Analysis of Vestibular-Ocular Reflex by Evolutionary Framework

D. Novák[1], A. Pilný[1], P. Kordík[1], Š. Holiga[1], P. Pošík[1],
R. Černý[2], and Richard Brzezny[2]

[1] Czech Technical University in Prague, Czech Republic
xnovakd1@labe.felk.cvut.cz
[2] 2nd Medical Faculty, Charles University, Czech Republic

Abstract. In this paper the problem of analysis of eye movements using sinusoidal head rotation test is presented. The goal of the method is to discard automatically the effect of the fast phase-saccades and consequently calculate the response of vestibular system in the form of phase shift and amplitude. The comparison of threshold detection and inductive models trained on saccades is carried out. After saccades detection 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 using the evolutionary strategy with covariance matrix adaptation. The performance of evolutionary approach is compared to least-square multimodal sinus fit. The experimental evaluation on real-world signals revealed that threshold saccades detection with combination of the evolutionary strategy is robust, scalable and reliable method.

1 Introduction

Vestibulo-ocular reflex (VOR) is responsible for maintaining retinal image stabilization in the eyes during relatively brief periods of head movement. 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 principle of the frequency response measurement using servocontrolled rotating mechanism 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 sum of sine (SOS) waves. 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.

The result of the stimuli is a prevailing pattern called nystagmus consisting of slow and fast phase - see Figure 2(a). First task of this work is the separation of slow and fast phases. Most previous algorithms used to detect fast phases were based on tresholding techniques (TH) [1]. We present a new approach based on Group of Adaptive Models Evolution (GAME) inductive modelling methodology [2].

The second task of this paper is a evolutionary based method with covariance matrix adaptation (CMA) for the direct estimation of the gain and phase lag of the individual sine components of the underlying SOS signal, i.e. for the measurement of several

V. Kůrková et al. (Eds.): ICANN 2008, Part I, LNCS 5163, pp. 452–461, 2008.

points of the frequency response at the same time. After the estimation, the VOR signal segments should match with the corresponding parts of the estimated SOS signal. This approach is compared to least-square method [1] fitted to sinusoid prototype.

The suggested methodology is first evaluated on artificially generated VOR-signal and next on the set of 7 healthy volunteers.

2 Methodology

A proposed methodology of VOR analysis is depicted in Figure 1. First, fast phases (or saccades as shown in Figure 2(a)) of eye movement signal S_{raw}, which is generated by vestibular organar (VO), are detected using GAME approach. After saccades detection, GAME topology selection is performed by inspecting FFT spectra of the reconstructed signal. GAMEs' results are compared to the thresholding method TH on artificially generated VOR signal using two evaluation criteria $E1$: performance rate in % and ration of false positive saccades detections to total number of saccades. Second, saccades are removed and eye signal S_{rem} is reconstructed by estimating vestibular organ's gains and phase shifts to a multimodal sinus technical signal $T(G_i)$. The evolutionary strategy with covariance matrix adaptation (CMA) is applied and its results are compared to linear square sinus fit (LSF) using again synthetic VOR signal. The performance $E2$ is evaluated by a ration of output gains of the processed signal $S(G_i)$ to input gains of the technical signal $T(G_i)$.

2.1 Data Acquisition

Eye movements were recorded using silver/silver chloride electrodes with a reference strip-chart recorder extended by digitized unit. For the eye movements capture the sampling frequency was set at 102.4 Hz that is consider to be enough for accurate capture of all eye movement characteristics. The rotating frequency of chair was constant throughout recording accelerating or decelerating and changing directions hence producing single and composed sinusoidal frequency signals with period $f_1 = 0.05$ Hz and $f_1 = 0.05$ Hz, $f_2 = 0.1$ Hz, resp. 7 patients have been measured, 5 with their head fixed, 2 without any head fixation.

