VOICE RANGE PROFILE EXAMINATION METHOD AND ITS APPLICATIONS

PAVEL NOVÝ, FRANTIŠEK VÁVRA, MICHALA KOTLÍKOVÁ

Abstract. The aim of the application of the voice range profile examination method in laryngeal disease diagnostics is to observe the development of the voice apparatus function after the surgery was performed. Since we deal with the examination of voice heavily affected by the hoarseness, we need to propose a robust algorithm for evaluating a fundamental frequency and a sound pressure level (phonation intensity of examined voice) in a real-time. The post-surgical progress is tested through the development of parameters derived from the statistical and pictorial features of the voice range profile.

Keywords: Voice Range Profile, Phonetogram, Fundamental Frequency, Sound Pressure Level, Autocorrelation Function, Fourier Descriptors
AMS classification: 62M10, 68T50, 92C50

1. Introduction

The human voice range profile (VRP) measurement is the functional acoustic examination method which provides information about the actual frequency and sound pressure level of examined VRP. This method originated in 1970 (Damsté), the recommendation for the examination standardization was published in 1983 (Schutte, Seidner) by the Union of European Phoniatricians (UEP). We noticed the modification of discussed method and the development of a phonetogram evaluation with respect to the set of parameters since 1991 (Pabon), 1992 (Titze), 1994 (Sulter, Shutte), 1999 (Bloothooft, Pabon), see [2], [5], [6].

The principle of the VRP examination is based on activity where during the vocal phonation (mostly "a") an examined person produces all achievable tones of his/her voice frequency range for both the minimal and maximal intensity level. The result of such examination is the phonetogram presenting the graphical interpretation of the vocal voice range. Apart from the vocal range testing the spoken voice range is occasionally tested by reading some standard text. In both tests all technical parameters of the acoustic scanning chain have to be guaranteed including the constant distance of the microphone and adequate limitation of the environmental noise.

The paper was supported by the Research Plan of Ministry of Education: "Information Systems and Technologies", No. MSM-235200005.

Particular applications of the voice range profile examination can be divided into five areas:

- assessment of the fundamental frequency and sound pressure level of trained voice,
- training of the transitions between vocal registers,
- logopaedia,
- load tests of the voice apparatus,
- functional diagnostics of the voice apparatus disease.

One of the first symptoms of the voice defect - dysphonia is the hoarseness. Hoarseness can be caused either by a change of the vocal cords physical properties (irregular oscillation) or by a vocal cords incompetence (manifested by the voice murmur). Due to the cause of defect, dysphonia can be divided into two fields: functional and organic. This division is however soft, because the functional dysphonia may lead up to secondary organic changes at vocal cords and the organic defect can be the secondary cause of the functional voice defect. See [3], [4].

Many organic vocal cords affections both primary and secondary require the surgery treatment. Following diagnoses belong to such findings at the vocal cords, see figure 1.1:

- polyp - organic diverticulum filled with ligament or organized blood tumour (haematoma),
- nodule - bounded mucous membrane pachytes which inhibit the voice cleft from the total closure,
- edema - (Reinke's edema), the voice cleft closes at the edema only,
- cyst - encapsulated formation inside the vocal cord which causes its rough limbus,
- tumour - benign and cancerous,
- ...

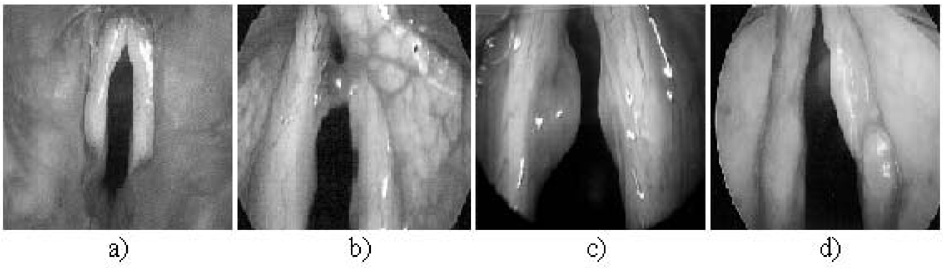
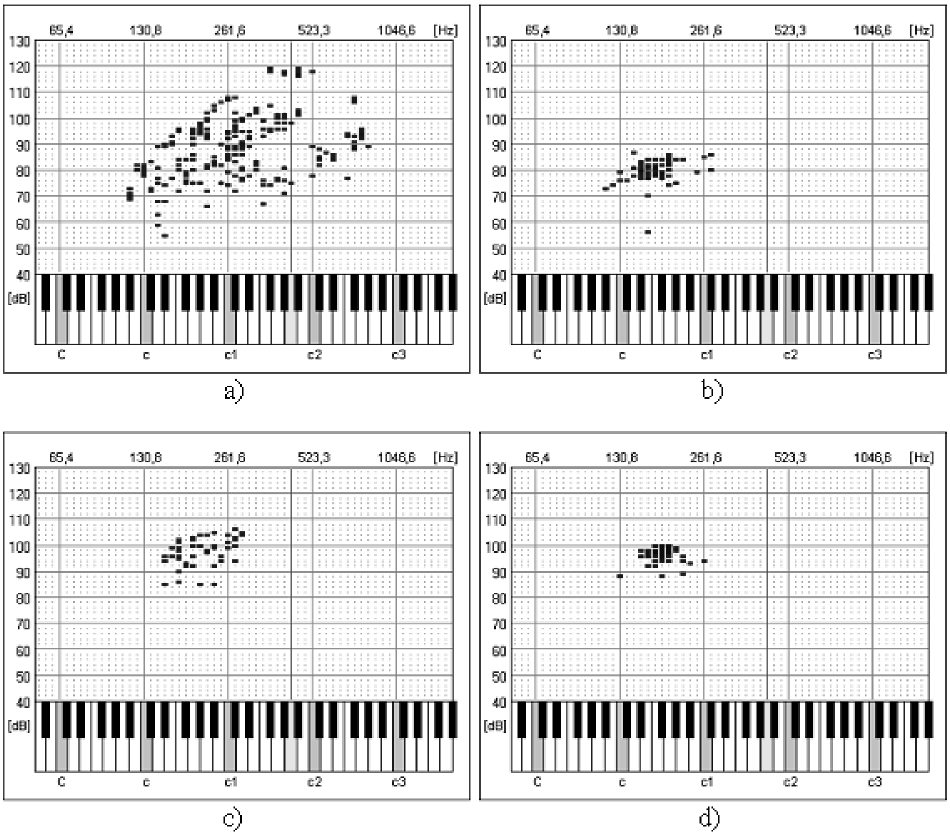


