VESTIBULO-OCULAR REFLEX SIGNAL PROCESSING USING EVOLUTIONARY ALGORITHM

Petr Pošík*

* Czech Technical University, Faculty of Electrical Engineering, Dept. of Cybernetics
Technická 2, 166 27 Prague, Czech Republic

posik@labe.felk.cvut.cz

Abstract: Vestibulo-ocular reflex (VOR) is an important source of diagnostic information for physicians. By analyzing it, they can recognize many disorders of the vestibular organ. The VOR signal is a response of the vestibular organ to the so called head rotation test. It is measured by tracking the eye movements, which are, however, distorted by saccades. After filtering the saccades out we are left with discontinuous signal segments. This paper presents an approach to align them to form a smooth signal with the same frequencies that were originally present in the source signal. The approach is based on a direct estimation of the signal component parameters. Two methods of direct search are compared—the Nelder-Mead simplex search and the evolutionary strategy with covariance matrix adaptation. The experimental evaluation on signals with 1 to 5 components revealed that the evolutionary strategy is more robust, scalable and reliable method.

Introduction

Vestibulo-ocular reflex (VOR) is responsible for maintaining retinal image stabilization in the eyes during relatively brief periods of head movement [1]. By analyzing the VOR signal, physicians can recognize some pathologies of the vestibular organ which may result in e.g. failures of the balance of a patient. The recognition of the pathologies is usually done by examining the slow-phase eye velocity and several points of the frequency response (gain and phase shift) of the vestibular system.

The principle of the frequency response measurement is relatively simple: the patient is situated in a chair which is then rotated in a defined way following a source signal—sine wave or a mixture of sine (MOS) waves. The chair with the patient is situated in the dark, the patient has his eyes open and performs some mental tasks which should distract him from mental visualization that could prevent the eye movements which are subsequently measured. This is called the head rotation test. Since the resulting eye signal is distorted by fast eye movements, so-called *saccades*, they must be removed from the signal. This is usually done by computing the angular velocity and the segments with the velocity higher than a predefined threshold are simply omitted from the signal. A method for discovering the right threshold was presented e.g. in [2]. The resulting signal consists of segments of slow phase movements which we are interested in. However, they are not aligned to form a smooth signal (see Fig. 1).

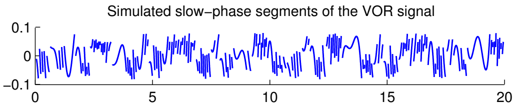

Figure 1: Simulated VOR signal with saccades removed. This is the input of the algorithm.

This VOR signal serves as a source for measuring the slow-phase velocity and the frequency response. The measurement of frequency response is usually done on the basis of interpolating these segments with some smooth curves and performing a Fourier transform of the resulting continuous signal. The frequency response created this way contains, however, some artifacts that come from the artificial interpolation curves and are not generated by the vestibular system.

This paper introduces a method for the direct estimation of the gain and phase lag of the individual sine components of the underlying MOS signal, i.e. for the measurement of several points of the frequency response at the same time. After the estimation, the VOR signal segments should match[1] with the corresponding parts of the estimated MOS signal, as shown in Fig. 2.

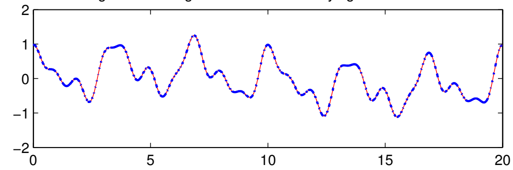

Figure 2: VOR signal segments aligned with the estimated MOS signal. The parameters of the MOS signal are output of the algorithm.

[1] In fact, the slow-phases of the measured VOR signal should go in the opposite direction than the source MOS signal. When the chair rotates to the right, the eyes should move to the left and vice versa.

ISSN: 1727-1983

Problem Specification

It is assumed that the source signal (which controls the rotation of the chair with the patient) is formed as a mixture of sine waves:

$$y^S(t) = \sum_{i=1}^{n} a_i^S \sin(2\pi f_i t + \phi_i^S)\,, \tag{1}$$

where $y(t)$ is the source signal and a_i, f_i and ϕ_i are the amplitude, the frequency and the phase shift of the individual sine components, respectively. The superscript S indicates the relation to the source signal. Note that the frequencies f_i are not marked with this superscript.

