Vocal Tremor Measurement Based on Autocorrelation of Contours

Markus Brückl, Fachgebiet Kommunikationswissenschaft, Technische Universität Berlin, Germany

Abstract

An algorithm to measure vocal tremor is presented, validated, and applied. The expected input is a sound file that captures a sustained phonation. The 6 output values are the frequency of frequency and amplitude tremor, intensity indices of frequency and amplitude tremor, and power indices of frequency and amplitude tremor. Basic principles of the algorithm are (1) autocorrelations of pitch and amplitude contours that are based on an autocorrelation of the input signal, (2) the correction for declination of (natural) contours as well as (3) a contour peak-picking and -averaging method for the determination of tremor intensities. The tremor power indices are new measures that weight tremor intensities by tremor frequency in order to receive bio- and psychologically more significant measures of tremor magnitude. The algorithm is implemented as a script of an open-source speech analysis program providing an most accurate pitch-detection (autocorrelation) method.

Motivation

Despite the well known and broad spectrum of applicability of acoustical tremor measurement in speech research, biological, and psychological domains, only few systems are available (e.g. [1]) which extract vocal tremor from the acoustic signal.

Aims

Therefore the aim is to provide an open-source – and thus inspectable, discussable, adaptable, and free of charge – algorithm (Praat [2] script) to automatically measuring vocal tremor.

Definitions

Vocal tremor is generally defined as an unintentional low-frequency modulation of the vocal fold vibration. However, unlike other tremors the acoustic representation of vocal tremor channels into two components, a frequency (F_0) and an amplitude (intensity) tremor, of which each has its frequency and intensity. Their definitions are adopted from MDVP [1]:

$FTrF := Fftr$: frequency of F_0 tremor

$ATrF := Fatr$: frequency of amplitude tremor

$$FTrI := FTRI := 100 \cdot \frac{absFTRI - \bar{F}_0}{\bar{F}_0} \quad (1)$$

: intensity index of F_0-tremor

$$ATrI := ATRI := 100 \cdot \frac{absATRI - \bar{A}}{\bar{A}} \quad (2)$$

: intensity index of amplitude-tremor

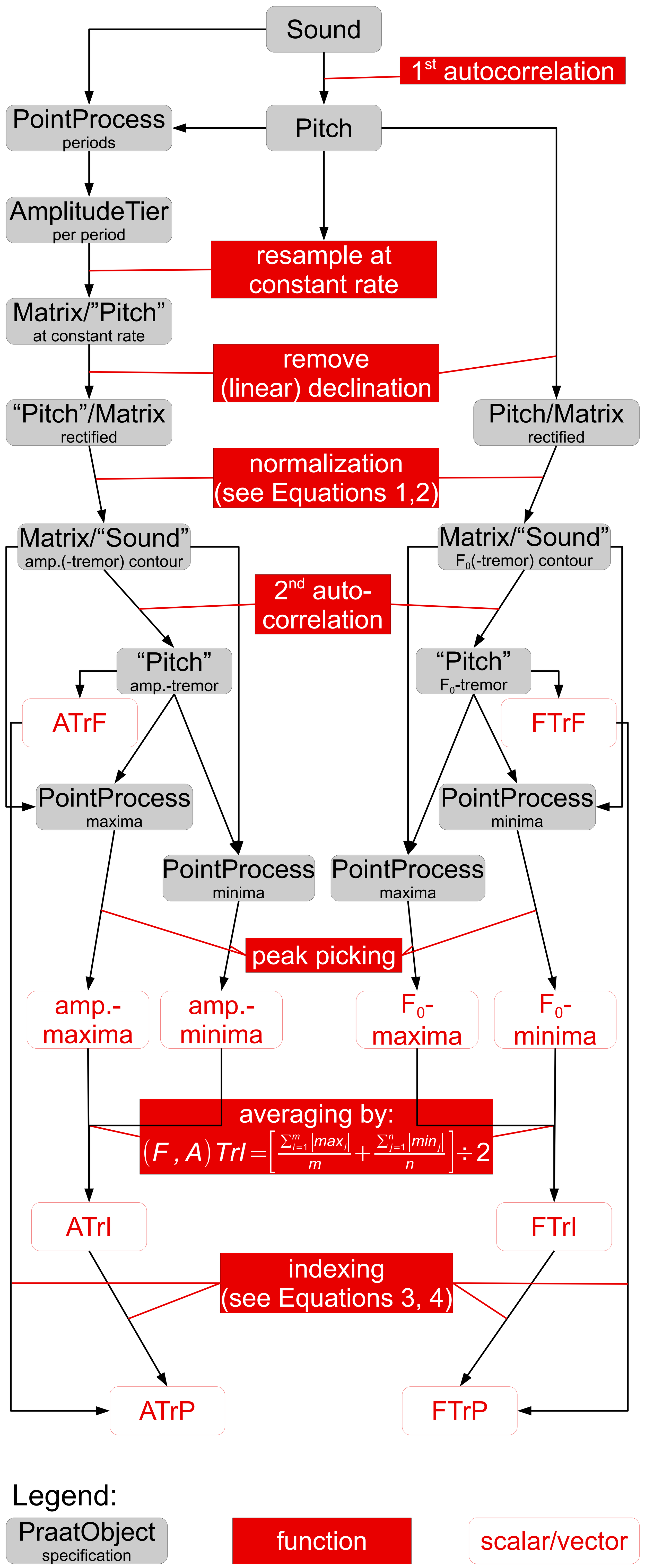
$$FTrP := FTrI \cdot \frac{FTrF}{FTrF+1} \quad (3)$$

