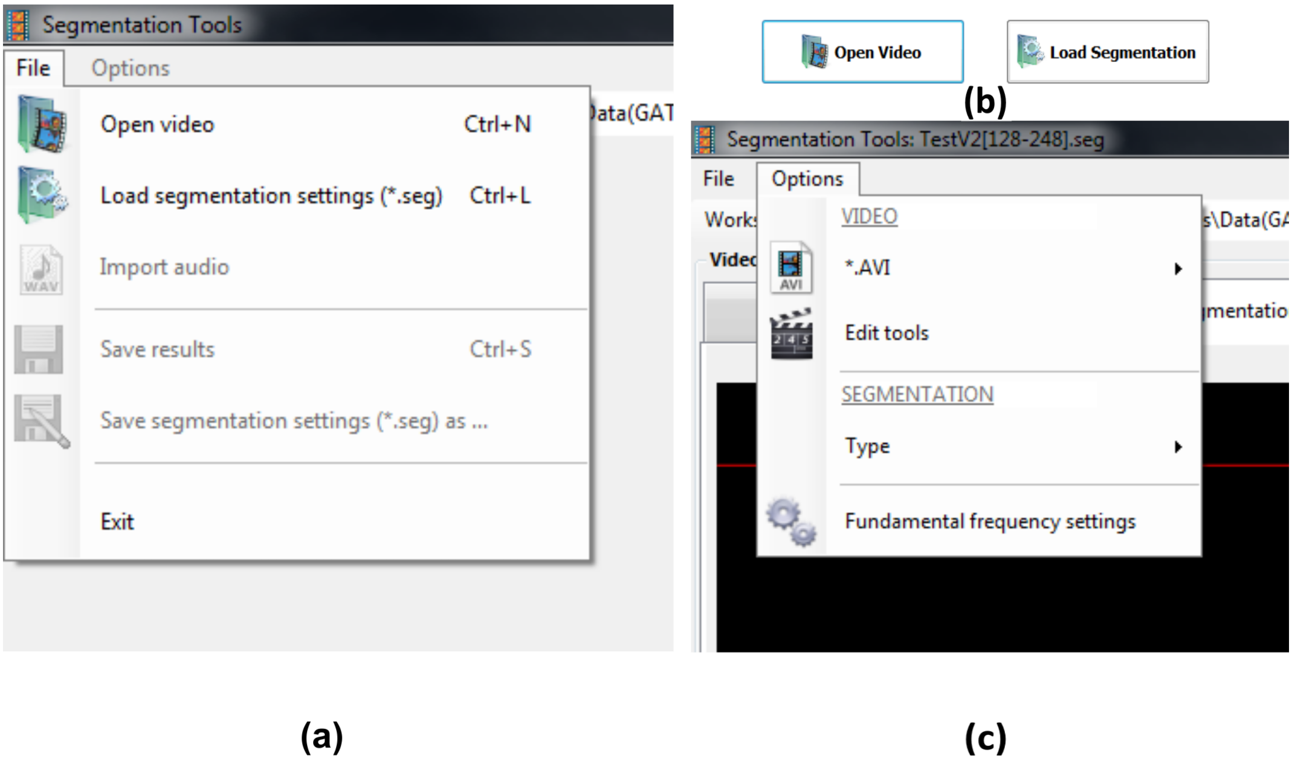


Figure 12 *– Segmentation Tools: Main menu. (a) File context menu, (b) shortcut buttons (c) Options context menu*

Main menu: File (fig. 12a)/b)

- **Open video** from context menu or button: opens a new video for segmentation.
- **Load segmentation settings (*.seg)** from context menu or button: loads an existing *Segmentation Settings File* (*.seg). This file contains only the settings required for the segmentation and does not contain the segmented glottis contours or computed PVGs.
- **Import audio**: Opens a menu (fig. 15b) that allows to import an audio file (".wav") from the computer and attaches it to the current video.

  Note: it is important to ensure that the correct audiodata is selected since currently no additional preventive measures for avoiding mismatches are implemented.
- **Save results**: (sec. 4.1.4 - C28).
- **Save segmentation settings (*.seg) as ...**: saves only the segmentation settings.
- **Exit**: closes the *Segmentation Tools* and returns to the *Glottis Analysis Tools*: Main menu (fig. 1).

Main menu: Options (fig. 12c)

- ***.AVI**:provides the actions available for AVI-Files (*avi).
  - $\rightarrow$ **images (*PNG)**: convertes a given video file into a sequences of images.

    Note: This feature is important for the correctness of the segmentation settings and for displaying segmentation results.
  - **Replace (maintaining settings)**: replaces the given video with another video (*.avi) maintaining the defined segmentation settings.
- **Video Editing Tool**: Opens the *Video Editing Tool* menu (sec. 6.2). Allows rotating and cutting the video as well as changing the brightness and contrast of the video. The edited video can saved as a new video file afterwards.
- **Type**: allows picking the type of the segmentation (sec. 4.1.4 - C22).
- **Fundamental frequency settings**: gives access to the settings used for the cycles detection and thereby fundamental frequency determination (sec. 2.2.2).

4.1.2 Region of Interest selection

The *Region of Interest selection* is opened after selecting a new video for segmentation and an initial region of interest is predicted using a neural network. It provides the following controls (fig. 13:

ROI1. Draw the region of interest (red rectangle) with holding the left mouse button down and moving the mouse. **Note:** this selection cannot be changed after the segmentation, i.e., ensure that all visibile parts of the vocal folds and glottis are in the selected region.

ROI2. **Playback**: plays the current videofile.

ROI3. Allows to select a frame interval for segmentation (see sec. 4.1.4 - C23-25).

ROI4. Shows the position, width and height of the currently drawn region of interest.

ROI5. **Clear selection**: removes the currently drawn region of interest.

ROI6. **Rotation**: allows to rotate the video in 90Â° steps.

ROI7. **Brightness/Contrast settings**: allows to adjust brightness and contrast and to activate a histogram equalization display of the video. **Note:** all changes applied here are purely visual and will not affect the calculated glottal area and contour in any way.

ROI8. **Mode**: allows to switch between manual (the classic segmentation that was implemented already in the previous versions of GAT) and automatic segmentation.

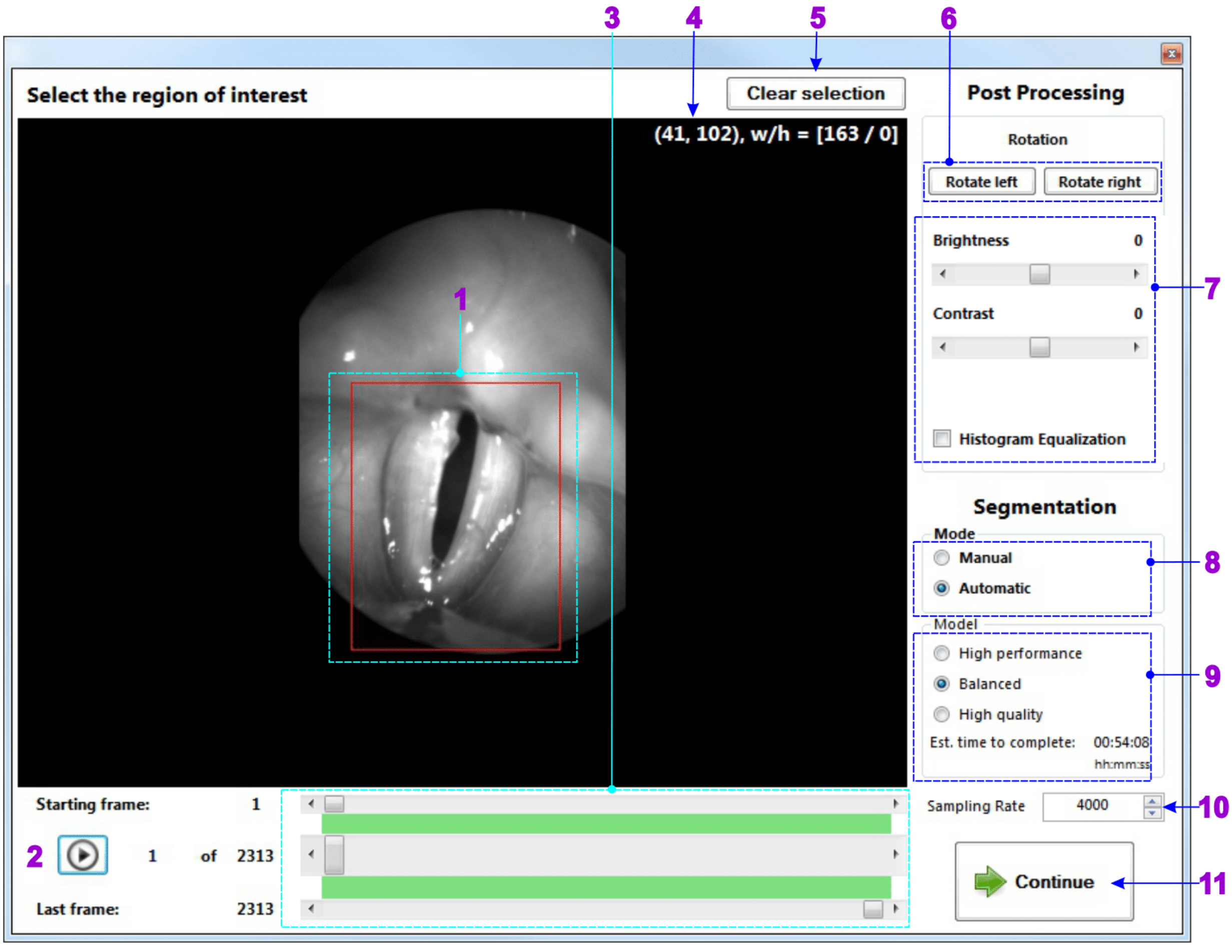


Figure 13 *– Segmentation Tools: Region of interest selection*

ROI9. **Model**: allows to switch between three types of models for the automatic segmentation. Depending on the model, the region of interest and the length of the video, the time needed to automatic segment is estimated.

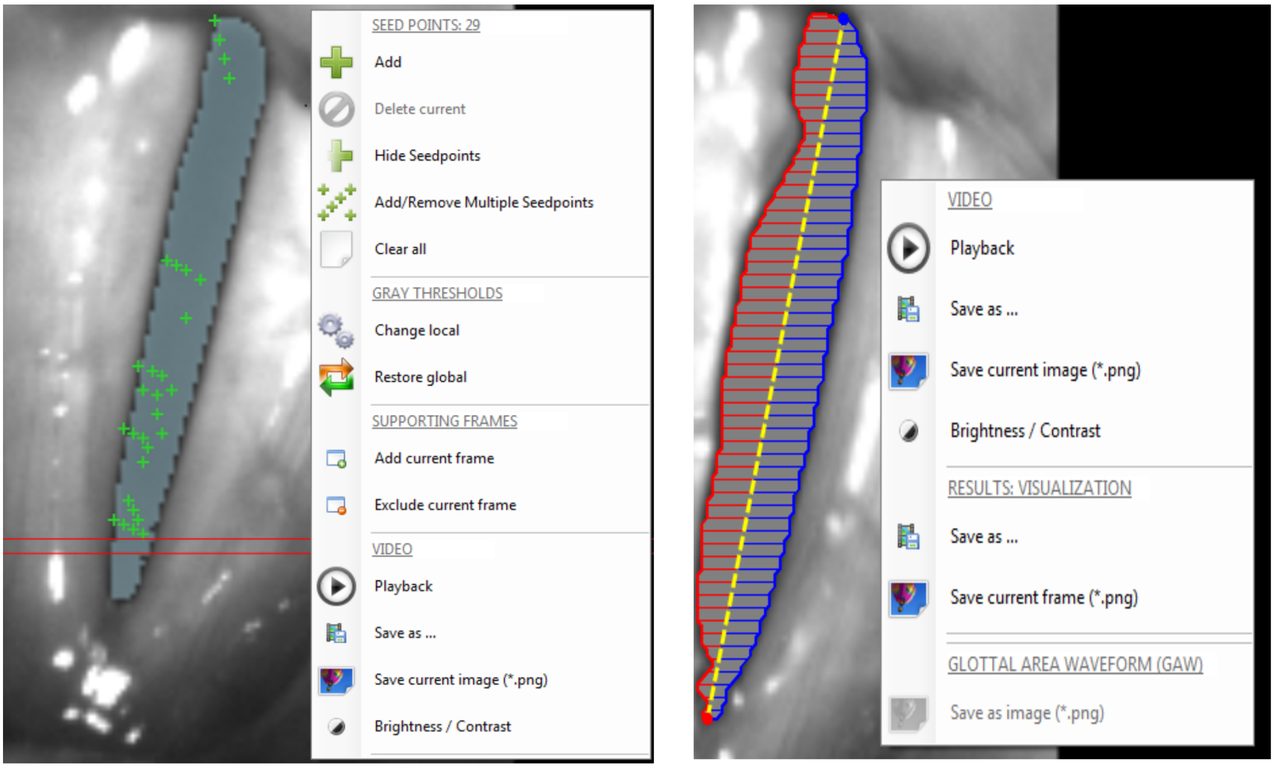
ROI10. **Sampling Rate**: select the sampling rate of the video. Default: 4000 frames per second..

ROI11. **Continue**: applies the selected Region of interest and, if selected, starts the automatic segmentation. If the **Mode** "manual" was chosen it switches to segmentation tools main form (fig. 16).

4.1.3 Context menus

Context menu: Segmentation settings This context menu (fig. 14a) is only available before the segmentation is completed.

- SEED POINTS (sec. 4.1.4 - C7)
  - **Add**: adds a new seed point.
  - **Delete current**: deletes the current seed point.
  - **Hide Seed Points**: hides the seed points but not the selected area.
  - **Add/Remove Multiple Seed Points**: switches between adding (default) and removing meshes of seed points. multiple seed points can be added/removed by holding the left mouse button and drawing a rectangle which is then filled with seed points/all seed points within it are deleted.
  - **Clear all**: delete all seed points.
- GRAY THRESHOLDS (sec. 4.1.4 - C6)



(a) **(b)**

Figure 14 *– Segmentation Tools: Contextmenus. (a) Segmentation settings, (b) Results settings*

- **Change local**: enables refinement of the gray threshold values opening a menu for detailed adjustments to the Threshold values.
  Note: After applying the changes and closing the menu the current frame has also to be added to the supporting frames (sec. 4.1.4 - C9), since otherwise the changes will be lost as soon as another video frame is selected.

- **Restore global**: restores the gray threshold values from the supporting frames.
  Note: The supporting frames are frames with defined gray thresholds: the gray thresholds between supporting frames will be interpolated linearly. Therefore generated gray thresholds are called the *Global thresholds* and in dark violet color (sec. 4.1.4 - C8).

• <u>SUPPORTING FRAMES</u> (sec. 4.1.4 - C9)

- **Add current frame** (sec. 4.1.4 - C10)

- **Exclude current frame** (sec. 4.1.4 - C11)

• <u>VIDEO</u>

- **Playback** (sec. 4.1.4 - C20).

- **Save as ...**: saves the given raw images as a movie or a sequence of images (fig. 15a)

- **Save current image (*.png)**: saves the current raw image as a PNG-file (*.png).

- **Brightness / Contrast**: allows changing the brightness and contrast of the images by opening the in (fig. 31) depicted menu.

  IMPORTANT: Changes made this way will **not** affect the segmentation, only the visualization of the video. However, if the video itself is changed and reloaded afterwards, it will affect the segmentation.

Contextmenu: Results options This contextmenu (fig. 14b) is only available after the segmentation.