Furthermore, it is assumed that the output signal of the vestibular organ is of the same form as the input one, i.e. it contains only sine components with the *same frequencies* as the source signal but possibly with different amplitudes and phase shifts. It should be of the form

$$y(t) = \sum_{i=1}^{n} a_i \sin(2\pi f_i t + \phi_i)\,. \tag{2}$$

If we knew the a_i and ϕ_i parameters of the output MOS signal components, we could calculate the amplification (a_i/a_i^S) and phase lag $(\phi_i - \phi_i^S)$ at individual frequencies and deduce the state of the vestibular organ. (Ideally, we should have amplification[2] of 1 and minimal phase lag at all frequencies, which is not possible. However, physicians can analyze the deviations and diagnose the states that are not OK.)

Unfortunately, we do not have access to the output MOS signal described by Eq. 2. We have only the measured VOR signal, i.e. the segments of the output MOS signal that are left after filtering out the saccades from the eye-tracking signal (see Fig. 1)[3]. However, we can search for the unknown parameters a_i and ϕ_i of the output MOS signal by solving the optimization task described in the following text.

Minimizing Loss Function

Let m be the number of segments of the VOR signal at hand, $v_j(t)$, $j = 1\ldots m$, be the actual j-th segment of the VOR signal and t_j^{ini} and t_j^{end} be the initial and the final time instants for the j-th signal segment. As stated above, we can find the parameters of the output MOS signal by searching the $2n$-dimensional space of points $\mathbf{x}$, $\mathbf{x} = (a_1, \phi_1, \ldots, a_n, \phi_n)$. Such a vector of parameters represents an estimate of the output MOS signal and we can compute the degree of fidelity with which the MOS corresponds to the VOR signal segments by constructing a loss function as follows:

$$L(\mathbf{x}) = \sum_{j=1}^{m} \int_{t_j^{ini}}^{t_j^{end}} ((v_j(t) - \bar{v}_j) - (y(t) - \bar{y}_j))^2 dt\,, \tag{3}$$

[2] Or, rather -1 with respect to the previous footnote.

[3] It is important to note that in this article only artificially generated (simulated) VOR signals were used. This allows for assessing the precision of the proposed method.

where $\bar{v}_j$ is the mean value of the j-th VOR signal segment and is computed as

$$\bar{v}_j = \frac{1}{t_j^{end} - t_j^{ini}} \int_{t_j^{ini}}^{t_j^{end}} v_j(t)dt\,, \tag{4}$$

and $\bar{y}_j$ is the mean value of the current estimate of the output MOS signal related to the j-th segment and is computed as

$$\bar{y}_j = \frac{1}{t_j^{end} - t_j^{ini}} \int_{t_j^{ini}}^{t_j^{end}} y(t)dt\,. \tag{5}$$

Subtracting the means $\bar{v}_j$ and $\bar{y}_j$ from the VOR signal segments $v_j(t)$ and MOS signal $y(t)$, respectively, we try to match the VOR signal segment to the corresponding part of the MOS signal. If they match, their difference is zero, otherwise it is a positive number quadratically increasing with the difference. This operation is carried out for all m VOR signal segments.

In practice we work with the discretized versions of the signals so that we usually substitute the integral with a sum. The equations are then[4]

$$L(\mathbf{x}) = \sum_{j=1}^{m} \sum_{i=t_j^{ini}}^{t_j^{end}} ((v_j(i) - \bar{v}_j) - (y(i) - \bar{y}_j))^2\,, \tag{6}$$

$$\bar{v}_j = \frac{1}{t_j^{end} - t_j^{ini}} \sum_{i=t_j^{ini}}^{t_j^{end}} v_j(i)\,, \tag{7}$$

$$\bar{y}_j = \frac{1}{t_j^{end} - t_j^{ini}} \sum_{i=t_j^{ini}}^{t_j^{end}} y(i)\,. \tag{8}$$

Nature of the Loss Function

In Figures 3, 4 and 5 it is shown what the landscape of the loss function $L(a_1, \phi_1, a_2, \phi_2)$ looks like if two of the parameters are kept fixed. For these figures, the optimal values of the parameters are set to $\mathbf{x} = (0.6, 1, 0.2, 0.2)$ (marked with a small cross in the figures). It seems that the loss function $L(\mathbf{x})$ exhibits many features which are considered to be hard for any optimization algorithm, namely:

Non-separability. It is not sufficient to optimize the parameters one after another. The profile of the loss function along one variable changes significantly with a change in another variable. See Figures 3, 4 and 5—the cross describing the optimum is not situated in the optimum of the cut if the other parameters are not optimal as well. The function cannot be decomposed to a set of simpler optimization tasks.