Fig. 1.1: Laryngostroboscopic view of vocal cord
a) standard finding,
b) nodule,
c) Reinke's edema,
d) left vocal cord polyp.

The aim of the application of the voice range profile examination method is the observation and assessment of the voice function development after performing the surgery and following therapy and voice training at the ORL (Oto-Rhino-Laryngology) clinic at the University hospital in Pilsen. VRP assessment is focused onto three main diagnostics types: the one- and double- sided polyps, the Reinke's edema and the vocal cord nodules. The vocal (phonation of "a") and spoken VRP has been used within the examination. See figure 1.2.



a) b)

c) d)

Fig. 1.2: Examples of observed voice range profiles
a) vocal VRP - normal finding: phonation of the vowel "a",
b) spoken VRP - normal finding: standard text reading,
c) vocal VRP - Diagnosis: polyp of the vocal cord,
d) spoken VRP - Diagnosis: polyp of the vocal cord, see c).

2. VOICE RANGE PROFILE

According to used phoniatrical terminology, the principle of the voice range profile examination is the assessment of the voice level and its loudness i.e. the sound pressure level. Voice level evaluation is based on the assessment of the fundamental frequency $F_0[dB]$ which corresponds to the voice cords oscillation, see [2], [3]. By means of the autocorrelation function (ACF) used in our analysis we will talk about the repetition frequency of the voice signal. To suit the purpose of the VRP, the fundamental frequency F_0 is quantized according to uniformly tempered twelve-stage tuning and it is usually measured for following amplitudes: $F_0(min) = 55$ Hz and $F_0(max) = 1661$ Hz, which means for five octaves. Loudness is then the voice intensity expressed in values of the acoustic pressure $SPL[dB]$. SPL is quantized within the 1dB step, usual interval of the sound pressure level is between $SPL_{min} = 30$ dB and $SPL_{max} = 150$ dB.

Let the voice range profile be represented by the matrix VR, where νr_{ij} stands for the periodicity of the couple $[SPL_i, f_j]$ during certain VRP examination. $SPL_i[dB]$ then the sound pressure level (voice intensity) measured at the fundamental frequency $f_j[dB]$.

We assume to have measured digitized voice signal $s(k)$ with the sampling period T_{Samp}. Slow changes of the voice signal allow us to use methods of short-term analysis, see [7], [8]. The task of the voice signal analysis is then solved for chosen section - microsegment of the measured voice signal. The microsegment contains N samples and its time length is T_M.

Then we can write:

$$k = 0, 1, ..., N-1, \qquad T_{Samp} = \frac{1}{F_{Samp}},$$
$$N = \frac{T_M}{T_{Samp}}, \quad T_M > T_0(max), \quad T_0(max) = \frac{1}{F_0(min)}.$$

If we assume (for particular application) that $F_{Samp} = 44100$ Hz, $T_M = 30$ ms and $F_0(min) = 55$ Hz, then following holds true: $N = 1323$ samples and $T_0(max) = 18,18$ ms.

3. VOICE SIGNAL ANALYSIS

The aim of the microsegment voice signal analysis is the determination of the couple $[SPL_i, f_j]$, which means the evaluation of the fundamental frequency $F_0[Hz]$ and the sound pressure level $SPL[dB]$. Short time autocorrelation function is used for the purpose of the voice signal periodicity analysis.

3.1. **AUTOCORRELATION FUNCTION.** Let autocorrelation function (ACF) be defined as:

$$R_n(m) = \frac{1}{N} \sum_{k=-\infty}^{+\infty} s(k) \cdot w(n-k) \cdot s(k+m) \cdot w(n-k+m), \tag{1}$$

where $s(k)$ is k^{TH} sample of the voice signal, $w(n)$ is weight function of the rectangular window type:

$$\begin{aligned} w(n) &= 1 \quad \text{for } 0 \le n \le N-1, \\ w(n) &= 0 \quad \text{for all other } n. \end{aligned}$$

Arising out of mentioned ideas, if we focus onto the single microsegment containing N samples, we could write for ACF (after small rearrangements):

$$R(m) = \frac{1}{N-m} \sum_{k=0}^{N-1-m} s(k) \cdot s(k+m), \text{ for } m = 0, 1, ..., N-1. \tag{2}$$

If the current voice signal is periodic with the period T_S, then ACF acquire maximal values for $m = 0, T_S, 2T_S, 3T_S, \ldots$. For the fundamental frequency F_0 then stays:

$$R(m^*) = max_{m=1,\ldots,N-1}[R(m)], \tag{3}$$

$$F_0 = \frac{F_{Samp}}{m^*}, \tag{4}$$

where m^* is the number of samples corresponding to the signal periodicity.