- VIDEO
  - **Playback** (sec. 4.1.4 - C20).
  - **Save as ...**: saves the images with the result of the segmentation as a movie or a sequence of images (fig. 15a)
  - **Save current image (*.png)**: saves the current image with the results of the segmentation as a PNG-file (*.png).
  - **Brightness** / **Contrast**: allows changing the brightness and contrast of the images (fig. 31).
- RESULTS: VISUALISATION
  - **Save as ...**: saves the images with the result of the segmentation as a movie or a sequence of images (fig. 15a) (including markers for the midline plus left and right contour)
  - **Save current image (*.png)**: saves the current image with the results of the segmentation as a PNG-file (*.png). (including markers for the midline plus left and right contour)
- GLOTTAL AREA WAVEFORM (GAW)
  - **Save as image (*.png)** saves the GAW of this segmentation as a PNG-file (*.png)..

4.1.4 Controls

The *Segmentation Tools* provide the following controls (fig. 16,17):

C1. Folder of current videofile.

C2. **Source tab**: Shows the current video frame without thresholds and selected area

C3. **Segmentation Settings tab**: Allows adjusting the segmentation settings

C4. **Results tab**: Shows the detected cycles and midline with current segmentation settings (will only be automatically selected after successful segmentation).

C5. **Set workspace** (sec. 2.2.1).

C6. **Red lines**: mark the vertical positions of the gray thresholds required for the detection of the segmented glottis contour using *Region-Growing*. Gray values between the lines will be interpolated linearly. The positions of these lines can be changed by pulling them vertically or using the context menu **Segmentation settings/options** (fig. 14a) GRAY THRESHOLDS - Change local).
(2 lines under (max. opening) and 2 lines above the glottis (max. opening) help to "crop" the image.)

C7. **Seed points**: are marked as green crosses. These points are required for the detection of the segmented glottis contour using *Region-Growing*. The seed points are among the global parameters. I.e. each seed points is used in the entire image sequence.
Seed points help to assess the accuracy of the gray values calibration. They can be added by double-clicking on the image or can be added and deleted using the contextmenu **Segmentation settings/options** (fig. 14a SEED POINTS: - Delete current, - Clear all)).

C8. **Gray thresholds**: indicates the global (dark violet colored) and local (current frame only, blue colored) gray thresholds. Their correctness is most important for an accurate segmentation of the given image sequence. The values of the thresholds can be changed by moving the 4 *"Marker Boxes"*.

C9. **Supporting frames**: The supporting frames are frames with defined gray thresholds settings: the gray thresholds between supporting frames will be interpolated linearly. These gray thresholds are called the *Global thresholds* and are used in the entire sequence of images.

C10. **Add**: Adds the current frame to the supporting frames: saves current local gray thresholds and updates the **Global thresholds** re-interpolating afterwards.

C11. **Exclude**: Excludes the current frame from the supporting frames: deletes the global gray thresholds of the current frame and updates the **Global thresholds** re-interpolating afterwards.

C12. **Audio visualization**: Allows the user to select between spectral and signal representation of the audio data.

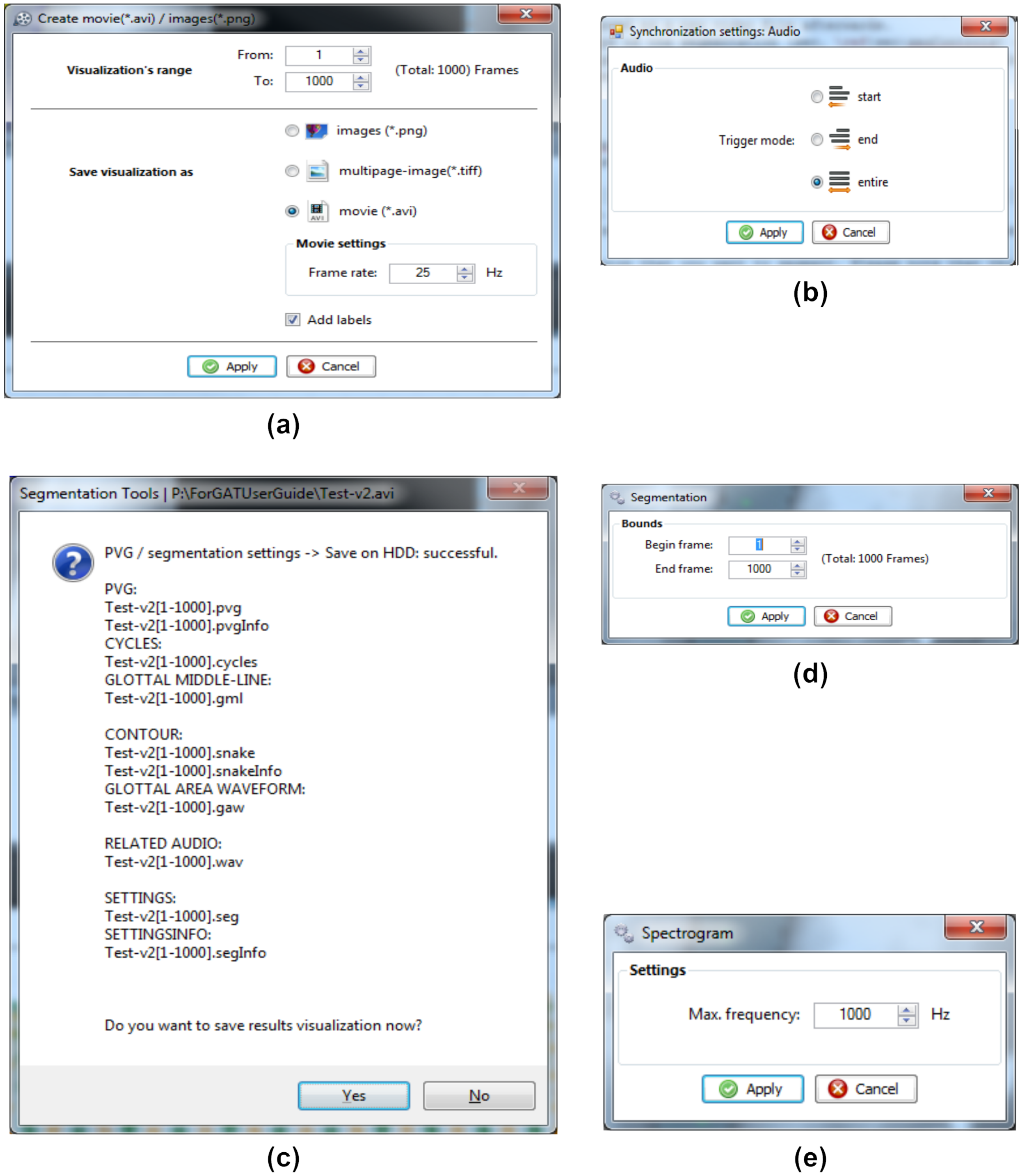


Figure 15 – *Segmentation Tools: Popup-windows. (a) Save visualization / raw images, (b) Synchronization settings audio ↔ video, (c) Save results, (d) Set segmentation bounds, and (e) Set maximum displayed frequency for spectrogram*

C13. Maximum displayed frequency in the spectrogram. This value can be changed with a mouse click (fig. 15e).

C14. **Spectrogram**: displays the spectrogram as well as the signal of corresponding audio sequence (*.wav), if it exists.

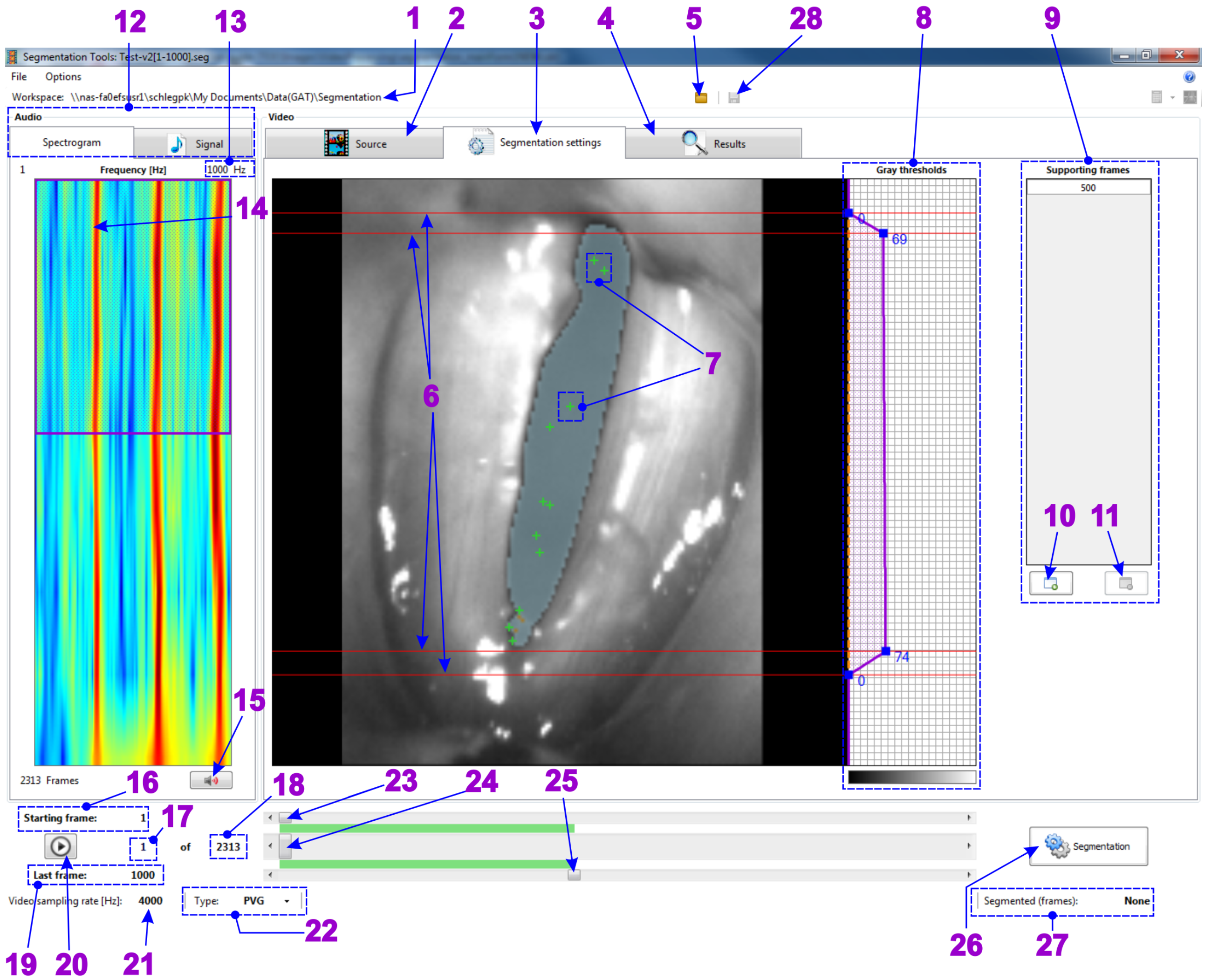


Figure 16 – *Segmentation Tools: Main form - Tabpage* "Segmentation"

Note: The user will be prompted to choose an acoustic signal corresponding to the videofile when opening a new video or loading an existing project (fig. 15b).

C15. **Audio playback**: plays the synchronized acoustic signal.

C16. First frame of the sequence to be segmented. This value can be changed by mouse click (fig. 15d).

C17. The number of the currently displayed frame.

C18. Total number of frames in the image sequence.

C19. Last frame of the sequence to be segmented. This value can be change by mouse click (fig. 15d).

C20. **Playback**: plays the current video back.

C21. **Video sampling rate**. This value can be changed with a mouse click.

C22. **Type** represents the type of segmentation.
The following types are available:

- **Contour**: The segmentation results contain only the segmented glottis contours (*.snake) and the complete glottal area waveform (*.gaw).

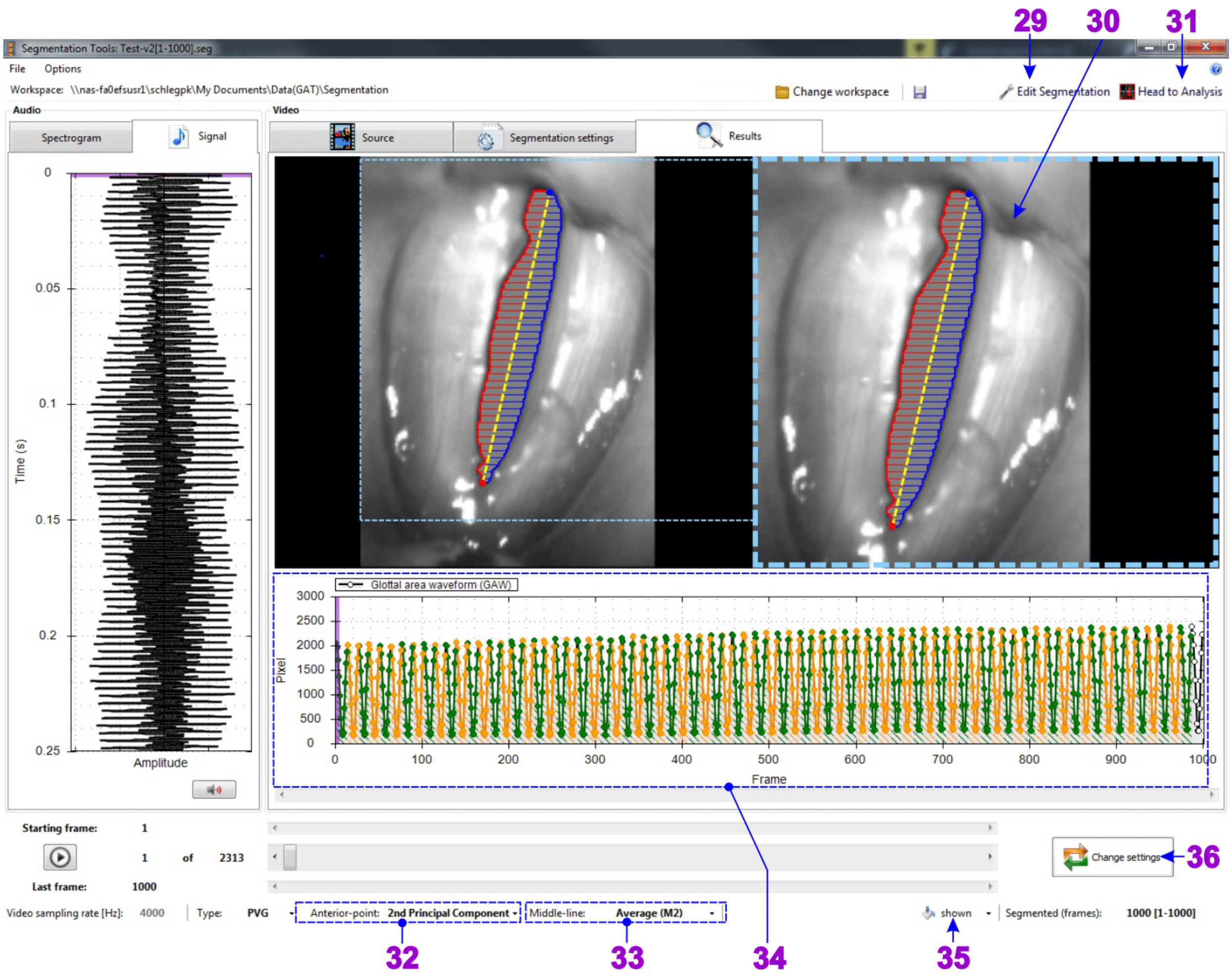


***Figure 17** – Segmentation Tools: Main form - Tabpage* "Results of segmentation"

- **PVG**: The segmentation results contain the phonovibrogram (*.pvg), the glottal main axis (*.gml), the segmented glottis contours (*.snake) and the complete as well as left and right glottal area waveform (*.gaw).

C23. Sets up the first frame of the sequence to be segmented by scrolling.