Long narrow valleys not aligned with coordinate axes. See Fig. 3. Even gradient based algorithms have problems finding minimum of such a landscape. They have to perform many small steps along the bottom of the valley before they hit the optimum.

[4] In the equations 6, 7 and 8, the arrays $v_j(i)$ are supposed to be indexed with i ranging from t_j^{ini} to t_j^{end}.

Multimodality. See Fig. 5. There are several local minima. In this case they are caused by the periodicity of the sine function with respect to the phase shift.

However, based on the experience when optimizing this function I hypothesize that this function could be unimodal, but with very narrow and perhaps tortuous vallyes leading to this global optima.

Optimization Methods

The parameter vector $\mathbf{x}$ is projected to the loss function via the estimate of the MOS signal $y(i)$ (and via the mean values $y_j(i)$). In principle, the loss function $L(\mathbf{x})$ is differentiable with respect to the individual parameters. Thus, to find the optimal values of the parameter vector $\mathbf{x}$ we could compute the partial derivatives of L and use a gradient-based optimization method. However, this approach is not pursued in this article.

Instead, two methods of direct black-box optimization are used: the well known Nelder-Mead downhill simplex search and the evolutionary strategy with covariance matrix adaptation (CMA-ES). There is not enough space in this article to describe them in detail, but a short simplified review is useful.

Nelder-Mead simplex search. It is a well-known and established deterministic optimization algorithm [3]. During the search in D-dimensional space it maintains a set of $D+1$ points, forming the so-called simplex. The search is performed along a line which goes through the worst point of the simplex and the average of the other points. After a better point is found, it replaces a point in the simplex and the algorithm iterates. Because of the simplex behavior during the optimization, the algorithm is sometimes also called the *amoeba* algorithm.

Evolutionary Strategy with Covariance Matrix Adaptation. It is very recent and progressive stochastic optimization algorithm [4]. It maintains a D-dimensional normal distribution from which it samples new data points. The distribution is then in turn adapted based on the loss function values for these new points. The algorithm performs a kind of iterative principal component analysis of the selected perturbation vectors.

One of the aims of this paper is to decide which of these two algorithms is more suitable for solving this particular optimization task.

Experimental Setup

The above described method was tested on artificially generated VOR signals to assess its success and precision and to decide which of the optimization algorithms is more suitable for this task. The tests were carried out on signals consisting of 1 to 5 sine components, i.e. the search was carried out in 2-, 4-, 6-, 8-, and 10-dimensional parameter spaces.

Generating VOR signal. First, for each sine component of the signal, the values of frequency, amplitude and phase shift were randomly generated. The ranges for individual parameters can be found in Table 1. Using these randomly generated values, a continuous MOS signal (which is to be estimated) is created. This signal then undergoes a disruption process which cuts it to individual segments with ‘pauses’ between them. This way the gaps created by filtering out the saccades are simulated. The segments are then placed to the same level (see Fig. 1).

Table 1: Settings for parameters of artificial VOR signal

Parameter	**Value (Range)**
f_i	$\langle 0.05, 2\rangle$
a_i	$\langle 0.2, 2\rangle$
ϕ_i	$\langle 0, \pi/2\rangle$
Sampling Freq.	500 Hz
Signal Duration	20 s
Saccade Duration	0.05 s

Experimental Evaluation. For each number of components, 9 different VOR signals were generated. For each of them the parameters of the underlying MOS were estimated by minimizing the loss function using both the Nelder-Mead simplex search and the CMA-ES. In each run, the algorithms were allowed to perform 10,000 evaluations of the loss function and a particular run was considered to be successful if the algorithm found a parameter set with the loss function value lower than 10^{-8}.

Results

In this section, the results are surveyed, described and discussed.

Success Rates

First, let us review the success rates of both algorithms when estimating the parameters of the MOS signal with the number of components ranging from 1 to 5 (see Tab. 2).

Table 2: Success rates (in percentages) of Simplex and CMA-ES algorithms

Components	**Simplex**	**CMA-ES**
1	100.0	100.0
2	100.0	100.0
3	100.0	100.0
4	44.4	100.0
5	0.0	100.0

As we can see, the simplex algorithm has difficulties with finding the optimum of the loss function in less than 10,000 evaluations when the underlying MOS signal has 4 or more components.