Equations (1) to (4) represent the standard approach of using the short time ACF. While applying the VRP examination to the voice affected by dysphonia (caused by the organic affection) we can assume that the voice signal $s(k)$ will not contain only formants (for vocals F_1 and F_2). It will also contain the breathiness additions, manifestation of irregular voice cords oscillation, especially diplophony and hoarseness caused by the change in amplitude or period of particular vibrations. The requirement of the task real-time processing represents another limitation in terms of time consumption of the final algorithm.

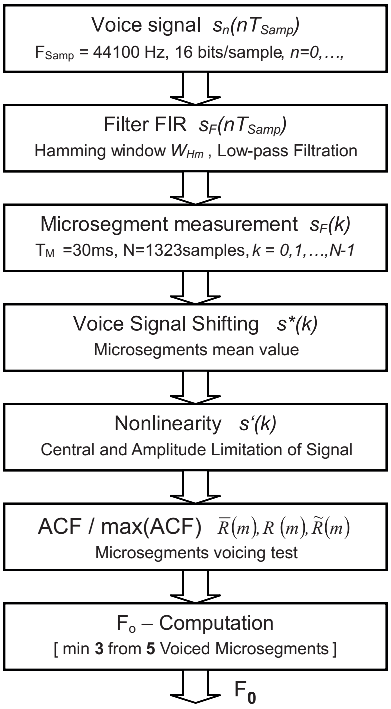


Fig. 3.1: Procedure of voice signal processing and fundamental freq. evaluation.

On the base of the ACF properties analysis (see [1]) and series of performed experiments the methodology of voice signal pre-processing and microsegment signal processing was established. Further on, the evaluation of autocorrelation function was supplemented with the microsegment voicing tests. Mentioned activities resulted in the following procedure of the voice signal processing $(s_n(k), s_F(k))$ and fundamental frequency evaluation (F_0), see fig. 3.1.

Let us assume the sequence of voice signal samples discrete in time and quantized in amplitude $s_n(nT_{Samp}) = s_n(k)$, where $k = nT_{Samp} = 0, 1, ..., \infty, F_{Samp} = 44100$ Hz, 16bits/sample.

In the first step, we perform the filtration using the non-recursive low-pass filter FIR. The filter has the symmetric Hamming window, where w_{Hm} stays for filter coefficients and $2N_{Hm}+1$ is their number (see [8]). The aim of this filtration is the suppression of higher frequency components, which can be generally invoked by previously mentioned hoarseness and its causes. Result of such filtering activity is the filtered voice signal $s_F(nT_{VZ}) = s_F(k)$:

$$s_F(k) = \sum_{j=-N_{Hm}}^{+N_{Hm}} s_n(k-j) \cdot w_{Hm}(j), \quad \text{where } k = nT_{Samp} = 0, 1, ..., \infty. \tag{5}$$

Let us define the microsegment of length N samples for such pre-processed signal. Filtered voice signal in microsegment will be further denoted as $s_F(k)$, but now $k = 0, 1, ..., N-1$.

Second step is shifting the microsegment signal over its mean value $E(s_F)$:

$$s^*(k) = s_F(k) - E(s_F), \quad E(s_F) = \frac{1}{N}\sum_{k=0}^{N-1} s_F(k). \tag{6}$$

Another step represents central and amplitude limitation of the $s^*(k)$ signal, see [7] with non-linearity of following type:

$$\begin{aligned} s'(k) &= +1, \text{ if } s^*(k) > \alpha \cdot min[Max_1, Max_2, Max_3], \\ s'(k) &= -1, \text{ if } s^*(k) < \alpha \cdot max[Min_1, Min_2, Min_3], \\ s'(k) &= 0, \text{ for other } s^*(k), \end{aligned} \tag{7}$$

where Max_1, Max_2 and Max_3 are observed maximal values of the signal $s^*(k)$ in three consequential segments the microsegment is divided into. Analogically Min_1, Min_2 and Min_3 are minimal values observed in such segments. Coefficient $\alpha \in (0, 1)$ stays for the non-linearity threshold level.

Having the adjusted signal $s'(k)$ we compute the autocorrelation function $R'(m)$ in accordance to equation (2) and we perform its standardization and restriction to non-negative values:

$$\bar{R}(m) = \frac{R'(m)}{R'(0)}, \quad R'(m) = \frac{1}{N-m}\sum_{k=0}^{N-1-m} s'(k) \cdot s'(k+m), \tag{8}$$

$$\begin{aligned} \tilde{R}(m) &= \bar{R}(m) \quad \text{for } \forall\, \bar{R}(m) \geq 0, \\ \tilde{R}(m) &= 0 \qquad\ \ \text{for } \forall\, \bar{R}(m) < 0, \\ &\quad \text{for } \forall\, m = 0, 1, ..., N-1. \end{aligned} \tag{9}$$

According to equations (3) and (4) we set the number of samples to m^*, which corresponds to periodicity of the signal $s(k)$. By means of the threshold level $\beta \in (0, 1)$ we can adjust the sensitivity of method for the frequency detection and microsegment voicing.

$$\bar{R}(m^*) = max_{m=1,...,N-1}[\bar{R}(m)], \quad \bar{R}(m^*) > \beta \Rightarrow \mu_{segment-voiced}. \tag{10}$$

Final fundamental frequency is then computed as a harmonic average of frequencies which corresponds at least to three voiced microsegments out of five successive microsegments analyzed:

$$F_0 = \frac{F_{Samp}}{m_{3/5}}, \quad m_{3/5} = \frac{1}{n_z}\sum_{i=1}^{n_z} m_i^*, \quad \text{for } n_z = 3, 4, 5. \tag{11}$$

Results of described procedure of the voice signal processing are displayed step by step on figures 3.2 and 3.3.