C24. Changes the current frame by scrolling.

C25. Sets up the last frame of the sequence to be segmented by scrolling.

C26. **Segmentation**: performs the segmentation with the defined global settings.

C27. **Number of segmented Frames**: Shows the numbers of the currently segmented Frames, changes after Segmentation

C28. **Save results**: saves the results of the segmentation as well as the segmentation settings (*.seg) and computes the phonovibrogram if the **Type** (sec. 4.1.4 - C22) is set to **PVG**

Note: If the **Type** (sec. 4.1.4 - C22) is set to **Contour**, the phonovibrogram will not be computed!

C29. **Edit Segmentation**: allows to edit the currently selected frame manually and opens the glottal area drawing tool described depicted in (fig:segmentMain3).

C30. Zooms in on the selected area of the current image. This area can be changed by moving the dashed line rectangle on the left side.

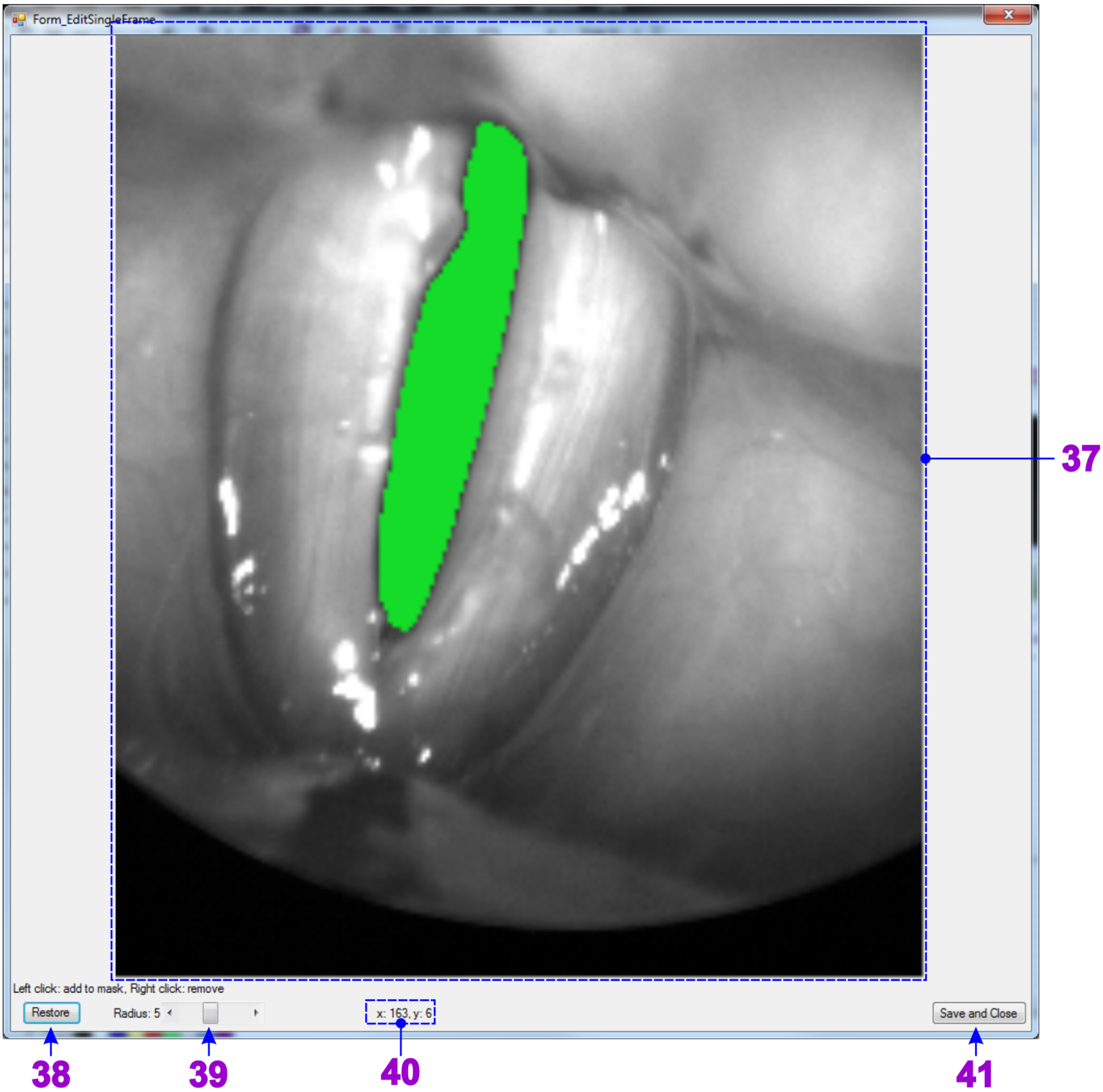


***Figure 18** – Segmentation Tools: Extra form* - "Edit Single Frame"

C31. **Run Analysis Tools (PVG)**: passes the results of the segmentation on to the *Analysis Tools (PVG)*. This option is only available if **Type** is set to **PVG** and the results of the segmentation are already saved.

C32. **Anterior-point**: allows picking the type of anterior commissure point (sec. 4.1.6).

C33. **Middle line**: allows picking the type of the glottal main axis (sec. 4.1.6).

C34. Displays the glottal area waveform calculated from the segmented glottis contours.

C35. **shown** / **hidden**: turns on / off the internal area filling of the segmented glottis contours.

C36. **Change settings**: resets the segmentation results to allow changing the settings.

Note: if the segmentation results are resetted the calculation of the glottal Area Waveform has to be performed again to get back to the segmentation results.

C37. in this area the segmentation mask for the current frame can be edited by drawing with the mouse (left click: add to segmentation mask, right click: remove from segmentation mask).

C38. **Restore**: restores the segmentation mask of the frame as it originally was after segmentation.

C39. **Radius selection**: changes the radius of the paintbrush.

C40. shows the current position of the cursor.

C41. **Save and Close**: saves the edited mask for the current frame.

4.1.5 First steps

The following steps describe how to use the *Segmentation Tools* for the first time:

Step-1. Click on the main menu buttons:

a. **Open video**, if you want to open a new video for segmentation.

b. **Load segmentation**, if you want to load an existing segmentation project.

Step-2. Set up the segmentation settings (**manual**):

a. select a region of interest (fig 13)

b. select manual segmentation mode (ROI 8)

c. click on continue (ROI 11)

d. Add seed points for the region growing inside the glottal area

e. Add supporting frames: Choose a frame, then set up the gray thresholds by pulling the red lines on the image or the "Marker Boxes" on the right side and click **Add (C10)**. To delete a supporting frame, choose the frame number in **Supporting Frames(C9)** and click **Exclude(C11)**.

f. Set the bounds of the sequence to be segmented by scrolling of **C23/C25**.

g. Click **Segmentation(C26)** to perform the segmentation.

h. Changes to the segmentation settings may be necessary to improve the segmentation results. If this is the case, click **Change settings(C36)**.

Step-3. Set up the segmentation settings (**automatic**):

a. select a region of interest (fig 13)

b. select automatic segmentation mode (ROI 8)

c. select which model you want to use (ROI 9)

d. click on continue (ROI 11)

Step-4. Click **C32** to change the anterior commissure point.

Step-5. Click **C33** to change the glottal main axis

Step-6. Click the **save icon(C28)** to save the results of the segmentation.

Step-7. Go to analysis for parameter calculation (C31)

4.1.6 Glottal main axis / Anterior commissure point: Options

Until now the anterior and posterior commissure points were assumed to be the lowest and highest points on the segmented glottis contour. This approach can be problematic in the case that the glottis is not oriented along the vertical image axis. A potential for improvement exists in detecting the glottal midline independently from the spatial orientation of the glottis contour within the frame. This can be achieved with multivariate statistics, namely the *Principal Component Analysis (PCA)*. The points on the segmented glottis contour are assumed to form a two-dimensional set of points with an orientation that can be described by two principal directions. Those principal directions can be calculated from the *PCA* (fig.19-1). The larger component is assigned to the longitudinal direction which is supposed to be the widest extension of the glottis. Respectively, the smaller component is assigned to the lateral direction which is perpendicular to the latter. This method is effectively a coordinate transformation and the origin of the principal components' coordinate system can be used as a *reference point* (fig.19-1).
These facts will be used for a dynamic computation of the glottal main axis and anterior commissure point for segmented glottis contour.

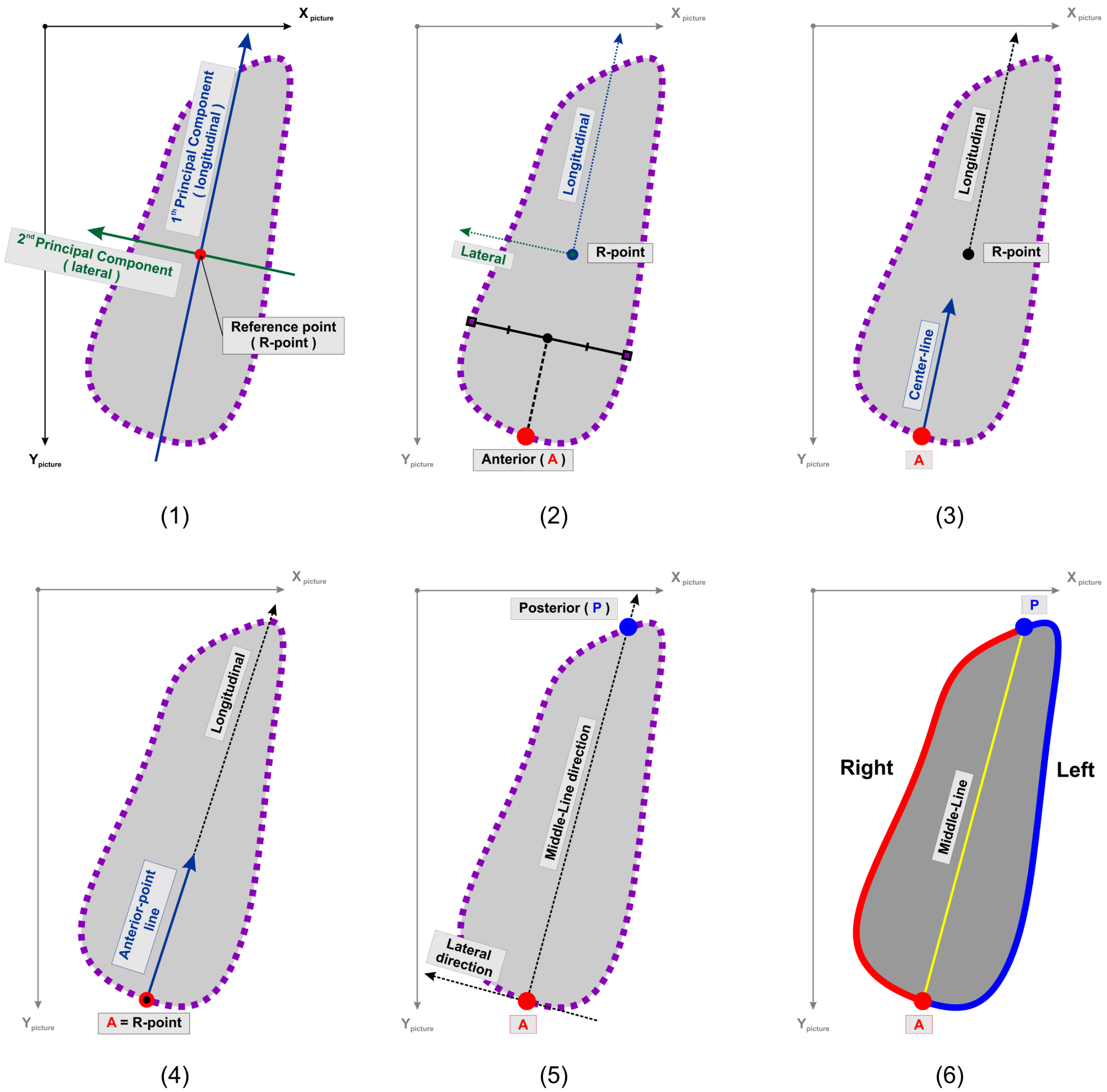


***Figure 19** – Schematic illustration of the functionality of the Principal Component Analysis applied to a glottis contour (1) and for glottal main axis detection: anterior commissure point detection (2), glottal main axis detection based on* Centroid-based (C) *(3) as well as Anterior-based (A) (4), posterior commissure point detection (5) and the combined result (6)*

Anterior commissure point In the *Segmentation Tools* the following options for the computation of the anterior commissure point are available:

- 2^{nd} **Principal component**: The anterior commissure point is determined by projecting the center of the largest lateral distance of the contour along the longitudinal axis onto the anterior region of the contour (fig. 19-2).
  Thereby, the principal components describing the longitudinal and lateral direction are calculated with respect to the *reference point* which is set to the median point of the glottis contour.
- **"Lower -Region" (10 pixel)**: The anterior commissure point is determined by projecting the center of the lateral distance (10 pixels above the lowest point) of the contour along the longitudinal axis onto the anterior region of the contour.
  Thereby, the principal components describing the longitudinal and lateral direction are calculated with respect to the *reference point* which is set to the median point of the glottis contour.

- **Lowest point**: The anterior commissure point is defined as the lowest point of the segmented glottis contour.

Glottal main axis In *Segmentation Tools* the following options for the computation of the glottal main axis are available:

- **Anterior-based (A)**: The glottal main axis direction is determined as the longitudinal component of the *PCA* with respect to the selected anterior commissure point as *reference point* (fig 19-4).
- **Average of (A) and (C)**: The glottal main axis direction is determined as mean direction between the Anterior-based(A) and Centroid-based (C) described above.
- **Centroid-based (C)**: The glottal main axis direction is determined as the longitudinal component of the *PCA* with respect to the median point of the segmented glottis contour as *reference point* (fig 19-3).
- **Anterior-Posterior Line (A - P)**: The glottal main axis is defined as connecting line between the selected anterior commissure point and highest point on the segmented glottis contour.

Posterior commissure point The posterior commissure point is determined as the intersection of the selected glottal main axis ensuing from the selected anterior commissure point and the glottis contour in the posterior region (fig.19-5). In case there is no such intersection the highest point on the segmented glottis contour is projected along the lateral direction on the selected glottal main axis ensuing from the selected anterior commissure point.