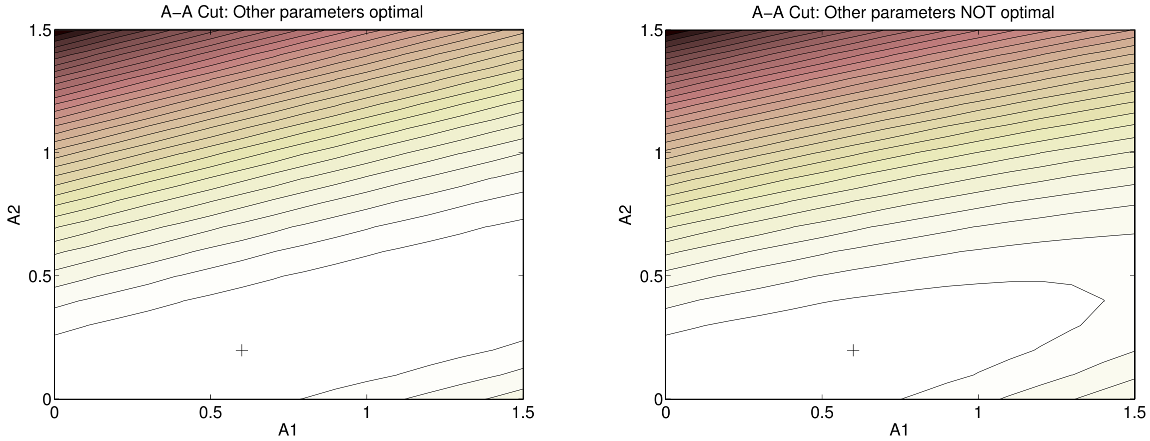

Figure 3: Cuts through the landscape of the loss function $L(a_1, \phi_1, a_2, \phi_2)$ with ϕ_1 and ϕ_2 fixed at their optimal values (left) and at values different from the optimal ones (right).

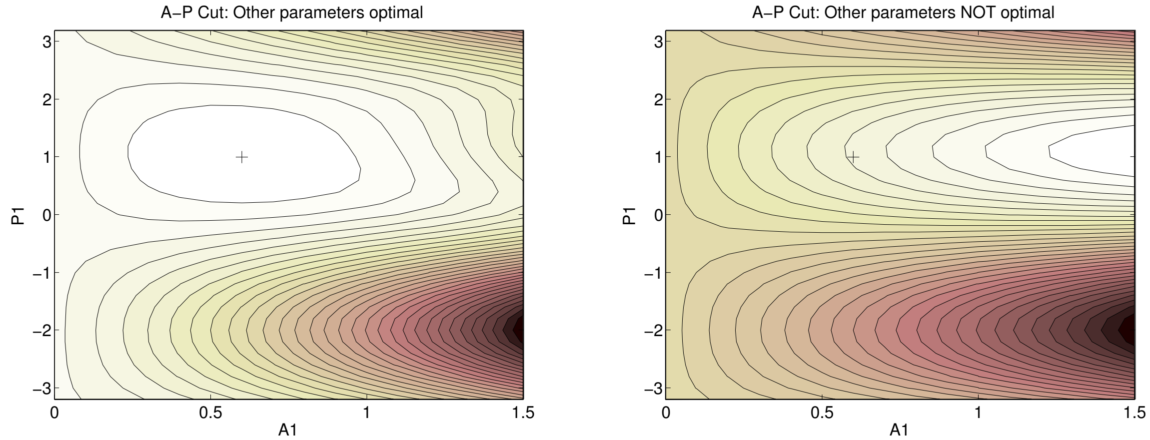

Figure 4: Cuts through the landscape of the loss function $L(a_1, \phi_1, a_2, \phi_2)$ with a_2 and ϕ_2 fixed at their optimal values (left) and at values different from the optimal ones (right).