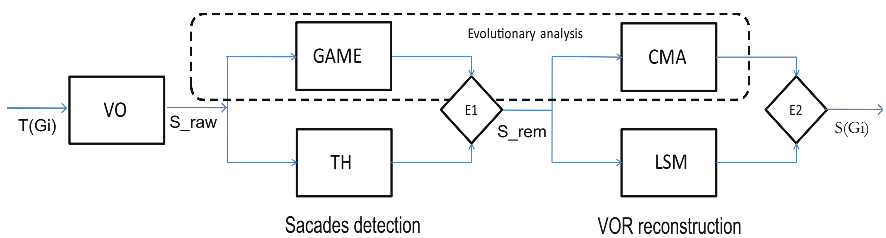
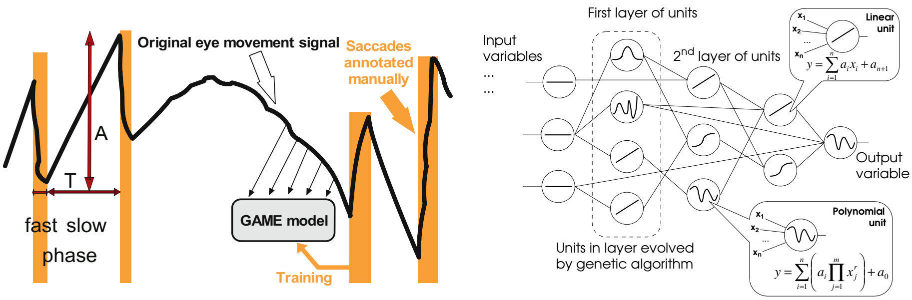


Fig. 1. Algorithm flow chart



(a) Slow phase GAME detection (b) GAME architecture

Fig. 2. (a): Slow phase velosity is defined as $SPV = A/T$. (b): GAME architecture.

2.2 Saccades GAME Detection

Group of Adaptive Models Evolution (GAME) [2] proceeds from the Group Model Data Handling (GMDH) theory. GMDH was designed to automatically generate model of the system in the form of polynomial equations. An example of inductive model created by GAME algorithm is depicted on the Figure 2(b). Similarly to Multi-Layered Perceptron (MLP) neural networks, GAME units (neurons) are connected in a feedforward network (model). The structure of the model is evolved by special niching genetic algorithm, layer by layer. Parameters of the model (coefficients of units' transfer functions) are optimized independently. Model can be composed from units of different transfer function type (e.g sigmoid, polynomial, sine, linear, exponential, rational, etc). Units with transfer function performing well on given data set survive the evolution process and form the model. Often, units of several different types survive in one model, making it hybrid. The eye movement signal is reconstructed in three steps: training data synthesis, identification of saccades and the signal reconstruction (removing the disturbances, filling out estimated values).

The first step is to prepare the training data. The key task is to annotate the saccades in the eye movement signal. For this purpose we implemented program which allows us to manually annotate the saccades in signal and save this information as a target signal for training - see Figure 2(a).

In the second step the data set is used to generate inductive models using the GAME method. Generated GAME model in fact acts as an automatic annotator of saccades. In the third step we reconstruct the signal using a script, which has two inputs (original eye movements signal and GAME model saccade estimation) and one output, the reconstructed signal. The method of reconstruction is based on information about signal character in the neighborhood of saccades. When saccade is estimated, the original signal is replaced with linear approximation of trends in the beginning and the end of saccade. The original eye movement signal with saccades (a), the output of the GAME model (b) and the reconstructed signal (c) is displayed on Figure 3.

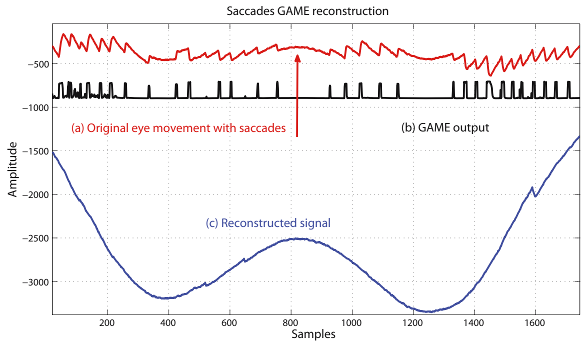
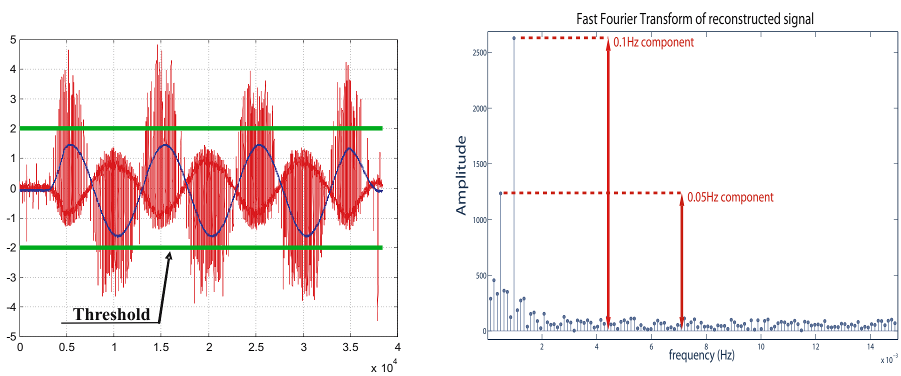