4.2 Analysis Tools (PVG)

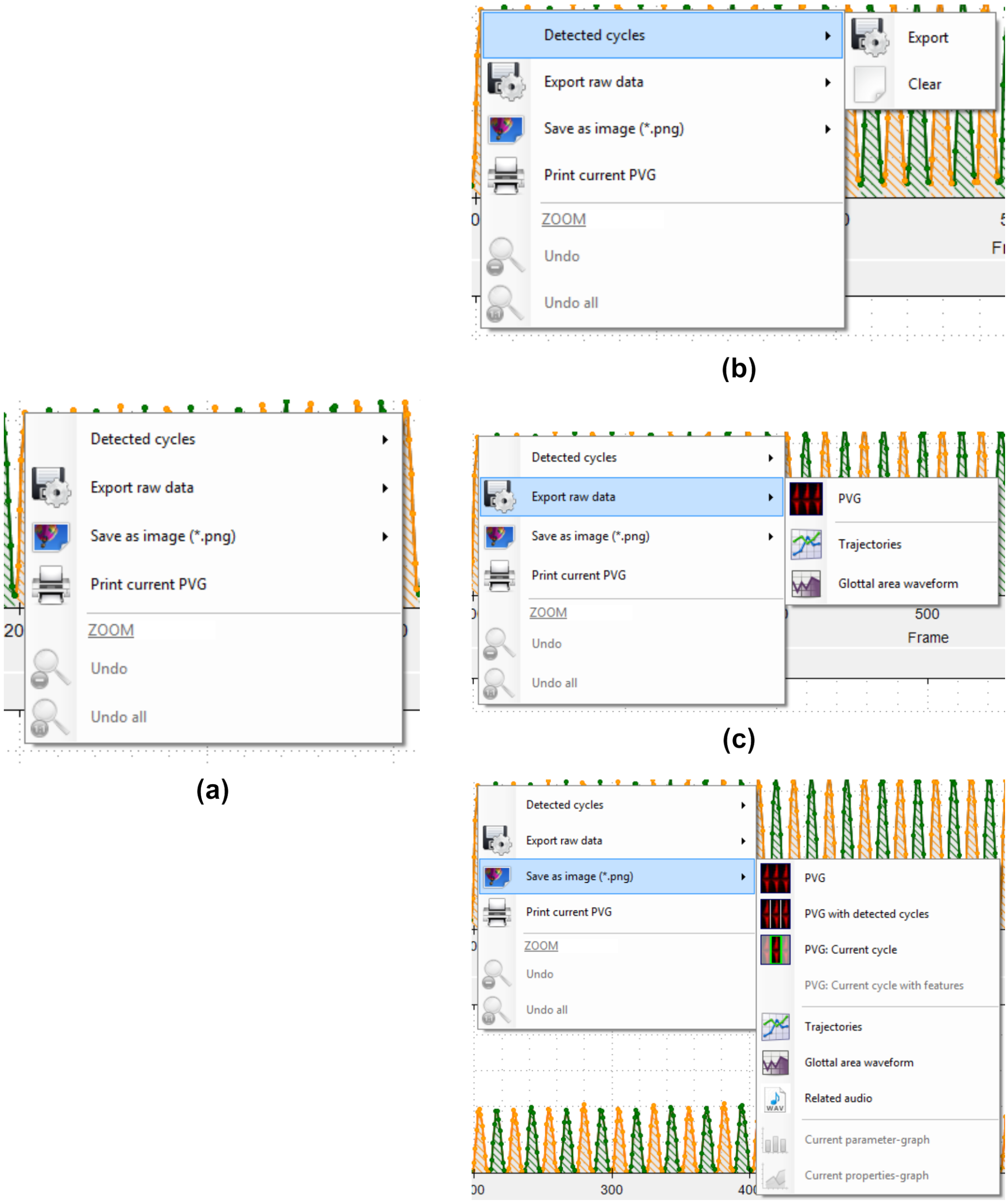
This tool allows the analysis of some characteristics of the vocal folds' vibratory motions obtained from the HSE recording after the segmentation as well as the acoustic signal synchronized to it.

4.2.1 Context menu

This context menu (fig. 20a) is enabled in all tabpages.

- **Detected cycles** (fig. 20b):
  - **Export**: saves the detected cycles to a file (*.multiCycles).

    Note: After the cycles are successfully saved, it is possible import the file later on (sec. 2.2.2) .
  - **Clear**: (sec. 4.2.2 - C2) deletes the detected cycles and if analysis was performed, clears the analysis results too
- **Export raw data** (fig. 20c): saves only specific results:
  - **PVG**: the phonovibrogram color coded in an image to a file (*.csv),
  - **Trajectories**: the current trajectories to a file (*.trajs),
  - **Glottal area waveform**: the glottal area waveform to a file (*.csv).
- **Save as image (*.png)** (fig. 20d): save an image of:
  - **PVG**: the entire phonovibrogram to a file (*.png)
  - **PVG with detected cycles**: the entire phonovibrogram with detected cycles to a file (*.png). This option is only available after the cycles detection.
  - **PVG - current cycle**: the current cycle of the PVG to a file (*.png). This option is only available after the cycles detection.
  - **PVG - current cycle with features**: the current cycle of the PVG with depicted contour angles (sec. 3.7) to a file (*.png). This option is only available if the analysis has been performed.
  - **Trajectories**: the trajectories-graph (sec. 4.2.2 - C16) to a file (*.png).



(a) (b) (c) (d)

Figure 20 – Analysis Tools(PVG): Context menu

- **Glottal area waveform**: the GAW-graph (sec. 4.2.2 - C12) to a file (*.png).
- **Related audio**: the audio-graph (sec. 4.2.2 - C18) to a file (*.png).
- **Current parameter-graph**: current parameter-graph (sec. 4.2.2 - C33) to a file (*.png). This option is only available if the analysis has been performed and **C33** is visible.
- **Current properties-graph**: current properties-graph (sec. 4.2.2 - C34) to a file (*.png). This option is only available if the analysis has been performed and **C34** is visible.

4.2.2 Controls

The *Analysis Tools (PVG)* contain the following controls (fig. 21,22,23,24, 25,26):

C1. **Load phonovibrogram**: opens the phonovibrogram file (*.pvg) for the analysis.

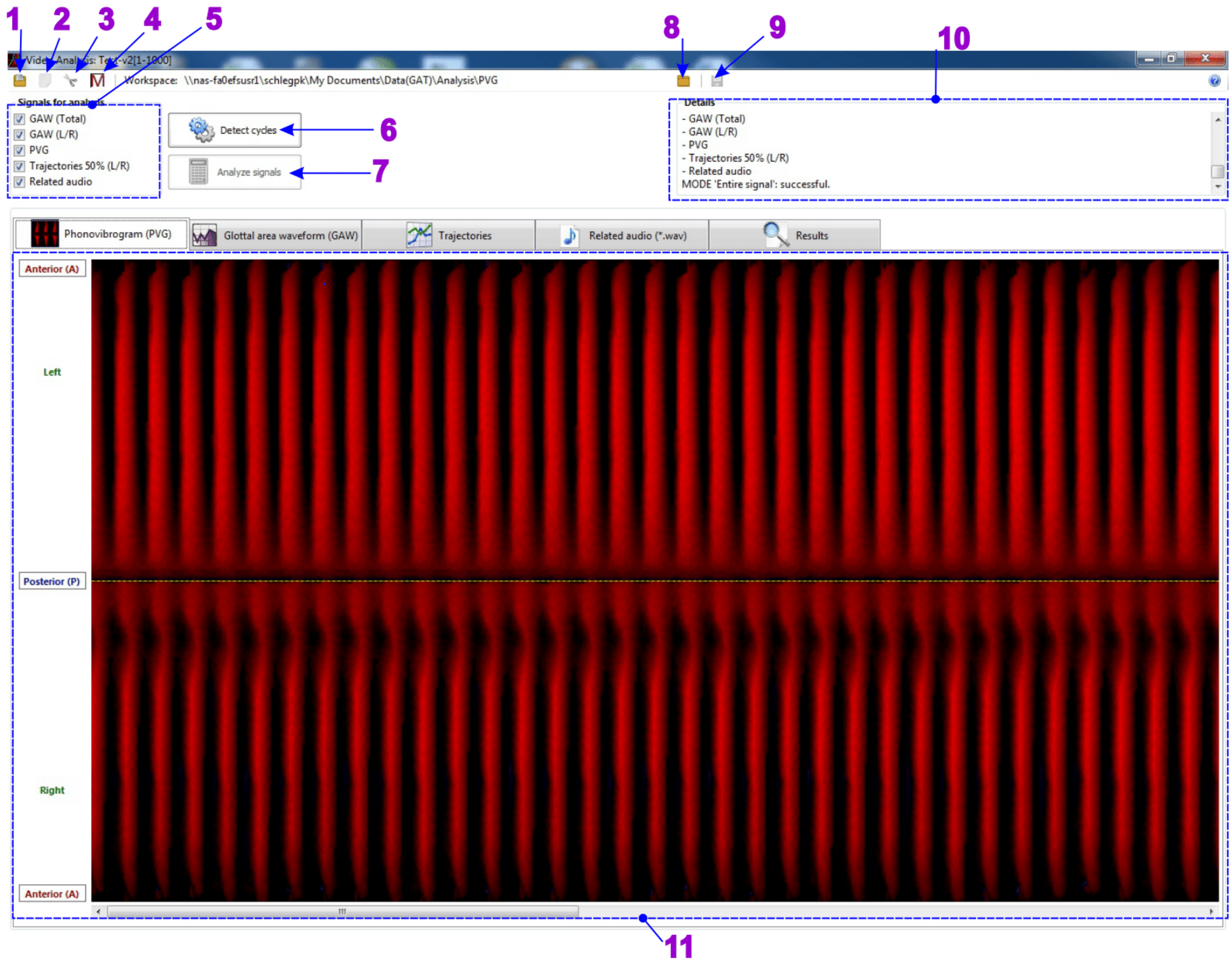


Figure 21 – *Analysis Tools (PVG): Main form - Tabpage* "Phonovibrogram (PVG)"

C2. **Clear**: deletes the detected cycles and if the analysis was performed, clears the analysis results too. This button is only available after the cycles detection.

C3. **Edit PVG / GAW**: Opens a menue for filtering and frame selection of the PVG and GAW signals for upcoming analysis (fig. 25)

C4. **Export to Matlab workspace file (*mat)**: Exports GAW and trajectory data plus information about video sampling rate and frames.

C5. **Signal for analysis**: represents all synchronized signals that could be loaded. It allows selecting the signals for cycles detection and add-on analysis, respectively. The analysis of the signal is only possible if the cycles for this signal are successfully detected.

C6. **Detect cycles**: performs the cycles detection for the selected signals using the defined settings (sec. 2.2.2).

C7. **Analyze signals**: calculate the selected parameters with the choosen analysis settings (sec. 2.2.3). This button is only available after cycles detection.

C8. **Set workspace** (sec. 2.2.1).

C9. **Save results**: saves the results of the analysis to files in the selected formats (sec. 2.2.3)

C10. **Details**: represents the detailed process report.

C11. **PVG**: represents the current phonovibrogram imaging.

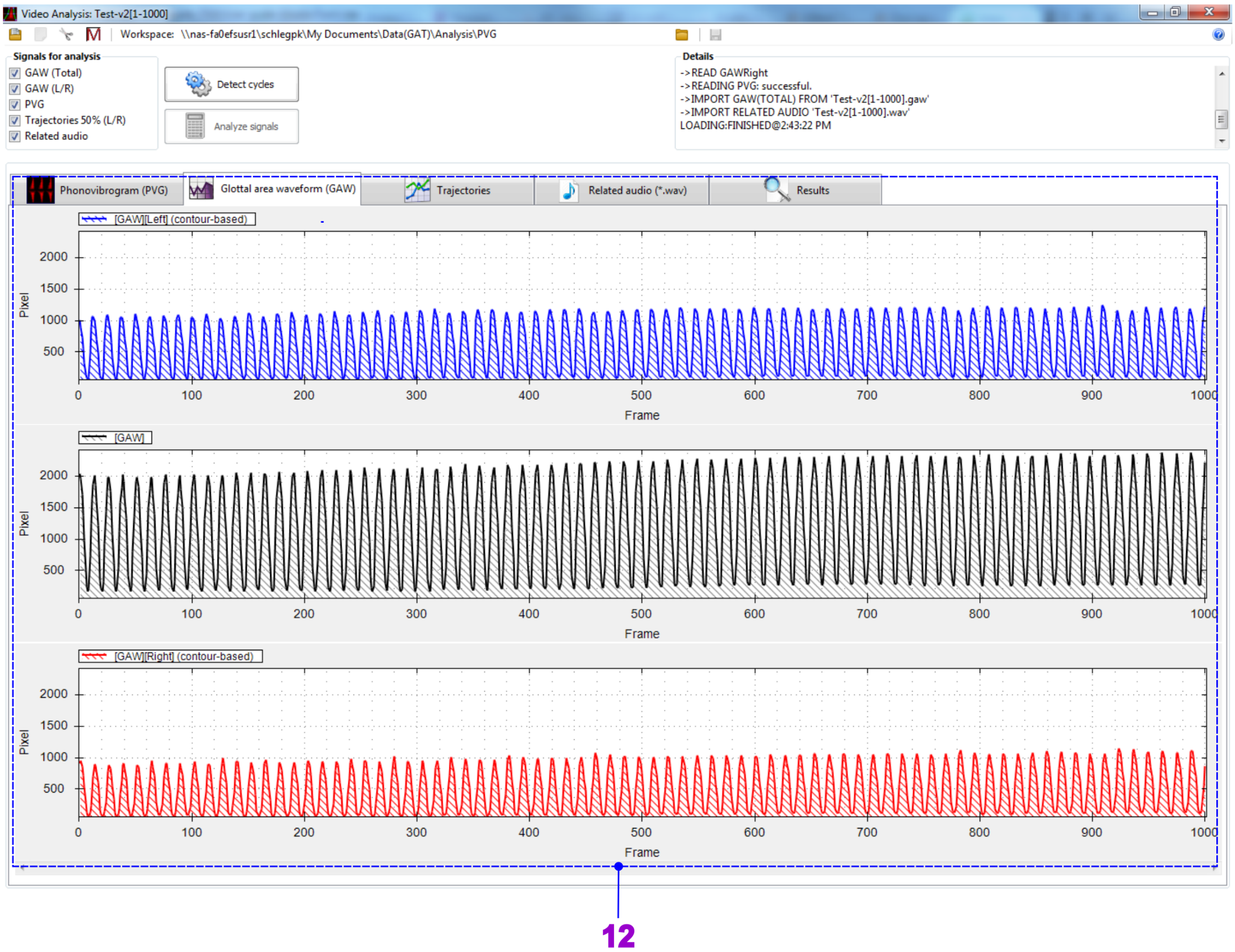


Figure 22 – *Analysis Tools (PVG): Main form - Tabpage* "Glottal area waveform (GAW)"

C12. **GAW-graph**: displays the glottal area waveforms (total, left and right respectively). **Important note**: since the total GAW is calculated by counting all segmented pixels per frame but the left and right GAWs are calculated by using the contour the total GAW is not exactly the sum of the partial GAWs. However, if no .gaw file is found, the total GAW is approximated as the sum of the partial GAWs. In such cases a warning message box is shown.

C13. **Trajectory position**: represents the percentage value of the current trajectorie's position on the Posterior-Anterior-line with respect to the posterior commissure point and its position graphically. The position of the trajectories can be changed by changing the percentage values or on the screen for the glottis reconstruction (**C12**) by moving the line representing the current position graphically. However, **changes can only be applied before cycle detection.**

C14. Displays the segmented glottis contours reconstruction as well as the current positioning of the trajectories.

C15. Allows scrolling through the frames.

C16. **Trajectories-graph**: displays the single values of the left as well as right selected trajectories over time.

C17. Plays an animation of the glottal contour.

C18. **Audio-graph**: displays the synchronized acoustic signal if one was provided.

C19. Maximum displayed frequency for spectrogram. This value can be changed with a mouse click (fig. 15e).

C20. **Spectrogram**: displays the spectrogram of the acoustic signal.

C21. **Audio playback**: plays the synchronized acoustic signal.