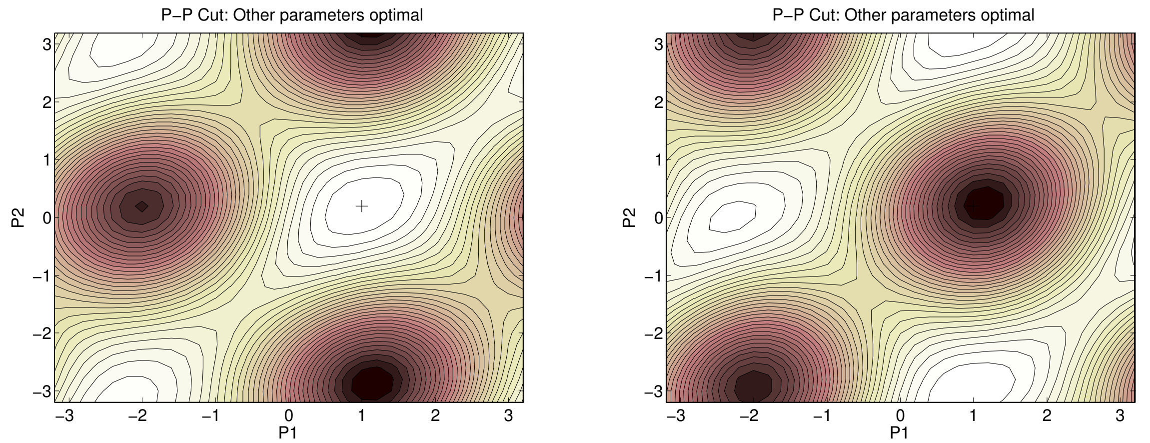

Figure 5: Cuts through the landscape of the loss function $L(a_1, \phi_1, a_2, \phi_2)$ with a_1 and a_2 fixed at their optimal values (left) and at values different from the optimal ones (right).

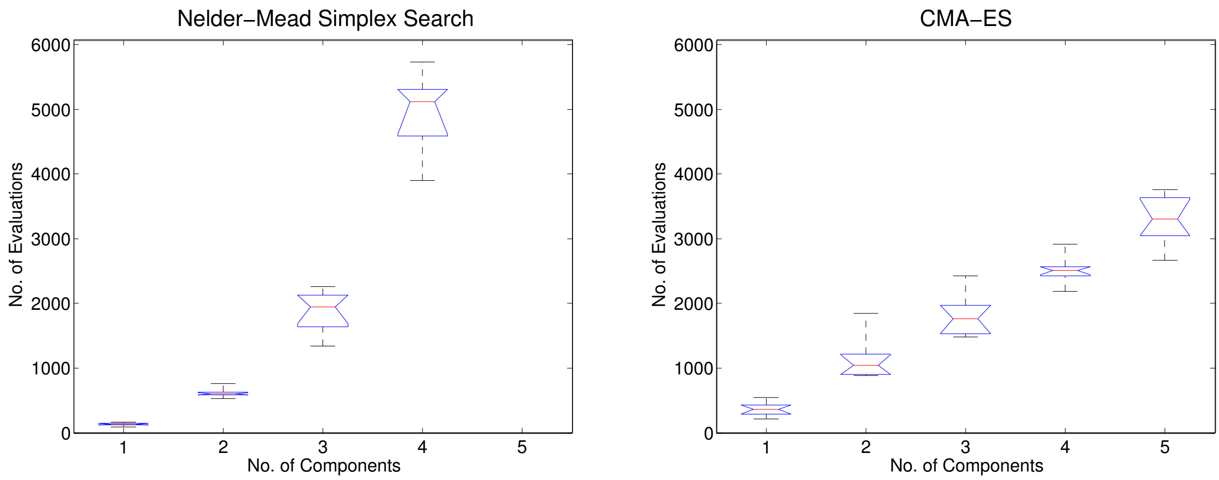

Figure 6: Number of evaluations needed to find a solution with quality better than 10^{-8} as a function of the number of components of the underlying MOS signal. *Middle line:* median, *box:* interquartile range, *whiskers:* minimum and maximum.

Speed of the Algorithms

The comparison of speed is based on the number of evaluations needed to find a solution with loss lower than 10^{-8}, i.e. only successful runs are considered. The results are summarized in Fig. 6. The two graphs reveal that the number of needed evaluations increases with the number of components (i.e. with the dimensionality of the search space) much faster for the simplex search method than for the CMA-ES where the increase is almost only linear (at least subquadratic). CMA-ES is clearly more scalable solution than the simplex search.

Evolution Profiles

The progress of evolution is depicted in Fig. 7. It presents the loss function value of the best solution found so far, averaged over all successful runs. Again, there is no line for the simplex method searching for parameters of 5 components.

Based on this graph, we could make a recommendation not to use the simplex search method when searching for parameters of the MOS signal with more than 2 components. The CMA-ES solves such tasks much better.