Fig. 3. The original signal (a), the output of GAME model (b) and reconstructed signal (c)

Thresholding approach. The GAME detection was evaluated against thresholding approach. Before the slow components were detected, the signal was preprocessed. First, the velocity was computed using derivative of horizontal eye movement channel. Afterwards the signal was filtered out implementing elliptic low pass filter. Taking advantage of signal statistical behaviour, the signal velocity was normalized in such a way that its mean was zero and standard deviation was one. In the next step the detection of slow phase was applied. It was based on horizontal velocity signal; the information of vertical signal was not regarded as useful. If we look at the typical example of eye movement development in time - see Figure 4(a), the prevailing pattern is nystagmus. In other words, the eye movement signal consists of series of triangles that are superposed on the isoline that is most similar to sinusoidal function. Artefacts presented in the signal such as noise are reduced by low pass filtering. Therefore after these assumptions the peaks of horizontal velocity signal detect the position of slow phases. Peaks were simply determinated by setting up the threshold as it is shown in Figure 4(a).



(a) Detection of Fast Phases

(b) Spectrum analysis

Fig. 4. (a): Thresholding method illustration. (b) FFT analysis after saccades reconstruction.

2.3 Signal CMA Reconstruction

The second part of VOR analysis focuses on signal reconstruction after fast phase removal. It is assumed that the source signal (which controls the rotation of the chair with the patient) is formed as a sum of sine waves (SOS):

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

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. 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 SOS 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.

Unfortunately, we do not have access to the output SOS signal described by Eq. 2. We have only the measured VOR signal, i.e. the segments of the output SOS signal that are left after filtering out the saccades from the eye-tracking signal. However, we can search for the unknown parameters a_i and ϕ_i of the output SOS 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, $v_j(t)$, $j = 1 \dots m$, 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 SOS signal by searching the $2n$-dimensional space of points $\mathbf{x}$, $\mathbf{x} = (a_1, \phi_1, \dots, a_n, \phi_n)$. Such a vector of parameters represents an estimate of the output SOS signal and we can compute the degree of fidelity with which the SOS corresponds to the VOR signal segments by constructing a loss function as follows:

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

where $\bar{v}_j$ is the mean value of the j-th VOR signal segment and $\bar{y}_j$ is the mean value of the current estimate of the output SOS signal related to the j-th segment

Subtracting the means $\bar{v}_j$ and $\bar{y}_j$ from the VOR signal segments $v_j(t)$ and SOS signal $y(t)$, respectively, we try to match the VOR signal segment to the corresponding part of the SOS 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.

Optimization Method. The parameter vector $\mathbf{x}$ is projected to the loss function via the estimate of the SOS signal $y(i)$ (and via the mean values $y_j(i)$). The evolutionary strategy with covariance matrix adaptation was chosen to minimize the objective function $L(x)$. CMA is very recent and progressive stochastic optimization algorithm [3]. 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.

3 Results

First, the proposed approach will be evaluated on synthetic data to decide which of these two algorithms (evalutionary and threshold approach) is more suitable for estimating VOR clinical important paratemers as gain or phase shift.

3.1 Synthetic Data

The tests were carried out on six signals ($S1-S6$) consisting of 1 to 6 sine components, i.e. the search was carried out in 2-, 4-, 6-, 8-,10- and 12-dimensional parameter spaces.