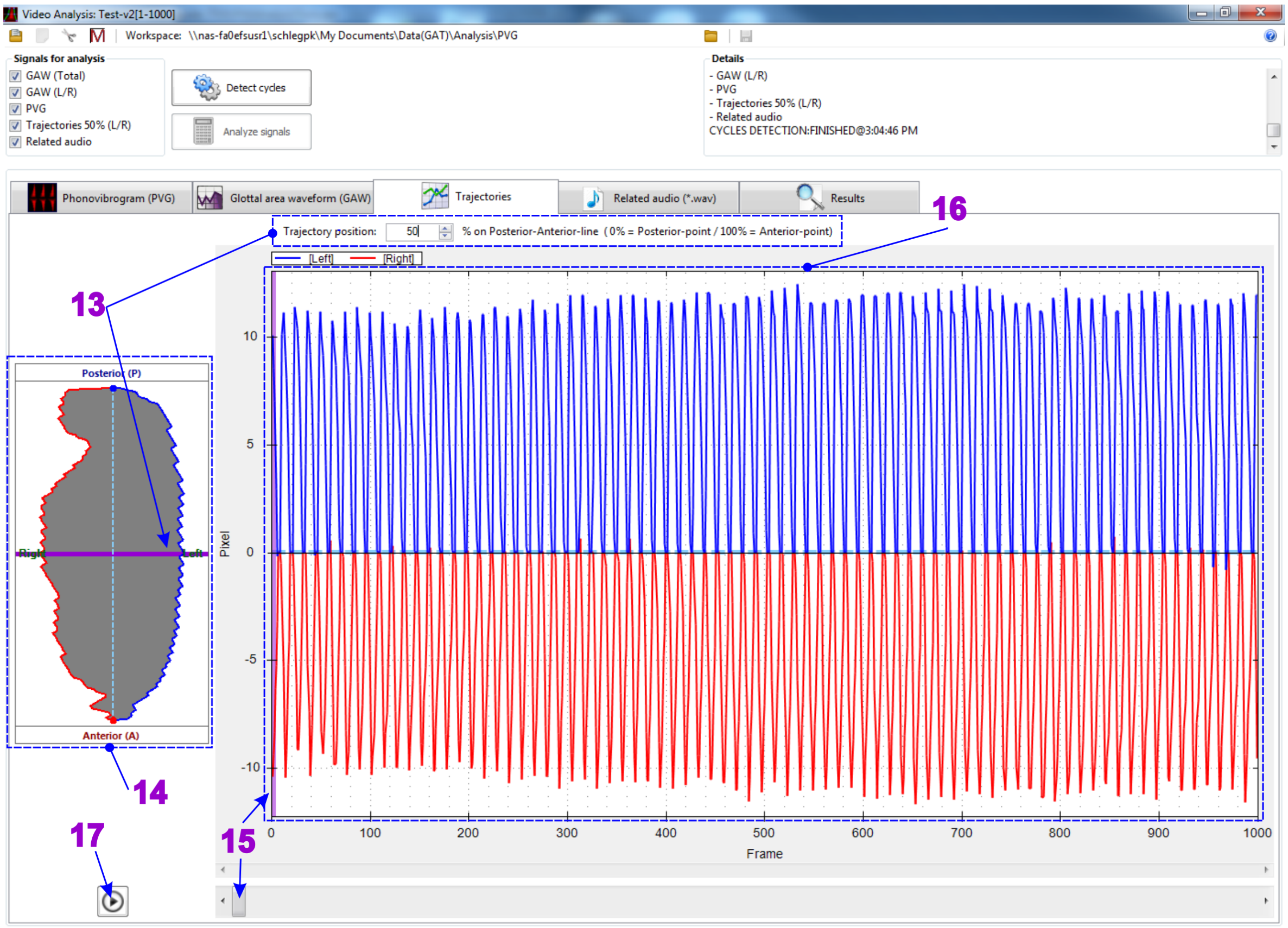


Figure 23 – *Analysis Tools (PVG): Main form - Tabpage* "Trajectories"

C22. Allows the selection of a range of frames.

C23. **Filter selection**: a dropdown menu which allows the selection of one filter for data preprocessing.

C24. Only visible if a filter is selected. Allows to adjust the filter by changing the filter order or scale factor.

C25. **Preview button**: calculates the influence the currently selected filter would have on the data, when clicked.

C26. **Preview graphs**: shows a preview of the influence the currently selected filter would have on the data.

C27. menu for applying the changes by preprocessing or exiting without changing the data for analysis.

Note: The original data is not changed and can easily be restored by opening the Edit menu again, selecting "none" in the filter selection and applying the changes.

C28. Current cycle of the PVG (maximum based detection was chosen).

C29. Controls for the navigation through the detected and analyzed PVG and total GAW cycles.

C30. Switches between the previosly selected analysis modes (sec. 2.2.3).

C31. **Parameter-graph**: dispays the values of the currently selected parameter. This graph is only shown if the selected parameter has more than one value (e.g. cycles-based parameters).

C32. the calculated thresholds of the pvg (only visible if Phonovibrogram measures were selected for parameter calculation) which are used for the PVG angle estimation.

C33. Represents the summarized values of the calculated parameters.

C34. Controls for the navigation through the detected and analyzed audio, partial GAW and Trajectory cycles.

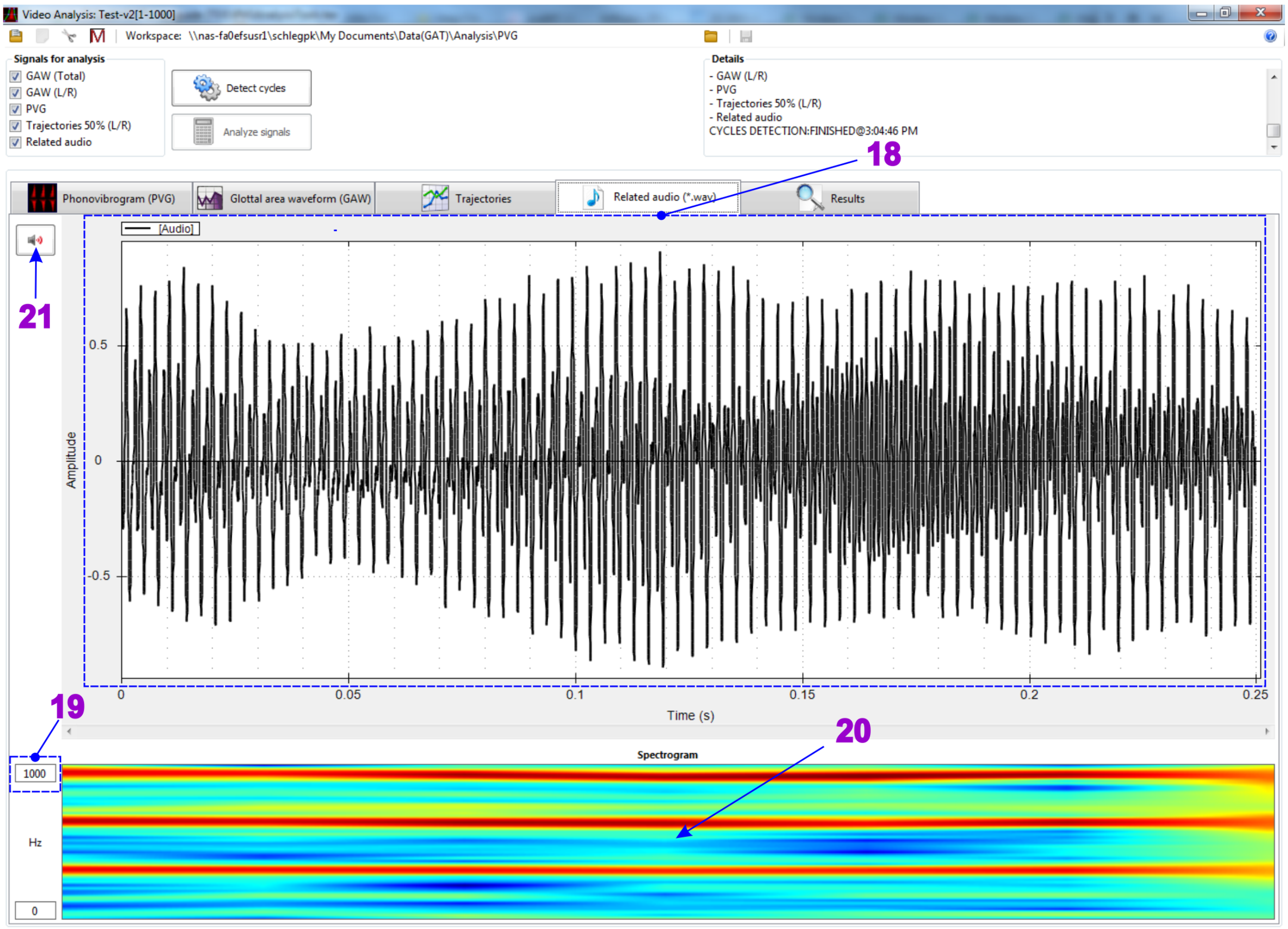


Figure 24 – Analysis Tools (PVG): Main form - Tabpage "Related audio (*.wav)"

C35. Displays the properties related to the currently selected parameter. This graph is only visible, if any properties exist.

4.2.3 First steps

The following steps describe how to use the *Analysis Tools (PVG)* for the first time:

Step-1. Click the **load icon (C1)** to load the signal.

Step-2. In **the selection table(C5)** pick the type of signal that you want to analyze.

Step-3. Click **detect cycles(C6)** and set up the settings for the cycles detection and video sampling rate.

Step-4. Click **analyze signals(C7)** and select the part of cycle sequence to be analyzed and the parameters, which you want to calculate for this sequence.

Step-5. Click the **save icon(C9)** to save the results of the analysis.

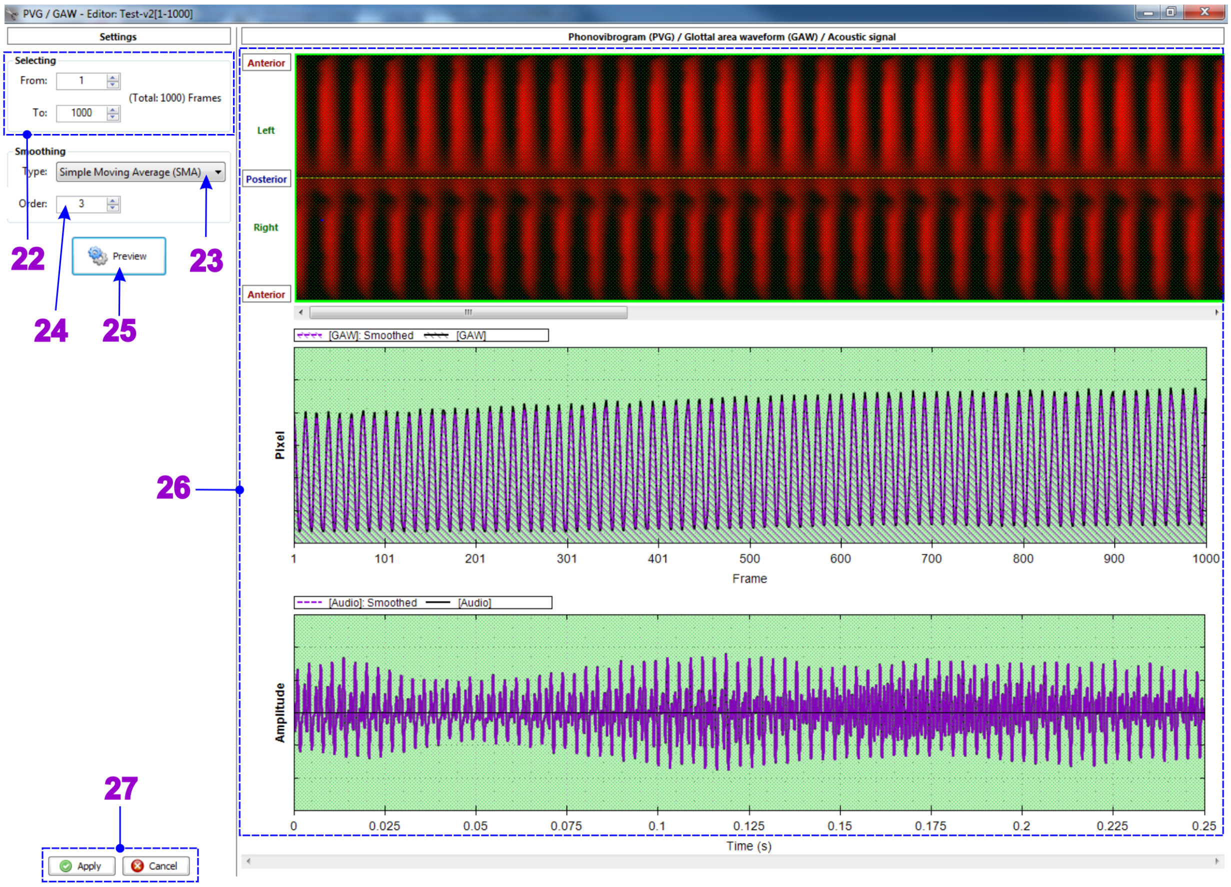


Figure 25 – *Analysis Tools (PVG): Main form - Submenu* "Edit PVG / GAW"

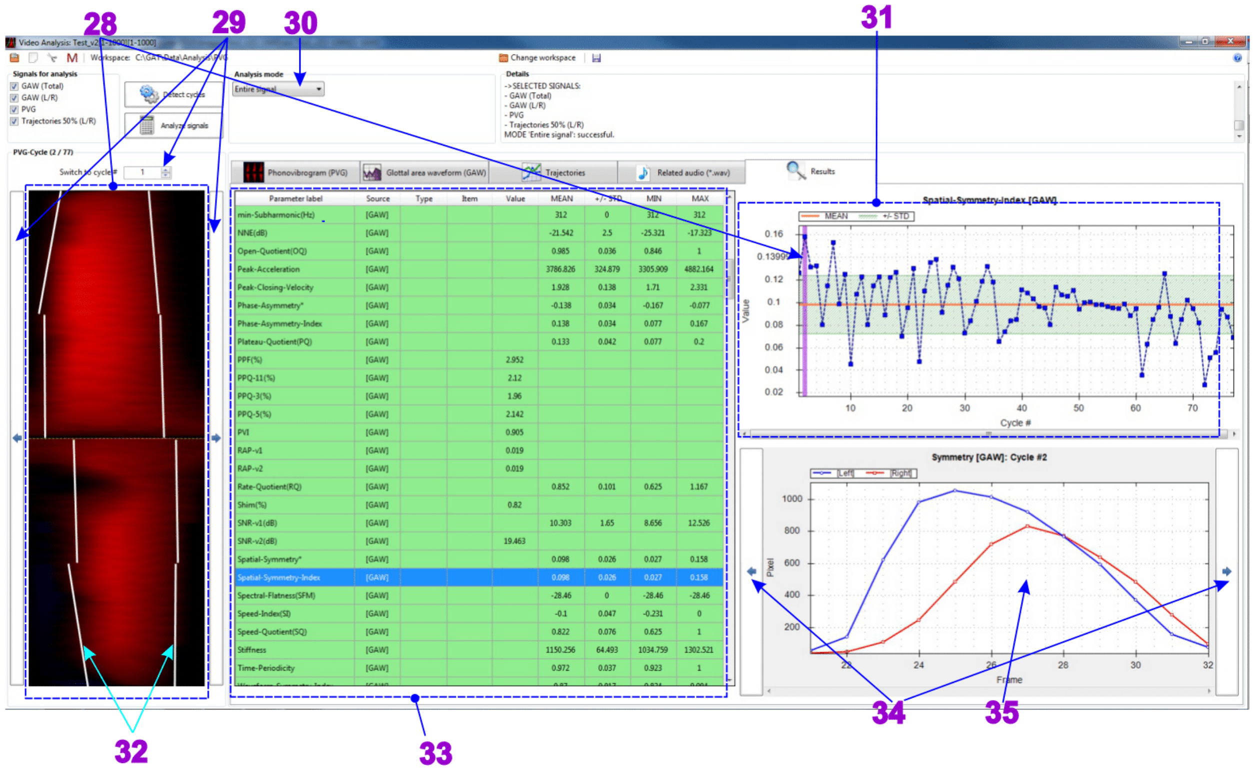


Figure 26 – *Analysis Tools (PVG): Main form - Tabpage* "Results of analysis"

5 Audio Processing

Audio Processing provides the following features:

- **Analysis Tools (Audio)** (sec. 5.1): performs the analysis of an acoustic signal without a corresponding HSE recording.

5.1 Analysis Tools (Audio)

This tool allows the analysis of audio files(*.wav) without a corresponding HSE recording.