Precision of the Estimates

The precision of the solutions provided by the algorithm must be assessed. We know that the segments of the simulated VOR signal come from a MOS signal with some specified parameters $\mathbf{x}' = (a'_1, \phi'_1, \ldots, a'_n, \phi'_n)$. The optimization algorithm provides the estimate of these parameters[5] $\mathbf{x} = (a_1, \phi_1, \ldots, a_n, \phi_n)$. The errors in estimates of the amplitudes are computed as

$$e_i^a = \frac{|a_i - a'_i|}{a'_i}, \tag{9}$$

[5]It is assumed that these estimates are 'normalized', i.e. that all $a_i \geq 0$ and all $\phi_i \in (-\pi, \pi\rangle$.

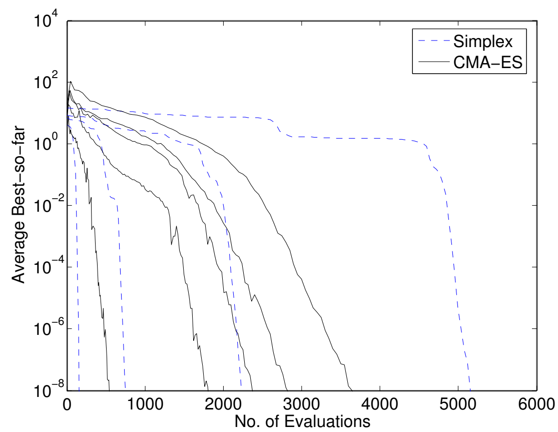

Figure 7: Typical progress of successful search runs as done by the simplex method and by the CMA-ES. Both the leftmost lines (solid and dashed) belong to 1 component, the rightmost dashed line belongs to simplex searching for parameters of 4 components while the rightmost solid line belongs to CMA-ES searching for parameters of 5 components.

and the errors in estimates of the phase shifts are computed as

$$e_i^\phi = \frac{|\phi_i - \phi'_i|}{\pi}. \tag{10}$$

The maximal error values for estimates of amplitudes and phase shifts across all components are presented Tab. 3. The unsuccessful runs are excluded. That is also the reason of missing data for the simplex search with 5 components—there were no successful runs.

Although other fitness thresholds for judging the success of individual runs were not tested, the threshold presents a way how to tune the quality of the parameter estimates.

Table 3: Maximal errors for successful runs

	Simplex		CMA-ES	
Comps	**max**(e^a)	**max**(e^ϕ)	**max**(e^a)	**max**(e^ϕ)
1	1.13e-4	5.71e-6	9.65e-5	2.21e-6
2	6.34e-5	8.41e-6	5.50e-5	5.49e-6
3	4.16e-5	5.74e-6	4.46e-5	6.57e-6
4	1.30e-5	4.11e-6	3.70e-5	1.17e-5
5	—	—	2.84e-5	3.43e-6

Summary, Conclusions and Future Work

In this paper, a new method of VOR signal processing was introduced and experimentally evaluated on artificially generated signals. Similarly to other conventionally used methods, it relies on the right identification of the fast eye movements, saccades, that must be filtered out of the signal in advance. After that, conventionally used methods interpolate the signal segments and carry out the Fourier transform to obtain the amplification and the phase shift on the original frequencies. On the contrary, this method directly estimates these parameters from the signal segments trying to align them with the underlying estimated mixture of sine waves.

In the experimental section, two direct search algorithms were compared. Although the simplex search is faster for signal with 1 or 2 sine components, it does not scale up well. For signals with 3 or more components, the CMA-ES is preferable—it is more robust, reliable and scalable. Depending on the selected threshold for the loss function value, if the individual run succeeds (which was always the case for the CMA-ES), the found solution is very precise compared to the results of Fourier transform.

Although this article can serve as a good proof of concept, this method is more time-demanding than the use of Fourier transform and could be used only as an off-line processing technique. Further investigation is needed to assess if this additional effort is actually worthwhile—if the added value in the precision can be used by physicians to perform better diagnostics.

Acknowledgment

The project was supported by the Ministry of Education, Youth and Sport of the Czech Republic with the grant No. MSM6840770012 entitled "Transdisciplinary Research in Biomedical Engineering II".

References

[1] C.W. STOCKWELL. Vestibular testing: past, present, future. *British Journal of Audiology*, 31(6):387–398, 1997.

[2] D. NOVÁK, D. CUESTA-FRAU, R. BRZEZNÝ, R. ČERNÝ, L. LHOTSKÁ, and V. ECK. Method for clinical analysis of eye movements induced by rotational test. In *Proc. of EMBEC*, 2002.

[3] J.A. NELDER and R. MEAD. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[4] N. HANSEN and A. OSTERMEIER. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.

ISSN: 1727-1983 © 2005 IFMBE