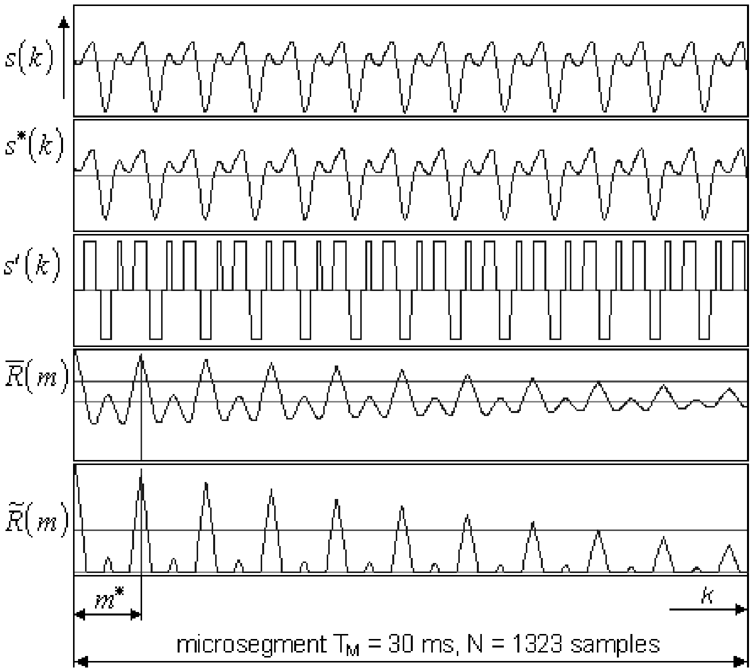


Fig. 3.2: Results of individual step of the voice signal processing for the purpose of fundamental frequency determination, record of the vowel "a", fundamental frequency $F_0 = 440$ Hz, so-called "chamber A", standard phonation.

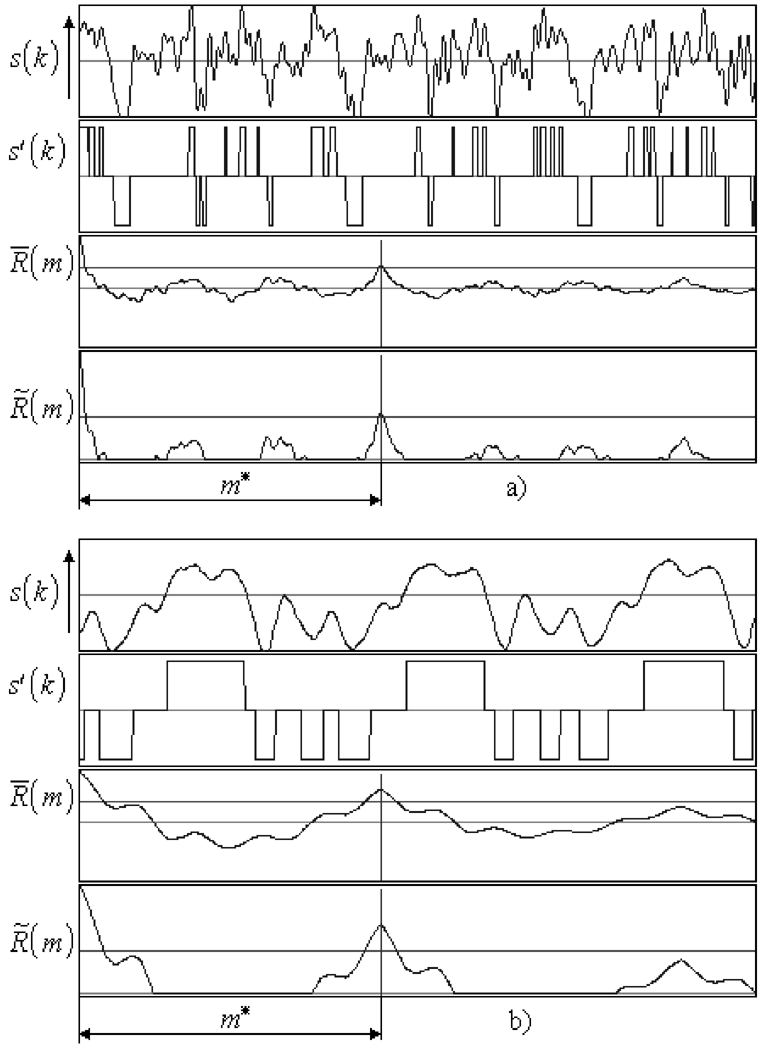


Fig. 3.3: Comparison of the voice signal processing and frequency determination results, record of the vowel "a", fundamental frequency $F_0 = 92,5$ Hz:

a) Diagnosis: polyp of the vocal fold,

b) standard phonation.