5.1.1 Controls

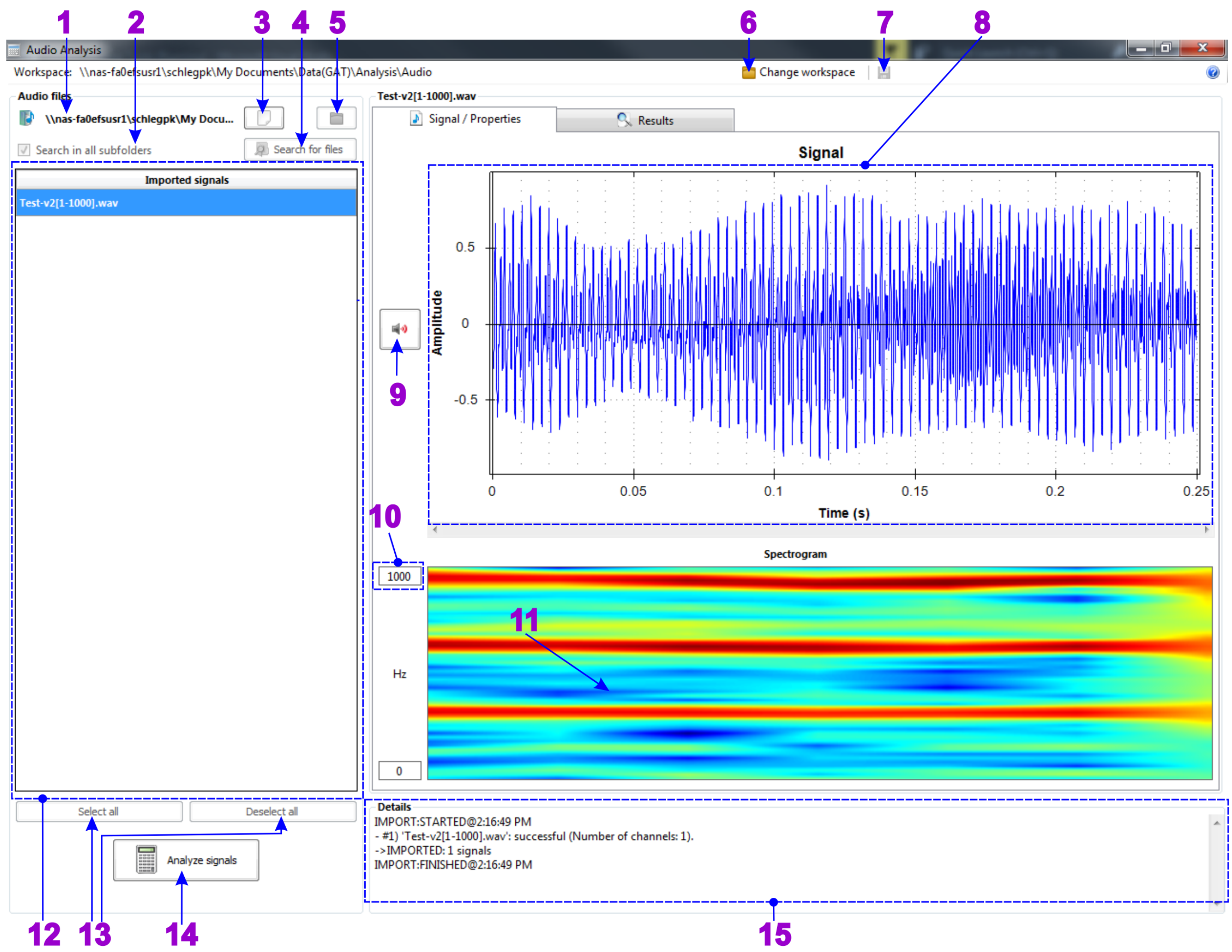


Figure 27 *– Analysis Tools (Audio): Main form - Tabpage* "Properties"

The *Analysis Tools (PVG)* provide the following controls (fig. 27,28):

C1. **Source folder**: displays the folder, where the audio files (*.wav) are stored.

C2. Checkbox to enable the automatic search subfolders.

C3. **Clear**: rejects the imported files and if analysis was performed, clears the analysis results too. This button is only available after files were imported.

C4. **Search for files**: browses the source folder for audio files.

C5. **Change source folder**: changes the source folder if desired.

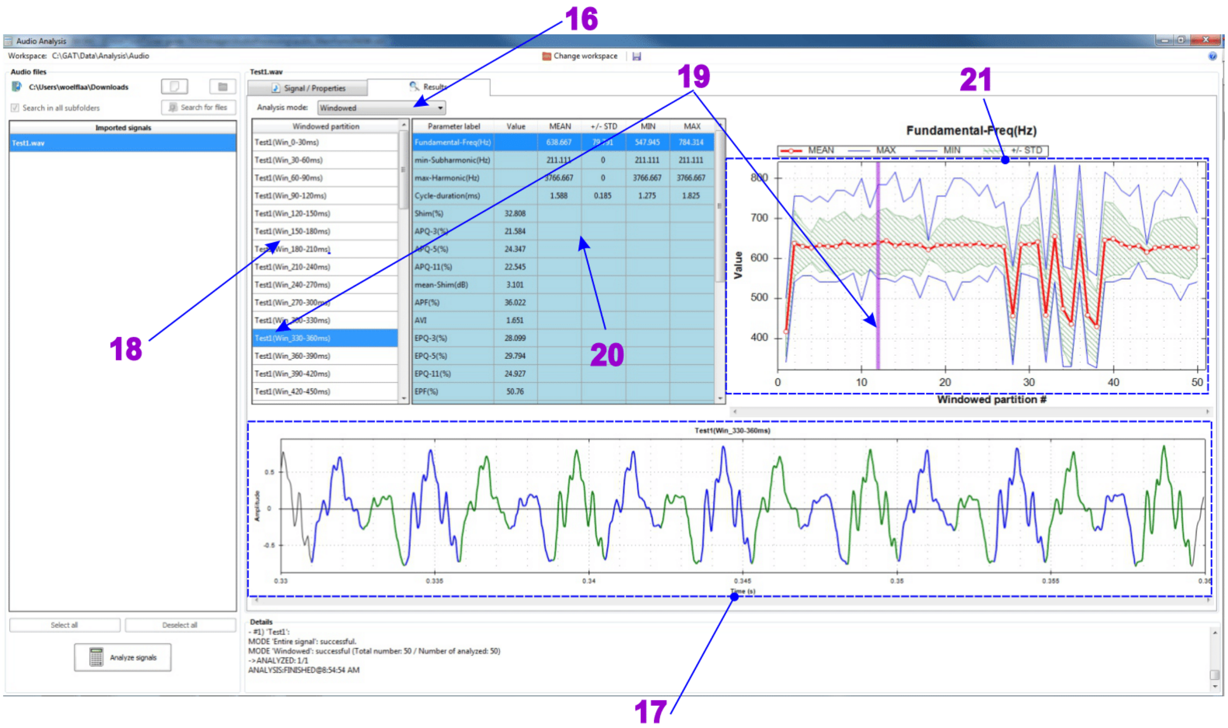


Figure 28 *– Analysis Tools (Audio): Main form - Tabpage* "Results of analysis"

C6. **Change workspace** (sec. 2.2.1).

C7. **Save results**: saves the results of the analysis to files in the selected formats (sec. 2.2.3)

C8. **Signal-graph**: displays the current signal.

C9. **Audio playback**: plays the selected acoustic signal.

C10. Maximum displayed frequency in the spectrogram. This value can be changed with a mouse click (fig. 15e).

C11. **Spectrogram**: displays the spectrogram of the current signal. If analysis is performed, it displays the distribution of the fundamental frequency and its harmonics over time.

C12. **Containing files** / **imported signals**: lists all files in the source folder and all successfully imported signals, respectively.

C13. **Select all** / **Deselect all buttons**: selects or deselects all files diplayed in the Containing files / imported signals box.

C14. **Import files** / **Analyze signals**: performs the import of the selected audio files (*.wav) and afterwards the full analysis of the selected files. The full analysis includes not only the calculation of the selected parameters with the choosen analysis mode (sec. 2.2.3), but also the cycles detection for the selected signals using the selected settings (sec. 2.2.2)

Note: During import each single channel of an audio file (*.wav) will be separately imported, if the given signal contains more than one channel.

C15. **Details**: displays the detailed process report.

C16. Switches between the previosly selected analysis modes (sec. 2.2.3).

C17. **Partition-graph**: displays the currently selected partition of the signal, if analysis mode **Windowed** was chosen (sec. 2.2.3).

C18. Displays all partitions of the signal, if analysis mode **Windowed** was chosen (sec. 2.2.3).

C19. Displays the currently selected partition on the parameter-graph.

C20. Displays the summarized values of the calculated parameters for the entire signal or the currently selected partition.

C21. **Parameter-graph**: displays the values of the currently selected parameter. If analysis mode **Windowed** was selected, this graph displays the distribution of the summarized values of the current parameter in all partitions.

5.1.2 First steps

The following steps describe how to use the *Analysis Tools (Audio)* for the first time:

Step-1. Click **change source folder(C1)** to choose the folder in which the audio files are stored.

Step-2. Click **search for audio files(C4)** to search for audio files (*.wav) in the selected source folder and import the desired files for further analysis.

Step-3. In the **files browser(C12)** check the acoustics signals you want to import. Use the right mouse button in the the **files browser(C12)** to deselect signals, which you do not want to analyze!

Step-4. Click **analyze signal(C14)** and set up the settings for the cycles detection. Furthermore, select the analysis mode and the parameters to be calculated.

Step-5. Click the **save icon (C7)** to save the results of the analysis.

6 Common Tools

Common Tools provides the following features:

- **Results Collector** (sec. 6.1): allows collecting computed analysis results.
- **Video Editing Tool** (sec. 6.2): allows editing of videos and images.

6.1 Results Collector

This tool collects the results of analyses done with the *Analysis Tools (PVG)* and *Analysis Tools (Audio)*, respectively. The program supports the following file formats:

- **XML-File** (*.xml),
- **CSV-File** (*.csv).

6.1.1 Controls

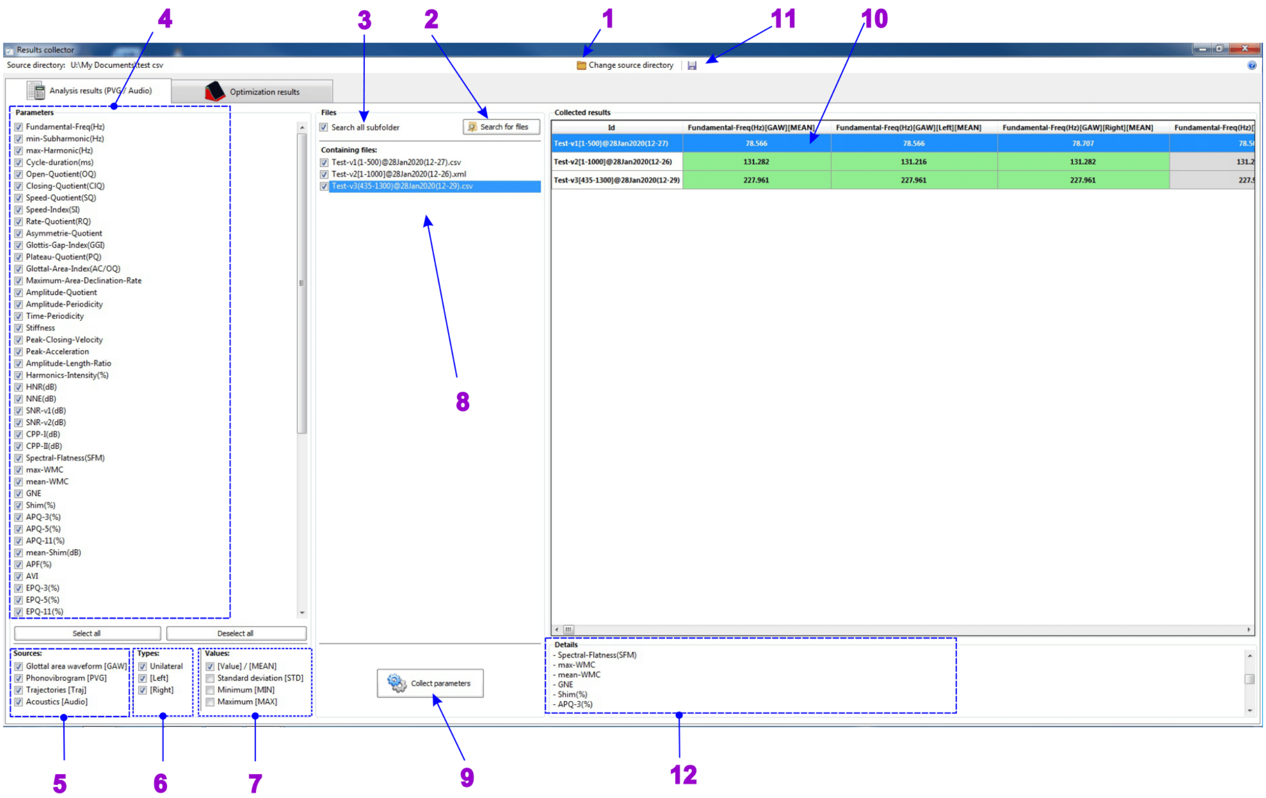


Figure 29 *– Results Collector: Main form*

The *Results Collector* provides the following controls (fig. 29):

C1. **Change source folder**: changes the source folder if desired.

C2. **Search for files**: browses the source folder for the desired files.

C3. Checkbox to enable the automatic search in all subfolders of the source folder.

C4. **Parameter browser**: lists all parameters that can be collected.

C5. **Source browser**: lists all sources of generated data, for which the parameters can be collected.

C6. **Type browser**: lists the different parameter types (total, left or right-sided) which should be collected.

C7. **Value browser**: lists all statistical values that can be computed.

C8. **File browser**: lists all acceptable files in the source folder.

C9. **Collect parameters**: performs the evaluation of the parameters using the previously selected options.

C10. **Results view**: displays the results of the successful evaluation.

C11. **Save collected results**: saves the results of the evaluation to a file.

C12. **Details**: displays the detailed process report.

6.1.2 First steps

The following steps describe how to use the *Results Collector* for the first time:

Step-1. Click the **open icon(C1)** to choose the source folder, where these files are stored.

Step-2. Click the **search for files(C2)** button to search for files to evaluate in the provided source folder.

Step-3. Check in the **parameter browser(C4)** the parameters for the evaluation.

Step-4. Check in the **source browser(C5)** the data sources for which you want to evaluate the parameters.

Step-5. Check in the **type browser(C6)** the parameter types which should be used.

Step-6. Check in the **values browser(C7)** the statistical evaluations you want to perform on the data.

Step-7. Check in the **file browser(C8)** the files containing the selected parameters.

Step-8. Press the **collect parameters(C9)** button to run the evaluation. After successfully collecting the selected parameters the overall results will be displayed in the **results view(C10)**

Step-9. Click the **save icon(C11)** to save the collected results.

6.2 Video Editing Tool

This tool allows changing of brightness and contrast of videos as well as rotating videos. The program supports the following file formats:

- **Videofile** (*.avi),
- **Tagged Image File Format** (*.tiff),
- **Portable Network Graphic** (*.png).

6.2.1 Controls

The *Video Editing Tool* provide the following controls (fig. 30):

C1. Shows the path of your loaded video.

C2. **Search for files**: Browses the source folder for the desired files.

C3. **Source window**: shows the unedited source file.

C4. **Result window**: shows the edited file.

C5. **Frame selection scroll bars**: select the minimum and maximum frame of the video and scroll through it, similar to (sec. 4.1.4 - C23-25).

C6. Shows the file format of your source file.

C7. **Video rotation**: allows to rotate your video in 90°steps.