Generating VOR signal. First, for each sine component of the signal, the values of frequency and amplitude were generated. The ranges for individual parameters can be found in Table 1. This SOS 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 Figure 5 where signal S_2 consisting of two sinus components is shown ($f_1 = 0.05, a_1 = 100, f_2 = 0.1, a_2 = 50$). Signal duration was set to $80s$, sampling frequency of synthetic signal to $102.4Hz$ and saccade duration to $0.1s$.

First, let us review the success rates of both algorithms when estimating the parameters of the SOS signal with the number of components ranging from 1 to 6. In Table 2, the ratio a_i/a_i^S of CMA and Least-Square Multimodal Sinus fit (LSF) is reported for all 6 testing sets. Figure 5 shows example of CMA reconstructed signals S_2.

Lest-square fit is sensitive to noise and outliers therefore the CMA approach significantly outperformed the least-square method. The CMA ratio a_i/a_i^S is equal to one in all cases. In each run, the algorithms were allowed to perform 10,000 evaluations of

Table 1. Settings for parameters of artificial VOR signals S1-S6

	Frequency	Amplitude
	f_i	a_i
S_1	$\langle 0.1\rangle$	$\langle 50\rangle$
S_2	$\langle 0.05, 0.1\rangle$	$\langle 25, 50\rangle$
S_3	$\langle 0.05, 0.1\rangle$	$\langle 25, 50\rangle$
S_4	$\langle 0.01, 0.04, 0.070.1\rangle$	$\langle 25, 38, 50\rangle$
S_5	$\langle 0.01, 0.033, 0.055, 0.076, 0.1\rangle$	$\langle 25, 33, 38, 43, 50\rangle$
S_6	$\langle 0.01, 0.028, 0.046, 0.064, 0.082, 0.1\rangle$	$\langle 25, 30, 35, 40, 45, 50\rangle$

Table 2. Comparision of artificial VOR signal gain estimation

		f1	f2	f3	f4	f5	f6
S1	CMA	1.00	-	-	-	-	-
	LSF	1.16	-	-	-	-	-
S2	CMA	1.00	1.00	-	-	-	-
	LSF	1.26	1.44	-	-	-	-
S3	CMA	1.00	1.00	1.00	-	-	-
	LSF	0.26	0.39	0.99	-	-	-
S4	CMA	1.00	1.00	1.00	1.00	-	-
	LSF	0.28	0.29	1.37	1.14	-	-
S5	CMA	1.00	1.00	1.00	1.00	1	-
	LSF	0.24	0.37	1.70	1.46	1.28	-
S6	CMA	1.00	1.00	1.00	1.00	1	1.00
	LSF	0.27	0.27	2.03	1.77	1.58	1.42

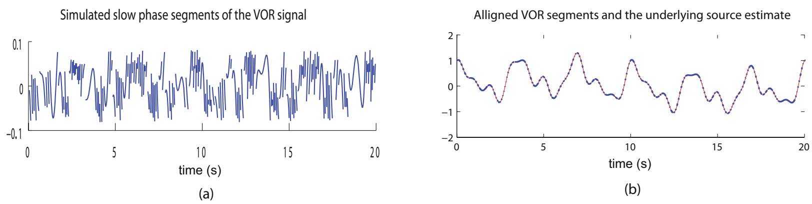


Fig. 5. (a) Simulated VOR signal with saccades removed. VOR signal segments aligned with the estimated SOS signal. (b) The parameters of the SOS signal are output of the algorithm.

the loss function and a particular run was considered to be successful if the algorithm found a parameter set with the loss function (3) value lower than 10^{-8}. Those learning parameters proved to be sufficient in order to precisely estimate the sinus components presented in the response of the vestibular organ.

3.2 Real-World Data

First of all, the fast phases were removed using GAME and thresholding approach (TH) for all 7 subjects $P1-P7$. Before comparing the perfomance of both methods we must address the task of evaluating the most suitable GAME architecture.