3.2. **SOUND PRESSURE LEVEL COMPUTATION.** The sound pressure level is computed for certain microsegment which is assessed as voiced with the help of autocorrelation function as follows:

$$SPL(T_M) = \frac{1}{T_M}\int_0^{T_M} s^2(t)dt \cong SPL^{TM}, \tag{12}$$

$$SPL^{TM} = \frac{T_{Samp}}{2T_M}\sum_{k=1}^{N-1}\Big(s^2(k-1)+s^2(k)\Big) = \frac{1}{N-1}\Big(\frac{s^2(0)}{2}+\frac{s^2(N-1)}{2}+\sum_{k=1}^{N-2}s^2(k)\Big),$$

where T_M is length of microsegment [ms], N is number of samples in microsegment and T_{Samp} is sampling period.

Sound pressure level expression in decibels generally holds:

$$SPL_{[dB]} = 10\log(SPL^{TM}).$$

Computation of final $SPL_{[dB]}$ for the purpose of $[SPL_i, f_j]$ substitution follows the similar rules as the assessment of fundamental frequency F_0 does. It means that at least for three microsegments out of five successive microsegments we compute the arithmetic average of the SPL^{TM} value.

$$(13)\qquad \begin{aligned} SPL_{[dB]} &= 10\log(SPL^{TM}_{3/5}),\\ SPL^{TM}_{3/5} &= \frac{1}{n_z}\sum_{i=1}^{n_z} SPL^{TM}_i, \text{ where } n_z = 3,4,5. \end{aligned}$$

4. VOICE RANGE PROFILE PARAMETERS

A set of parameters is defined to observe and assess the development of the voice range function (by means of the VRP examination method) after the operation and followed voice training were performed.

For such purpose we assume the voice range profile to be represented by a matrix VR, where $\nu r_{i,j}$ stands for the periodicity of the couple $[SPL_i, f_j]$ occurrence during performed examination. Voice range profile can be then understood as:

- two-dimensional statistical set,
- image matrix,
  - binary voice range profile for $\nu r_{i,j} \in \{0,1\}$,
    $\nu r_{i,j} = 0 \ldots$ particular couple $[SPL_i, f_j]$ was not observed during examination,
    $\nu r_{i,j} = 1 \ldots$ at least one particular couple $[SPL_i, f_j]$ was observed during examination.
  - voice range profile for $\nu r_{i,j} \in N$, i.e. for $\forall\, \nu r_{i,j}, \nu r_{i,j} \geq 0$ holds true.

In following, only some examination parameters of the whole set will be introduced, see figures 4.1 and 4.2.

- Parameters of the VRP range, where for $\forall\, \nu r_{i,j}, \nu r_{i,j} \in \{0,1\}$ holds true
  - maximal and minimal values of achieved fundamental frequency

$$(14)\qquad F_{min}, F_{max}, \quad \Delta F_0 = F_{max} - F_{min}$$

  - frequency range $[oct]$

$$(15)\qquad F_{range} = 6\,\frac{\log(F_{max}) - \log(F_{min})}{\log(2093) - \log(32.7)}[oct]$$

  - maximal a minimal values of achieved sound pressure level

$$(16)\qquad SPL_{min}, SPL_{max}, \quad \Delta SPL = SPL_{max} - SPL_{min}$$

- Parameters derived from the statistical features of VRP
  - average occurrence values of the voice frequency F_μ and intensity SPL_μ in the VRP, where for $\forall\, \nu r_{i,j}, \nu r_{i,j} \in \{0,1\}$ holds true.

$$F_\mu = \frac{1}{N_{VRP}} \sum_{k=1}^{N_{VRP}} f_k, \quad SPL_\mu = \frac{1}{N_{VRP}} \sum_{k=1}^{N_{VRP}} SPL_k, \tag{17}$$

N_{VRP} total amount of couples $[SPL_i, f_j]$ in examined VRP
f_k fundamental frequency in the k^{TH} couple $[SPL_i, f_j]$,
SPL_K sound pressure level in the k^{TH} couple $[SPL_i, f_j]$.

  - Coordinates of the VRP centre F_T and SPL_T regardless of the number of occurrence, i.e. $\nu r_{i,j} \in \{0,1\}$ holds true.

$$\begin{aligned}
M_{00} &= \sum_{SPL_i} \sum_{f_j} SPL_i^0 f_j^0 \nu r_{i,j} = \sum_{SPL_i} \sum_{f_j} \nu r_{i,j} = N_{VRP}, \\
M_{01} &= \sum_{SPL_i} \sum_{f_j} SPL_i^0 f_j^1 \nu r_{i,j} = \sum_{SPL_i} \sum_{f_j} f_j \nu r_{i,j}, \\
M_{10} &= \sum_{SPL_i} \sum_{f_j} SPL_i^1 f_j^0 \nu r_{i,j} = \sum_{SPL_i} \sum_{f_j} SPL_i \nu r_{i,j},
\end{aligned}$$

$$F_T = \frac{M_{01}}{M_{00}} = F_\mu, \quad SPL_T = \frac{M_{10}}{M_{00}} = SPL_\mu, \tag{18}$$

M_{00} is the zeroth moment and M_{01}, M_{10} are the first moments of the VRP area.