C8. **Brightness/Contrast adaption**: opens the Brightness/Contrast adaption menu (fig. 31). It allows changing the brightness and contrast of the images.

C9. Shows the start and end numbers of the converted frames after saving of the converted video

C10. **Format selection**: allows to select a different video format for your edited video file.

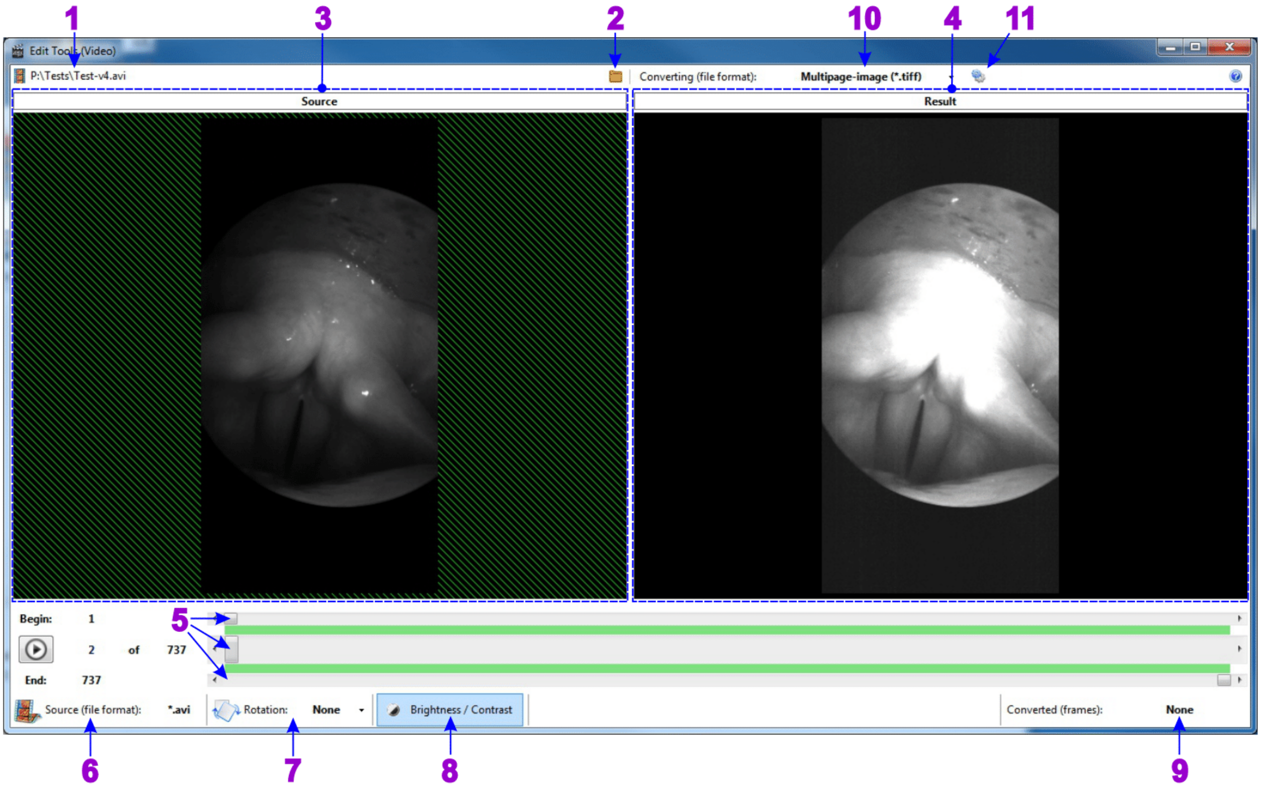


Figure 30 *– Video Editing Tool: Main form*

C11. **Convert**: Performs cutting and converting of the video as it was set in the *Video Editing Tool* menu. To the file name the number of selected frames are added per default to avoid accidentally overwriting your source video.
Note: Even if unedited the memory space the result video consumes can increase significantly. This is due to the high degree of precision which is used for saving the data.

6.2.2 First steps

The following steps describe how to use the *Video Editing Tool* for the first time:

Step-1. Click the **search for files (C2)** button to select the video file you want to edit.

Step-2. Select the range of frames you want to convert using the **frame selection scroll bars (C5)**.

Step-3. Click on the **video rotation (C7)** menu if you want to rotate your video.

Step-4. Click on the **Brightness**/**Contrast adaption (C8)** button to open the menu depicted in (fig. 31) and adjust the brightness and contrast of your video.

Step-5. Select your desired target format using the **Format selection (C10)** menu.

Step-6. Click on the **Convert (C11)** button to save your edited video as a new video file.

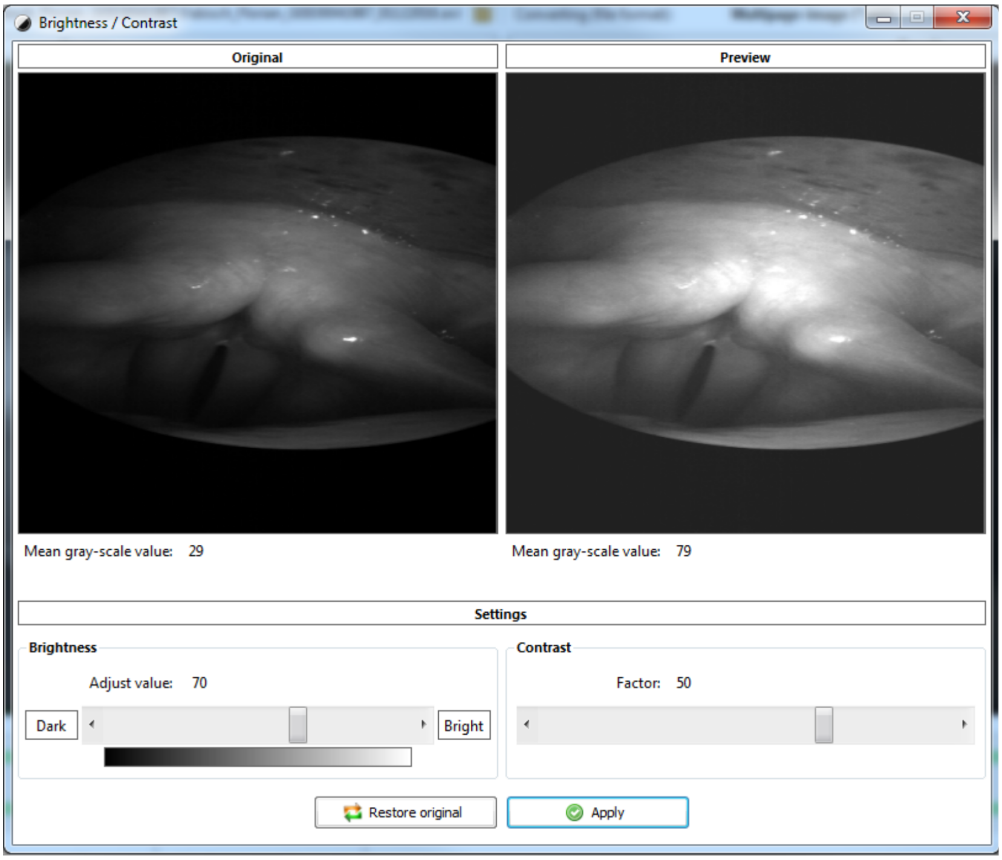


Figure 31 *– Video Editing Tool: Brightness / Contrast*

7 Parameter Information

The following tables contain a short overview of the properties of all parameters.

7.1 Fundamental frequency

***Table 1** – Fundamental Frequency Measures*

name	Fundamental Frequency
unit	Hertz (Hz)
range	$F(x) = [0, \infty)$
comment	Fundamental frequency of signal periods
expected values	About 125 Hz for men, 250 Hz for women
name	Cycle Duration
unit	Milliseconds (ms)
range	$F(x) = [0, \infty)$
comment	Duration of one signal period
expected values	About 8 ms for men, 4 ms for women
name	Maximum Harmonic
unit	Hertz (Hz)
range	$F(x) = [0, \infty)$
comment	Maximal detected multiple of fundamental frequency
expected values	Integer multiple of fundamental frequency
name	Minimum Subharmonic
unit	Hertz (Hz)
range	$F(x) = [0, \infty)$
comment	Minimal detected fraction of fundamental frequency
expected values	Fractions of fundamental frequency

7.2 Amplitude Perturbation

***Table 2** – Amplitude Perturbation Measures*

name	Mean Shimmer
unit	Decibel (dB)
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of amplitudes between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Shimmer(%)
unit	Dimensionless
range	$F(x) = [0, 300]$
comment	Describes normalized mean fluctuation of cycle amplitudes between two cycles **new version since GAT 2019**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Amplitude Perturbation Quotient
unit	Dimensionless
range	$F(x) = [0, \infty) \quad \lvert k \in \{3, 5, 11\}$
comment	Describes mean fluctuation of amplitudes between k cycles, is insensitive to constant linear deviations in entire signal **use not recommended, instead use Mean Shimmer**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Amplitude Perturbation Factor
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of amplitudes between two cycles
expected values	values close to zero; reaches zero if there is no fluctuation
name	Amplitude Variability Index
unit	Decibel (dB)
range	$F(x) = (-\infty, \infty)$
comment	Describes mean fluctuation of amplitudes
expected values	Values close to zero, or negative values; reaches $-\infty$ if there is no fluctuation

7.3 Period Perturbation

***Table 3** – Period Perturbation Measures*

name	Mean Jitter
unit	Milliseconds (ms)
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of cycle durations between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Jitter(%)
unit	Dimensionless
range	$F(x) = [0, 300]$
comment	Describes normalized mean fluctuation of cycle durations between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Jitter Ratio
unit	Dimensionless
range	$F(x) = [0, 3000]$
comment	Describes normalized mean fluctuation of cycle durations between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Jitter-Factor
unit	Dimensionless
range	$F(x) = [0, 300]$
comment	Describes normalized mean fluctuation of inverted cycle durations between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation. Difference between Jitter-Factor and Jitter(%): Jitter-Factor reaches high values especially if the signal contains shortened cycles, Jitter(%) if it contains prolonged ones
name	Period Perturbation Quotient
unit	Dimensionless
range	$F(x) = [0, \infty) \quad \vert k \in \{3, 5, 11\}$
comment	Describes mean fluctuation of cycle durations between k cycles; is insensitive to constant linear deviations in entire signal **use not recommended, instead use Mean Jitter**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Period Perturbation Factor
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of cycle durations between two cycles
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Relative Average Perturbation V1
unit	Dimensionless
range	$F(x) = [0, \frac{4}{3}]$
comment	Describes mean fluctuation of cycle durations between three cycles; insensitive to overall linear deviations **use not recommended, instead use Mean Jitter**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Relative Average Perturbation V2
unit	Dimensionless
range	$F(x) = [0, \frac{20}{9}]$
comment	Describes mean fluctuation of cycle durations between three cycles; insensitive to overall linear deviations, upper boundary is dependent on number of cycles **use not recommended, instead use Mean Jitter**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Period Variability Index
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of cycle durations
expected values	Values close to zero; reaches zero if there is no fluctuation

7.4 Energy Perturbation

Table 4 *– Energy Perturbation Measures*

name	Energy Perturbation Quotient
unit	Dimensionless
range	$F(x) = [0, \infty) \quad \vert k \in \{3, 5, 11\}$
comment	Describes mean fluctuation of cycle energy between k cycles; is insensitive to constant linear deviations in entire signal **use not recommended, instead use Energy Perturbation Factor**
expected values	Values close to zero; reaches zero if there is no fluctuation
name	Energy Perturbation Factor
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes mean fluctuation of cycle energy
expected values	Values close to zero; reaches zero if there is no fluctuation

7.5 Noise

Table 5 *– Noise Measures 1*

name	Harmonics Intensity
unit	Dimensionless
range	$F(x) = [0, 100]$
comment	Describes the harmonics (without fundamental frequency) share of the total Fourier spectrum; makes the problematic assumption of infinite narrow harmonics. **use not recommended, instead use Signal-to-Noise Ratio V1 or Normalized Noise Energy**
expected values	Values distinct higher than zero; reaches 100 for infinite number of harmonics without noise; reaches zero for signal without harmonics (besides fundamental frequency)
name	Harmonics-to-Noise Ratio
unit	Decibel (dB)
range	$F(x) = (-\infty, \infty)$
comment	Describes mean deviation of all signal cycles from the average signal cycle; cannot reach negative values for positive signals
expected values	High values; reaches $-\infty$ for alternating cycles, mirrored on the x-axis (e.g. positive and negative cycles); reaches zero for high dissimilarity of cycles (e.g. one zero-only-cycle and one with at least one entry > 0); reaches ∞ if all cycles are exact identical
name	Normalized Noise Energy
unit	Decibel(dB)
range	$F(x) = (-\infty, 0]$
comment	Sets the Fourier coefficients of an estimated overall noise-signal in relation to the total Fourier coefficients of the signal in Fourier spectrum. The noise-Signal is estimated using the signal in between the harmonics. This gives a ratio between estimated noise and entire signal
expected values	High, negative values; reaches $-\infty$ for an estimated noise of zero; reaches zero if the estimated noise and the actual signal are identical.
name	Spectral Flatness
unit	GAW and trajectories: log_{10}(pixel)^2 audio: log_{10}(dimensionless)^2
range	$F(x) = (-\infty, 0]$
comment	Describes similarity between measured signal and a signal with a "flat" spectrum (i.e. a white noise signal)
expected values	High, negative values; approaches $-\infty$ for a pure tone; approaches zero for white noise

Table 6 – Noise Measures 2

name	Glottal-to-Noise Excitation Ratio
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Measures noise in impulse-train-like signals. Thereto the Hilbert envelopes of the inverse filtered signal are calculated and correlated. For a pure impulse-train this envelopes are identical, for pure white noise they differ maximal from each other.
expected values	Values close to 1; reaches -1 for anti correlated envelopes; 0 for white noise; 1 for a perfect impulse train
name	Waveform Matching Coefficient
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Describes the maximum pairwise similarity of signal periods. Calculates therefore the cosine of the angle between all pairs of periods, which are defined as vectors in a multidimensional space; cannot reach negative values for positive signals
expected values	Values close to 1; reaches 0 if all period-vectors are perpendicular to one another; 1 if there are at least two identical periods.
name	Mean Waveform Matching Coefficient
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Describes the mean pairwise similarity of signal periods. Calculates therefore the cosine of the angle between all pairs of periods, which are defined as vectors in a multidimensional space; cannot reach negative values for positive signals
expected values	Values close to 1; reaches 0 if all period-vectors are perpendicular to one another; 1 if all periods are identical.
name	Signal-to-Noise Ratio V1
unit	Decibel (dB)
range	$F(x) = [0, 36]$
comment	Describes the relative energy of a reconstructed harmonic spectrum in relation to the total energy of the signal Fourier spectrum; Assumes finite narrow harmonics (between 12 and 24 Hz per harmonic).
expected values	High values; reaches 0 for a signal without any harmonics; reaches 36 dB for a pure harmonic signal
name	Signal-to-Noise Ratio V2
unit	Decibel (dB)
range	$F(x) = (-\infty, \infty)$
comment	Calculates ratio between energy spectrum of the signal and a via linear prediction decorrelated approximated noise signal.
expected values	High values,reaches $-\infty$ if the estimated noise is identical to the signal, reaches ∞ for indefinitely faint noise.
name	Cepstral Peak Prominence I
unit	(Dezibel (dB))
range	$F(x) = (-\infty, \infty)$
comment	Calculates the difference between the first rahmonic and a, by linear regression, estimated reference value for this peak.
expected values	Values > 0
name	Cepstral Peak Prominence II
unit	(Dezibel (dB))
range	$F(x) = (-\infty, \infty)$
comment	Calculates an average log-Fourier spectrum if two or more spectra of equal length are computed. Calculates thereafter the difference between the first Rahmonic of this mean spectrum and a, by linear regression, estimated reference value for this peak.
expected values	Values > 0

7.6 Mechanical

Table 7 *– Mechanical Measures*

name	Stiffness
unit	$\frac{1}{seconds}$
range	$F(x) = [0, \infty)$
comment	Describes maximum speed, with which the area of the glottis is changing during opening or closing, normalized to total varying glottal area
expected values	Values higher than zero; reaches zero for a not moving glottis; reaches ∞ for infinite fast oping or closing glottis.
name	Peak Closing Velocity
unit	$\frac{Megapixel}{seconds}$
range	$F(x) = [0, \infty)$
comment	Describes maximum speed, with which the area of the glottis would be changing during opening or closing, if the glottis movement is approximated as sinus curve
expected values	Values higher than zero; reaches zero for a not moving glottis; reaches ∞ if the signal period is infinitely small.
name	Peak Acceleration
unit	$\frac{Megapixel}{seconds^2}$
range	$F(x) = [0, \infty)$
comment	Describes maximum acceleration, with which the area of the glottis would be changing during opening or closing, if the glottis movement is approximated as sinus curve
expected values	Values higher than zero; reaches zero for a not moving glottis; reaches ∞ if the signal period is infinitely small.
name	Amplitude-to-Length Ratio
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes the ratio between changing glottal area and length of the glottis. This means the higher this ratio is, the wider the glottis opens in lateral direction.
expected values	Values higher than zero; reaches zero for a not moving glottis; reaches ∞ if the glottis becomes infinitely broad.