: power index of F_0-tremor

$$ATrP := ATrI \cdot \frac{ATrF}{ATrF+1} \quad (4)$$

: power index of amplitude-tremor

Implementation of the algorithm



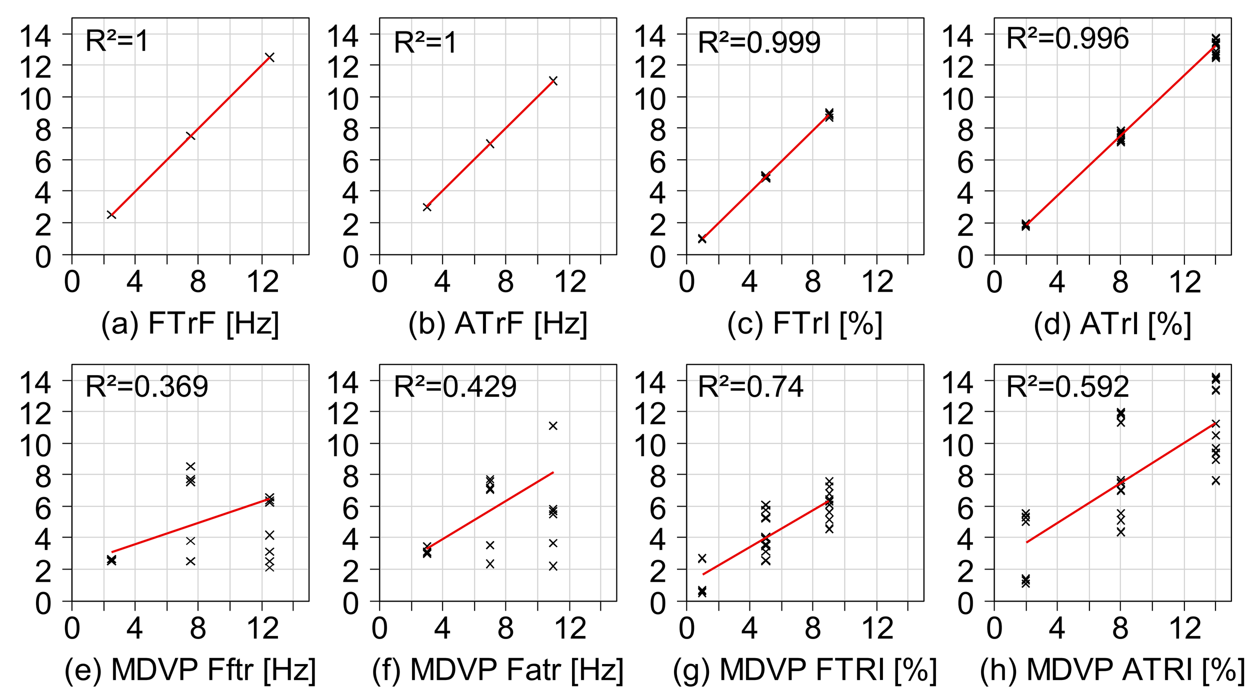
Validation

This algorithm and the 4 MDVP [1] measures are validated by measuring tremor in 729 sounds with known properties, generated according to the following equation:

$$s(t) = \sin\left[2\pi t \cdot \frac{bf - decf \cdot t}{2} + \frac{ftri \cdot amf}{ftrf} \cdot \sin(2\pi t \cdot ftrf)\right] \cdot \left[ba - deca \cdot t + atri \cdot ama \cdot \sin(2\pi t \cdot atrf)\right] \quad (5)$$

t: a certain point in time [s]
$s(t)$: (intensity [Pa] of the) signal at time t
ama: mean intensity [Pa]
amf: mean fundamental frequency [Hz]
ba: intensity [Pa] at time t=0
bf: fundamental frequency [Hz] at t=0
$deca$: (linear) declination of intensity [Pa/s]
$decf$: (linear) declination of fundamental frequency [Hz/s]
$ftrf$: (synth.) frequency tremor frequency [Hz]
$atrf$: (synth.) amplitude tremor frequency [Hz]
$ftri$: (synth.) frequency tremor intensity index
$atri$: (synth.) amplitude tremor intensity index