  - slope of the VRP (slope of the regression line) $SPL = a_{VRP} F_0 + b_{VRP}$

$$a_{VRP} = \frac{\sum_{SPL_i} \sum_{f_j} (f_j - F_\mu)(SPL_i - SPL_\mu)}{\sum_{SPL_i} \sum_{f_j} (f_j - F_\mu)^2}, \tag{19}$$

for $\forall$ couples $[SPL_i, f_j]$, where $\nu r_{i,j}$ <> 0 and we assume that $\nu r_{i,j} = 1$.

- Parameters derived from the shape characteristics of the VRP
  - perimeter $P_{erimeter}$ and area A_{rea} of the VRP area
    - ∗ P_{max}, A_{max} - perimeter and area of the circumscribed rectangle which is determined by coordinates $([SPL_{max}, F_{min}], [SPL_{min}, F_{max}])$,
    - ∗ P_{kvx}, A_{kvx} - perimeter and area of the VRP convex hull
    - ∗ P_{VRP}, A_{VRP} - perimeter and area of the examined VRP.
  - circularity (Formfactor) and rectangularity of the VRP area

$$Ff_{VRP} = 4\pi \frac{A_{VRP}}{P_{VRP}^2}, \quad \text{for } P_{VRP} <> 0, \tag{20}$$

$$Rt_{VRP} = \frac{A_{VRP}}{A_{max}}, \quad Rt_{k\nu x} = \frac{A_{k\nu x}}{A_{max}}, \quad \text{for } A_{max} <> 0 \tag{21}$$

  - Fourier description of the VRP area
    Each point of the VRP border line is set by the $[SPL_i, f_j]$ couple.

Closed border line will be further considered as periodical and oriented. For certain coordinates we can then derive the coordinate curves $SPL_i[k]$ and $f_j[k]$, where $k = 1, 2, 3, ..., L$ and L stays for the length of certain VRP border line. Each coordinate curve can be expressed by means of the Fourier series, see [11], [12]:

$$SPL_i[k] = \frac{1}{2}a_0^{SPL} + \sum_{t=1}^{harm}\left(a_t^{SPL}\cos\left(t\frac{2\pi k}{L}\right) + b_t^{SPL}\sin\left(t\frac{2\pi k}{L}\right)\right), \tag{22}$$

$$f_j[k] = \frac{1}{2}a_0^{F} + \sum_{t=1}^{harm}\left(a_t^{F}\cos\left(t\frac{2\pi k}{L}\right) + b_t^{F}\sin\left(t\frac{2\pi k}{L}\right)\right),$$

k order of point at the VRP border line,
L length of the border line,
a_t^{SPL}, b_t^{SPL} coefficients of Fourier series according to the coordinate curve $SPL_i[k]$,
a_t^{F}, b_t^{F} coefficients of Fourier series according to the coordinate curve $f_j[k]$.

Particular terms of Fourier series for $t = 0, 1, 2, ..., Harm$ can be computed as follows:

$$a_t^{SPL} = \sum_{k=1}^{L} SPL_i[k]\cos\left(t\frac{2\pi k}{L}\right),\ b_t^{SPL} = \sum_{k=1}^{L} SPL_i[k]\sin\left(t\frac{2\pi k}{L}\right), \tag{23}$$

$$a_t^{F} = \sum_{k=1}^{L} f_j[k]\cos\left(t\frac{2\pi k}{L}\right),\ b_t^{F} = \sum_{k=1}^{L} f_j[k]\sin\left(t\frac{2\pi k}{L}\right). \tag{24}$$

Having the finite number of Fourier series terms we obtain the approximation of the VRP border line in means of the minimal squared error. To describe the shape of the VRP, several characteristics (FD_t - Fourier Descriptors) are derived. These descriptors are invariant towards shifting, rotation and scaling of the voice range profile, where $t = 1, 2, ..., Harm$. For results see figure 4.2.

$$FD_t = \sqrt{(DD_t^{SPL})^2 + (DD_t^{F})^2},\quad DD_t^{SPL} = \frac{D_t^{SPL}}{D_1^{SPL}},\quad DD_t^{F} = \frac{D_t^{F}}{D_1^{F}}, \tag{25}$$