GAME topology selection. Three different models were used for saccades reconstruction. The model performance were evaluated on the base of spectrum analysis calculated by 4096 point FFT - one example is shown in Figure 4(b). Regarding neurons' transfer functions, Linear Config consists of linear units (linear and linearGJ neuron) only. NonLinConfig is composed of the followings transfer functions: PolySimpleNeuron,ExpNeuron, SigmNeuron, SinusNeuron a GaussianNeuron. The last NonLinearConfig ALL contains moreover polynomial units. 5 models were evaluated on the data of volunteer $P6$. The GAME topology comparison is depicted in Table 3 in

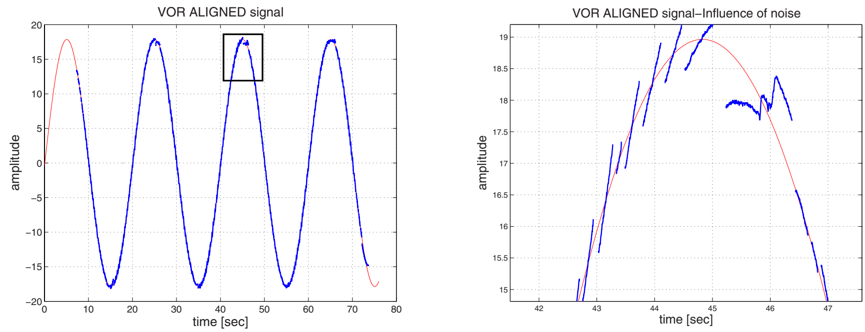
Table 3. Spectral results for three different GAMEs topology

c.	LinearConfig		NonLinearConfig		NonLinearConfig ALL	
model number	frequency[Hz]	amplitude	frequency[Hz]	amplitude	frequency[Hz]	amplitude
1	0,05	4,00E+005	0,05	1,60E+006	**0,05**	**1,20E+006**
	0,1	5,00E+005	0,1	1,50E+006	**0,1**	**1,16E+006**
	0,2	6,30E+004	0,18	8,50E-004	**0,22**	**2,60E+004**
	0,55	9,55E+004	0,23	4,00E-004	-	-
2	0,05	3,40E+005	0,05	1,30E+006	0,05	1,65E+006
	0,1	5,00E+005	0,1	1,32E+006	0,1	1,52E+006
	0,2	3,50E+004	0,15	8,00E+004	0,23	1,60E+004
	0,3	2,50E+004	0,23	3,50E+004	0,27	1,80E+004
	0,35	1,50E+004	0,3	2,50E+004	-	-
3	0,05	5,70E+005	0,05	9,90E+005	0,05	1,40E+006
	0,1	7,80E+005	0,1	1,04E+006	0,1	1,10E+006
	0,15	1,10E+005	0,28	5,00E+004	0,23	7,00E+004
	0,2	1,23E+005	-	-	0,28	4,00E+004
4	-	-	0,05	1,37E+006	0,05	1,08E+006
	0,1	3,40E+005	0,1	1,46E+006	0,1	1,13E+006
	0,23	1,00E+004	0,15	4,20E+004	0,23	1,13E+004
	0,35	2,00E+004	0,3	3,50E+004	0,28	2,80E+004
5	0,05	1,80E+005	0,05	1,23E+006	0,05	1,42E+006
	0,1	2,60E+005	0,1	1,18E+006	0,1	1,39E+006
	0,17	6,60E+004	0,15	5,70E+004	0,15	8,00E+004
	0,21	5,90E+004	0,23	2,80E+004	0,23	2,40E+004
	0,35	1,20E+004	0,35	3,40E+004	0,27	2,40E+004

terms of frequencies and their corresponding amplitudes found in FFT reconstructed spectrum. Only the fourth model with Liner Configuratin did not reveal 0.05 Hz component. GAME reconstruction also revealed other frequency components that were not presented in the SOS signal controlling the chair movements. However, their amplitudes are significantly less than amplitudes of 0.05 and 0.1 Hz components. In case of non-linear configurations, the difference is even in order of $10e^{-2}$. First model with NonLinearConfig ALLwas chosen for next analysis.

Next, the performance of saccades GAME and TH detection was evaluating using percentage rate of Sensitivity (Sn) and Specificity (Sp) for real signal $P6$. The saccades were annotated by four expert who determined the probability of saccades appearance. The average of those four annotation was taken as a final annotation. In case of GAME, $Sn = 87.8\%$, $Sp = 90.9\%$, in case of TH method, $Sn = xx$, $Sp = xx$. [!!Traditional thresholding techniques yielded better results in both performance criterions. GAME method requires more data in order to detect saccades more accurately.!!] Threshold method used also a priori information about the direction of chair rotation: when the chairs rotated to the right, only saccades exceeding positive threshold were processed (see Figure 4(a)) resulting in zero rate of false positives detections.