7.7 Glottal area waveform: Quotients

Table 8 *– Glottal Area Measures 1*

name	Open Quotient
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the share of time, during which the glottis is open
expected values	Values close to 1; reaches zero for a all time closed glottis; reaches 1 for a all time opened glottis.
name	Closing Quotient
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the share of the time, during which the glottis is closing.
expected values	Values close to 0.5; reaches zero for a never closing glottis; reaches 1 for a only closing glottis

Table 9 *– Glottal Area Measures 2*

name	Speed Quotient
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes the ratio between opening and closing time of the glottis
expected values	Values close to 1; reaches zero if the closing time of the glottis is infinitely longer than the opening time; reaches ∞ if the opening time of the glottis is infinitely longer than the closing time
name	Speed-Index
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Normalized form of speed quotient
expected values	Values close to 0; negative values implicate a longer closing time than opening time; positive values a longer opening time than closing time of the glottis.
name	Rate Quotient
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes the ratio between the time interval, during which the glottis is opening or closed, to the time interval, during which it is closing.
expected values	Values close to 0.5; reaches ∞ for a infinitely fast closing glottis; reaches 0 for a infinitely slow closing glottis.
name	Asymmetry Quotient
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Alternative normalized form of speed quotient
expected values	Values close to 0.33; lower values implicate a longer closing time than opening time; higher values a longer opening time than closing time of the glottis.
name	Glottis Gap Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the ratio between minimum and maximum opened glottis
expected values	Values close to zero; reaches zero for a entirely closing glottis; reaches 1 for a not moving open glottis
name	Plateau Quotient
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes how much the GAW is narrowed to the top
expected values	Values close (but not too close) to zero; reaches zero if the GAW has infinitely narrow peaks; reaches 1 for a constant GAW.
name	Glottal Area Index
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes how much of the area of the GAW is located close to it's position of maximal displacement.
expected values	Values not close to zero or ∞ ; the higher the value, the narrower the GAW; the lower the value the broader the GAW

7.8 Glottal area waveform: Periodicity

Table 10 *– GAW Periodicity Measures*

name	Amplitude-Periodicity
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes fluctuation of amplitudes between two cycles
expected values	Values close to zero; reaches zero if one amplitude is zero; reaches 1 for identical amplitude size.
name	Time-Periodicity
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes fluctuation of cycle duration between two cycles
expected values	Values close to zero; reaches zero if one cycle length is zero; reaches 1 for identical cycle lengths.

7.9 Glottal area waveform: Derivatives

Table 11 *– Derivative Measures*

name	Maximum Area Declination Rate
unit	$\frac{Megapixel}{seconds}$
range	$F(x) = [0, \infty)$
comment	Describes the maximum closing velocity of the GAW
expected values	High values; reaches zero for a not moving glottis; reaches ∞ for infinitely fast closing glottis.
name	Amplitude Quotient
unit	Seconds (s)
range	$F(x) = [0, \infty)$
comment	Describes maximum closing velocity of the GAW normalized to its amplitude
expected values	Values higher than zero; reaches zero for a infinitely fast closing glottis; reaches 1 for a not moving glottis (i. e. $GA_i = const. \Rightarrow MADR = 0,\ A_i = 0$); reaches ∞ for a infinitely long period.

7.10 Symmetry

Table 12 *– Symmetry Measures 1*

name	Phase Asymmetry Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the difference in phase of left and right GAW. Does not differ between left and right
expected values	Values close to zero; reaches zero if left and right corresponding GAW cycles are in phase; reaches 1 if both are as far apart as possible.
name	Spatial Symmetry Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the difference in area of left and right GAW. Does not differ between left and right
expected values	Values close to zero; reaches zero if the left and right corresponding GAW cycles have the same area; reaches 1 if one cycle is identical to the total GAW cycle and the other one is constant zero.

***Table 13** – Symmetry Measures 2*

name	Dynamic Range Symmetry Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the difference in size of the changing glottal area of left and right GAW. Does not differ between left and right
expected values	Values close to 1; reaches zero if one dynamic range is zero; reaches 1 if the left and right corresponding GAW dynamic ranges are identical
name	Amplitude Symmetry Index
unit	Dimensionless
range	GAW: $F(x) = [0, 1]$ Trajectories: $F(x) = [-1, 1]$
comment	Describes the difference in size of the maximal glottal area of left and right GAW. Does not differ between left and right
expected values	Values close to 1; reaches -1 if one Trajectory-cycle-maximum is -1 times the other; reaches zero if one maximal size is zero; reaches 1 if the left and right corresponding GAWs are even in maximal size;
name	Waveform Symmetry Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the difference in overall shape of left and right GAW. Calculates therefore the cosine of the angle between pairs of GAW cycles, which are defined as vectors in a multidimensional space. Because the GAW is strict positive, the maximum angle between the vectors is 90°. Therefore the minimum of WSI is 0.5. Does not differ between left and right
expected values	Values close to 1; reaches 0 if one Trajectory-cycle-vector is -1 times the other; reaches 0.5 if one GAW-cycle-vector is perpendicular to the other; reaches 1 if the left and right corresponding GAWs are identical
name	Phase Asymmetry
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Describes the difference in phase of left and right GAW. Differs between left and right
expected values	Values close to zero; reaches zero if left and right corresponding GAW cycles are in phase; reaches 1 or -1 if the distance between both is maximal.
name	Spatial Symmetry
unit	Dimensionless
range	$F(x) = [-1, 1]$
comment	Describes the difference in area of left and right GAW. Differs between left and right
expected values	Values close to zero; reaches zero if the left and right corresponding GAW cycles have the same area; reaches -1 if the right cycle is identical to the total GAW cycle and the left one is constant zero; reaches 1 if the left cycle is identical to the total GAW cycle and the right one is constant zero .
name	Dynamic Range Symmetry
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes the difference in size of the dynamic glottal area of left and right GAW. Differs between left and right
expected values	Values close to 1; reaches zero if one dynamic range is zero; reaches 1 if the left and right corresponding GAW dynamic ranges are identical
name	Amplitude Symmetry
unit	Dimensionless
range	GAW: $F(x) = [0, \infty)$ Trajectories: $F(x) = (-\infty, \infty)$
comment	Describes the difference in total size of the glottal area of left and right GAW. Differs between left and right
expected values	Values close to 1; reaches $-\infty$ if the left maximum is negative and the other one is infinitely close to zero; reaches zero if the total size of the left glottal area is zero; reaches ∞ if the total size of the right glottal area is infinitely close to zero

7.11 Phonovibrogram

***Table 14** – Phonovibrogram Measures*

name	Contour Angles
unit	Degrees (deg.)
range	$F(x) = [0, 180]$
comment	Different angles which are calculated between border areas of open and closed conditions in the Phonovibrogram
expected values	Values close to 90°
name	Contour Angles Symmetry Index
unit	Dimensionless
range	$F(x) = [0, 1]$
comment	Describes the ratio between one angle of the left and one of the right Phonovibrogram side. Normalizes to the greater angle
expected values	Values close to 1; reaches 0 for an angle of 0°; reaches 1 for equal angles
name	Contour Angles Symmetry
unit	Dimensionless
range	$F(x) = [0, \infty)$
comment	Describes the ratio between one angle of the left and one of the right Phonovibrogram side
expected values	Values close to 1; reaches 0 for a left sided angle of 0°; reaches ∞ for a right sided angle infinitly close to 0°

8 Literature

[1] S. Bielamowicz, J. Kreiman, B. R. Gerratt, M. S. Dauer, and G. S. Berke. “Comparison of voice analysis systems for perturbation measurement.” eng. In: *J Speech Hear Res* 39.1 (1996), pp. 126–134.

[2] R. J. Baken and R. F. Orlikoff. *Clinical measurement of speech and voice*. Second edition. Cengage Learning, 1999.

[3] I. R. Titze, Y. Horii, and R. C. Scherer. “Some technical considerations in voice perturbation measurements”. eng. In: *Journal of Speech, Language and Hearing Research* 30.2 (1987), pp. 252–260.

[4] D. Michaelis. “Das Göttinger Heiserkeits-Diagramm-Entwicklung und Prüfung eines akustischen Verfahrens zur objektiven Simmgütebeurteilung pathologischer Stimmen”. PhD thesis. Georg-August-Universität Göttingen, Mathematisch-Naturwissenschaftlich Fakultät, 1999.

[5] J. Lessing. “Entwicklung einer Klassifikationsmethode zur akustischen Analyse fortlaufender Sprache unterschiedlicher Stimmgüte mittels Neuronaler Netze und deren Anwendung”. PhD thesis. Georg-August-Universität Göttingen, Mathematisch-Naturwissenschaftlich Fakultät, 2007.

[6] H. Kasuya, Y. Endo, and S. Salin. “Novel Acoustic Measurements of Jitter and Shimmer Characteristics from Pathological Voice”. In: *EUROSPEECH’93* 1 (1993), pp. 1973–1976.

[7] Y. Koike. “Application of Some Acoustic Measures for the Evaluation of Laryngeal Dysfunction”. In: *Studia Phonoligica* 7 (1973), pp. 17–23.

[8] J. Hillenbrand and R. A. Houde. “Acoustic Correlates of Breathy Vocal Quality: Dysphonic Voices and Continuous Speech”. In: *Journal of Speech and Hearing Research* 39 (1996), pp. 311–321.

[9] E. Yumoto, W. J. Gould, and T. Baer. “Harmonics-to-noise ratio as an index of the degree of hoarseness.” eng. In: *J Acoust Soc Am* 71.6 (1982), pp. 1544–1549.

[10] H. Kasuya, S. Ogawa, K. Mashima, and S. Ebihara. “Normalized noise energy as an acoustic measure to evaluate pathologic voice.” eng. In: *J Acoust Soc Am* 80.5 (1986), pp. 1329–1334.

[11] D. Michaelis, M. Fröhlich, and H. W. Strube. “Selection and combination of acoustic features for the description of pathologic voices.” eng. In: *J Acoust Soc Am* 103.3 (1998), pp. 1628–1639.

[12] F. Klingholz. “The Measurement of the signal-to-noise ratio (SNR) in Continuos Speech”. In: *Speech Communication* 6 (1987), pp. 15–26.

[13] F. Klingholz. “Acoustic representation of speaking-voice quality.” In: *J Voice* 4.3 (1990), pp. 213–219.

[14] Y. Qi, R. E. Hillman, and C. Milstein. “The estimation of signal-to-noise ratio in continuous speech for disordered voices.” eng. In: *J Acoust Soc Am* 105.4 (1999), pp. 2532–2535.

[15] J. Hillenbrand, R. A. Cleveland, and R. L. Erickson. “Acoustic correlates of breathy vocal quality.” eng. In: *J Speech Hear Res* 37.4 (1994), pp. 769–778.

[16] K. G. Munhall, D. J. Ostry, and A. Parush. “Characteristics of velocity profiles of speech movements.” eng. In: *J Exp Psychol Hum Percept Perform* 11.4 (1985), pp. 457–474.

[17] I. R. Titze. “Mechanical stress in phonation.” eng. In: *J Voice* 8.2 (1994), pp. 99–105.

[18] E. B. Holmberg, R. E. Hillman, and J. S. Perkell. “Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice.” eng. In: *J Acoust Soc Am* 84.2 (1988), pp. 511–529.

[19] N. Henrich, G. Sundin, D. Ambroise, C. d’Alessandro, M. Castellengo, and B. Doval. “Just noticeable differences of open quotient and asymmetry coefficient in singing voice.” eng. In: *J Voice* 17.4 (2003), pp. 481–494.

[20] R. Patel, D. Dubrovskiy, and M. Döllinger. “Characterizing vibratory kinematics in children and adults with high-speed digital imaging.” eng. In: *J Speech Lang Hear Res* 57.2 (2014), S674–S686.

[21] D. D. Mehta, M. Zaéartu, T. F. Quatieri, D. D. Deliyski, and R. E. Hillman. “Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy.” eng. In: *J Acoust Soc Am* 130.6 (2011), pp. 3999–4009.

[22] G. Chen, J. Kreiman, B. R. Gerratt, J. Neubauer, Y.-L. Shue, and A. Alwan. “Development of a glottal area index that integrates glottal gap size and open quotient.” eng. In: *J Acoust Soc Am* 133.3 (2013), pp. 1656–1666.

[23] Q. Qiu, H. K. Schutte, L. Gu, and Q. Yu. “An automatic method to quantify the vibration properties of human vocal folds via videokymography.” eng. In: *Folia Phoniatr Logop* 55.3 (2003), pp. 128–136.

[24] P. Alku, M. Airas, E. Björkner, and J. Sundberg. “An amplitude quotient based method to analyze changes in the shape of the glottal pulse in the regulation of vocal intensity.” eng. In: *J Acoust Soc Am* 120.2 (2006), pp. 1052–1062.

[25] I. R. Titze. “Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation.” eng. In: *J Speech Lang Hear Res* 49.2 (2006), pp. 439–447.

[26] R. R. Patel, D. Dubrovskiy, and M. Döllinger. “Measurement of glottal cycle characteristics between children and adults: physiological variations.” eng. In: *J Voice* 28.4 (2014), pp. 476–486.

[27] P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, 2005.

[28] D. Voigt, M. Döllinger, A. Yang, U. Eysholdt, and J. Lohscheller. “Automatic diagnosis of vocal fold paresis by employing phonovibrogram features and machine learning methods.” eng. In: *Comput Methods Programs Biomed* 99.3 (2010), pp. 275–288.

[29] D. Voigt, M. Döllinger, T. Braunschweig, A. Yang, U. Eysholdt, and J. Lohscheller. “Classification of functional voice disorders based on phonovibrograms.” eng. In: *Artif Intell Med* 49.1 (2010), pp. 51–59.

[30] M. Döllinger, J. Lohscheller, A. McWhorter, and M. Kunduk. “Variability of normal vocal fold dynamics for different vocal loading in one healthy subject investigated by phonovibrograms.” eng. In: *J Voice* 23.2 (2009), pp. 175–181.

[31] M. Döllinger, D. Dubrovskiy, and R. Patel. “Spatiotemporal analysis of vocal fold vibrations between children and adults.” eng. In: *Laryngoscope* 122.11 (2012), pp. 2511–2518.

List of Figures