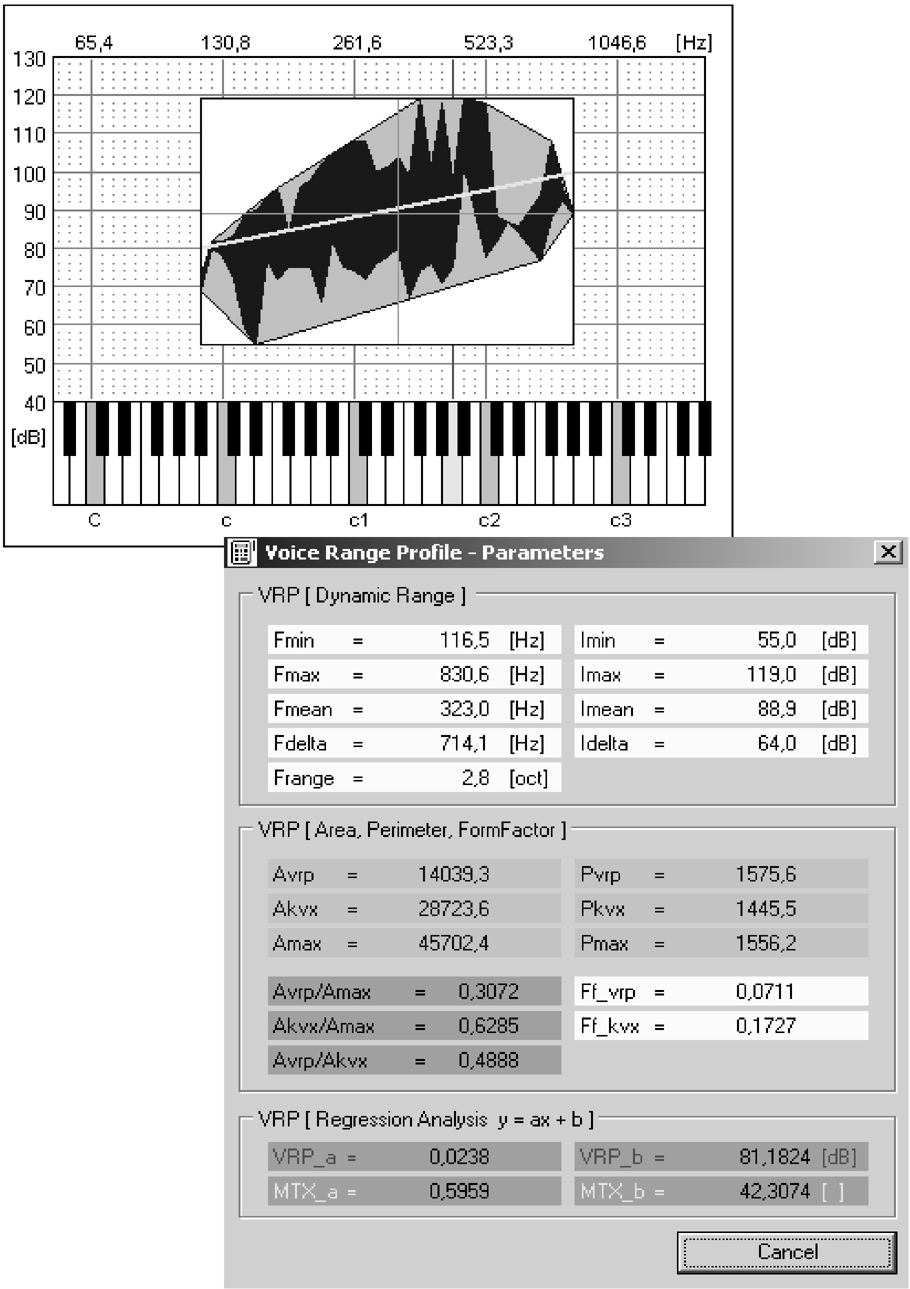
$$D_t^{SPL} = \sqrt{(a_t^{SPL})^2 + (b_t^{SPL})^2},\quad D_t^{F} = \sqrt{(a_t^{F})^2 + (b_t^{F})^2}.$$

5. CONCLUSION

Te aim of this article was to show an application of the voice examination by means of the Voice range profile examination method. We have also introduced one possible approach to fundamental voice frequency assessment for the purpose of testing the voice affected by hoarseness (caused by polyp, Reinke's edema, nodulus, etc.). Verification of the parameters selection and the evaluation method (while observing the voice development after operation) is performed by comparison of achieved results and results of the multidimensional voice analysis, see [4]. This verification is the subject of further cooperation with the ORL clinic at the University hospital in Pilsen. At the time when this contribution was written, there were 24 examined patients waiting for the operation, see Table 5.1.

	polyp	nodulu	edema	basal cell tumour
man	12	0	0	1
woman	11	0	0	0
total	23	0	0	1

Tab. 5.1: Review of patients waiting for the operation who has been examined by the VRP examination method. (Data observed from May to June 2003.)



VRP [Dynamic Range]

Fmin = 116,5 [Hz]
Fmax = 830,6 [Hz]
Fmean = 323,0 [Hz]
Fdelta = 714,1 [Hz]
Frange = 2,8 [oct]

Imin = 55,0 [dB]
Imax = 119,0 [dB]
Imean = 88,9 [dB]
Idelta = 64,0 [dB]

VRP [Area, Perimeter, FormFactor]

Avrp = 14039,3
Akvx = 28723,6
Amax = 45702,4

Pvrp = 1575,6
Pkvx = 1445,5
Pmax = 1556,2

Avrp/Amax = 0,3072
Akvx/Amax = 0,6285
Avrp/Akvx = 0,4888

Ff_vrp = 0,0711
Ff_kvx = 0,1727

VRP [Regression Analysis y = ax + b]

VRP_a = 0,0238
MTX_a = 0,5959

VRP_b = 81,1824 [dB]
MTX_b = 42,3074 [ ]

Fig. 4.1: Parameters of the voice range profile drown from its scope, shape and statistical features. Objects displayed on figure: maximal circumscribed rectangle, convex hull and regression line.