Furthermore, CMA reconstruction was carried out. One example of the resulting signal with one frequency component $f_1 = 0.05$ can be seen in Figure 6(a). We can see



(a) VOR alligned (b) Outliers Influence

Fig. 6. (a): Real-World VOR signal with slow components aligned. (b)Influence of noise in VOR aligned signal.

that all slow phases were precisely aligned in spite of some noisy segments that have been left after thresholding procedure. Even if not all fast phases were not detected producing some kind of noise, the evolutionary algorithms was able to cope with the outliers - see Figure 6(b).

Table 4 summarizes clinical parameters of real two frequency component signal ($f_1 = 0.05, f_2 = 0.1$): *(i)* gain G_i and phase shift ϕ_i for each component $f_i, i = 1,2$, *(ii)* total slow phase velocity $SPV = A/T$ defined in Figure 2(a) when the chair rotates to the left SPV_{left} and right SPV_{Right} and total mean SPV.

CMA ratio estimation worked correctly with exception of phase shift reconstruction. The minus sign means that the CMA phase estimation was shifted by π compared to original phase value. Volunteers P_6 and P_7 did not have their head fixed during VOR measurement. SPV's parameters of $P7$ suggest that for the used frequency components 0.05 and 0.1Hz the fixation was not so relevant. However, head fixations proved to be important for P_6 only pointing out to big intra-personality variability in VOR data.

Table 4. Comparision of real VOR signal parameters estimation: GAME, CMA and LSF

	P_1	P_2	P_3	P_4	P_5	P_6	P_7
G_1CMA	1.78	-3.25	-6.78	-4.70	-0.96	-6.52	-1.87
G_1^{LSF}	0.31	0.55	0.55	0.58	0.34	0.56	0.55
ϕ_1^{CMA}	1.42	-1.69	-1.31	-1.46	-1.66	-1.60	-1.97
ϕ_1^{LSF}	3.69	3.16	3.48	3.09	3.45	3.51	3.16
G_2^{CMA}	-0.63	1.45	2.97	2.65	0.41	3.20	0.90
G_2^{LSF}	0.25	0.48	0.55	0.54	0.30	0.55	0.48
ϕ_2^{CMA}	-1.88	1.14	1.43	1.31	1.14	1.32	0.92
ϕ_2^{LSF}	2.99	3.51	3.44	3.18	3.21	3.37	3.51
SPV_{left}	5.38	13.13	16.16	14.67	17.52	19.35	13.13
SPV_{rigth}	2.85	-9.41	-14.41	-10.14	-13.08	-22.39	-9.41
SPV_{mean}	4.03	11.14	15.23	12.25	15.15	20.97	11.14

There can be seen significant SPV's asymmetry of volunteer P1 when chair rotated to the left.

4 Conclusions

In this paper, a new method of VOR signal processing was introduced and experimentally evaluated both on artificially generated signals and on clinical real-world data. It relies on the right identification of the fast eye movements, saccades, that must be filtered out of the signal in advance. The simple and fast method based on GAME methodology for saccades removal was developed. The main advantage of the procedure lies in the simplicity of the basic algorithms and effectivity, even in the evaluation of irregular data such as patients suffering from vestibular organ disorders. However, a more examples for GAME learning is needed in order to make the GAME saccades detection suitable for use in the clinical situations.

Regarding phase and gain calculation, 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, the CMA proposed method directly estimates these parameters from the signal segments trying to align them with the underlying estimated mixture of sine waves.

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. Gordon, J., Furman, J.M.R., Kamen, E.W.: System identification of the vestibulo-ocular reflex: application of the recursive least-squares algorithm. In: Proc. of the Fifteenth Annual Northeast Bioengineering Conference, pp. 199–200 (1989)
2. Kordik, P.: Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution. PhD thesis, CTU Prague (2006)
3. Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation 9(2), 159–195 (2001)