The results of measurements on these input signals are visualized in the following scatterplots showing the measured values (ordinates) as functions of the synthetically given ones (abscissae). Figures (a)-(d) are based on data of the new measures, Figures (e)-(h) on data of measures of the MDVP.



(a) FTrF [Hz] (b) ATrF [Hz] (c) FTrI [%] (d) ATrI [%]
(e) MDVP Fftr [Hz] (f) MDVP Fatr [Hz] (g) MDVP FTRI [%] (h) MDVP ATRI [%]

An examination of the results on the new measures reveals that tremor frequency is detected exactly – each time (see Figures (a) and (b)). The intensity indices (see Figures (c) and (d)) correlate very highly with the synthetically given values. The tremor intensity (and power) measurements become more accurate if the analysis time step is reduced.

In contrast the MDVP parameters (Figures (e)-(h)) show considerably greater measurement errors, especially in the frequency extraction. The measured tremor frequency and tremor intensity values are generally too low.

Application: indicating speaker age

Both tremor measuring procedures are also applied to natural voices from 88 women comprising all adult ages (chronological age (CA): AM=50.42a SD=17.64a).

They sustained the vowel /a/ as long as possible. Three 2.2s lasting segments were extracted: a start segment including the vocal onset, a quasi-stationary (middle) segment and the end including the offset. These were rated by approximately 30 Listeners in order to estimate the speakers' age. These values [a] were corrected for listener bias and averaged to perceptive ages (PA) per vowel-segment. For details on these data please refer to [3].

The results of correlation analyses (Pearson's r) between tremor intensity and power measures and these age scales are reported in the following table:

	/a/ start		q.-s. /a/		/a/ end	
	CA	PA	CA	PA	CA	PA
FTRI	**0,511**	**0,476**	**0,410**	**0,377**	*0,250*	**0,393**
Ln(FTRI)	**0,484**	**0,419**	**0,461**	**0,315**	**0,308**	**0,455**
ATRI	-0,055	0,011	0,137	0,209	*0,289*	*0,305*
Ln(ATRI)	0,053	0,049	0,197	0,147	0,176	0,189
FTrI	**0,340**	**0,359**	**0,216**	**0,450**	0,103	**0,290**
Ln(FTrI)	**0,422**	**0,468**	**0,389**	**0,622**	**0,290**	**0,513**
ATrI	**0,544**	**0,504**	**0,364**	**0,505**	0,119	**0,438**
Ln(ATrI)	**0,588**	**0,509**	**0,526**	**0,554**	*0,228*	**0,517**
FTrP	**0,326**	**0,359**	*0,236*	**0,475**	0,135	**0,325**
Ln(FTrP)	**0,404**	**0,454**	**0,403**	**0,660**	**0,315**	**0,544**
ATrP	**0,550**	**0,530**	**0,424**	**0,558**	*0,236*	**0,492**
Ln(ATrP)	**0,573**	**0,520**	**0,577**	**0,601**	**0,318**	**0,555**

Conclusions

The new algorithm provides considerably more valid measures of vocal tremor than the compared MDVP measures. Its measures also serve as very good indicators of female speakers' age extracted from sustained vowel input.

The notion of the need to consider pitch and intensity declinations in tremor measurement proved to be awarding as well as the invention of tremor power indices. Both innovations can be considered as steps towards reliably measuring glottal tremor in connected speech.

References

[1] Kay Elemetrics Corp., "Multi-Dimensional Voice Program (MDVP), Model 5105" (Version 2.6.2), [Computer program], 1993/2003.
[2] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer" (Version 5.3.04), [Computer program], University of Amsterdam. Online: http://www.praat.org/, accessed on 19 Feb 2012.
[3] Brückl, M., "Altersbedingte Veränderungen der Stimme und Sprechweise von Frauen", W. Sendlmeier [Ed], Mündliche Kommunikation, Vol. 7, Logos Verlag, Berlin, 2011.