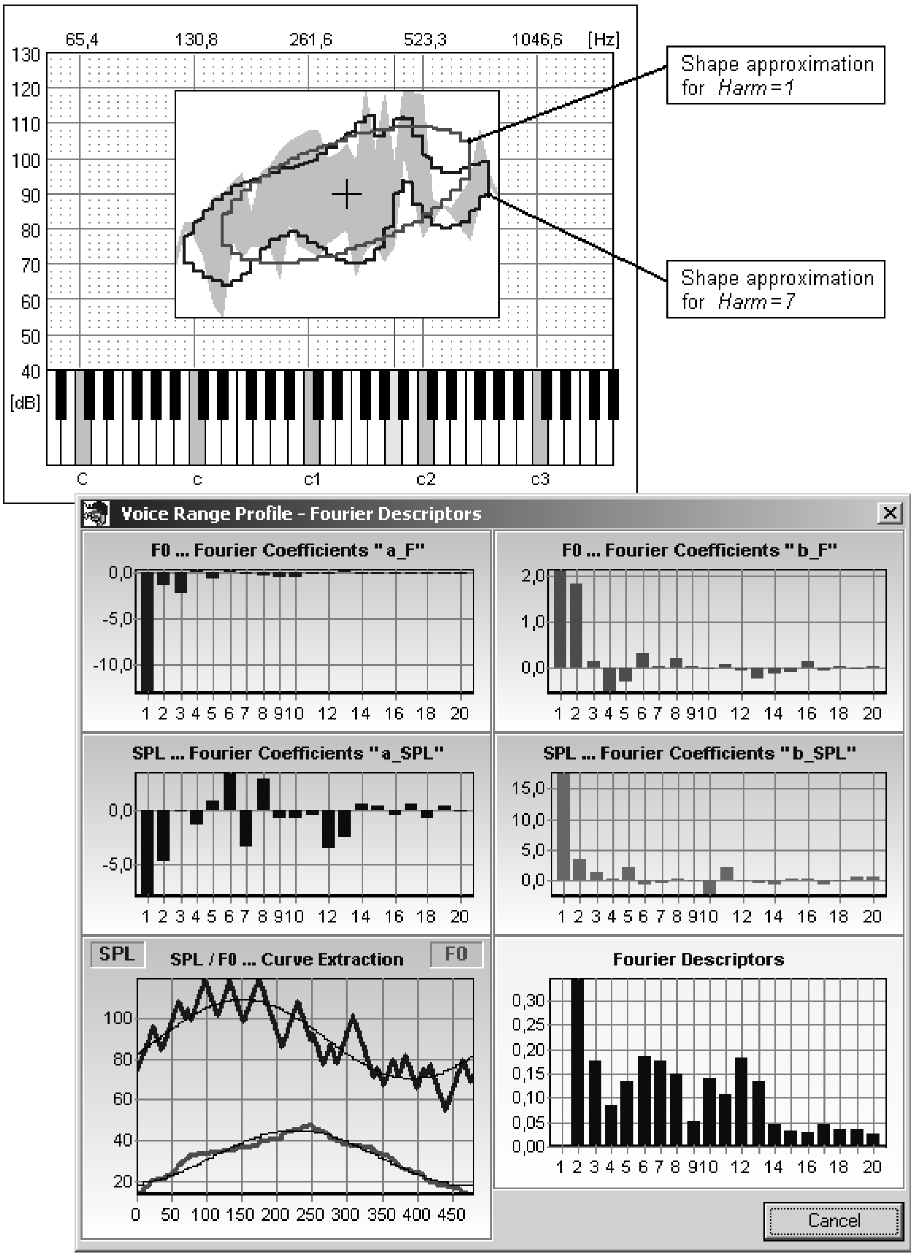


Fig. 4.2: Parameters of Fourier voice range descriptors drawn from the voice range profile. Objects displayed on figure: shape approximation for the number of Fourier series terms $Harm = 1$ and $Harm = 7$. (F_0 is quantized according to uniformly tempered twelve-stage tuning and SPL is quantized within the $1dB$ step.)

REFERENCES

[1] Vávra, F., Nový, P., Mašková, H.: Voice Range Profile and Problem of Fundamental Frequency, Summer School DATASTAT'03, Masaryk University Brno, Svratka, 2003.

[2] Švec, J.: Studium mechanicko-akustických vlastností zdroje lidského hlasu, Doktorská disertační práce, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, 1996.

[3] Vokřál, J.: Akustické parametry chraptivosti, Doktorská disertační práce, ČVUT Praha, Fakulta elektrotechnická, Praha, 1998.

[4] Novák, A.: Foniatrie a pedaudiologie, UNITISK s.r.o., Praha, 1996.

[5] Sulter, A. M.: A structured approach to voice range profile analysis, Journal of Speech and Hearing Research, 1994, 37, 1076-1085.

[6] Bloothooft, G., Pabon, P.: Vocal Registers Revisited, Proceedings Eurospeech '99, Budapest, 423-426, 1999.

[7] Psutka, J.: Komunikace s počítačem mluvenou řečí, Academia Praha, 1995.

[8] Hlaváč, V., Sedláček, M.: Zpracování signálů a obrazů, skripta ČVUT, Praha, 2001.

[9] Pánek, L.: Informační systém pro analýzu hlasového pole, Diplomová práce, Západočeská univerzita, Fakulta aplikovaných věd, Plzeň, 2002.

[10] Zuzák, F.: Compact Standardized Shape Description, Ph.D. Thesis, University of West Bohemia, Department of Computer Science and Engineering, Pilsen, 2000.

[11] Zhang, D., Lu, G.: A Comparative Study of Fourier Descriptors for Shape Representation and Retrieval, The 5th Asian Conference on Computer Vision, Melbourne, Australia, 2002.

[12] Zahn, C. T., Roskies, R. Z.: Fourier Descriptors for Plane Closed Curves, IEEE Trans. on Computers 21, 269–281, 1972.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, FACULTY OF APPLIED SCIENCES, UNIVERSITY OF WEST BOHEMIA IN PILSEN

E-mail address: `novyp@kiv.zcu.cz`, `vavra@kiv.zcu.cz`, `kotlikov@kiv.zcu.cz`