EVALUATION OF AN ACOUSTIC INTERFACE FOR TREMOR ANALYSIS

Marian Weger, David Pirrò, Robert Höldrich
Institute for Electronic Music and Acoustics (IEM)
University of Music and Performing Arts, Graz, Austria
{weger,pirro,hoeldrich}@iem.at

ABSTRACT

We present the evaluation of a sonification approach for the acoustic analysis of tremor diseases. The previously developed interactive tool offers two methods for sonification of measured 3-axes acceleration data of patients' hands. Both sonifications involve a bank of oscillators whose amplitudes and frequencies are controlled by either frequency analysis similar to a vocoder or Empirical Mode Decomposition (EMD) analysis. In order to enhance the distinct rhythmic qualities of tremor signals, additional amplitude modulation based on measures of instantaneous energy is applied. The sonifications were evaluated in two experiments based on pre-recorded data of patients suffering from different tremor diseases. In Experiment 1, we tested the ability to identify a patient's disease by using the interactive sonification tool. In Experiment 2, we examined the perceptual difference between acoustic representations of different tremor diseases. Results indicate that both sonifications provide relevant information on tremor data and may complement already available diagnostic tools.

1. INTRODUCTION

Tremor is a movement disorder which produces involuntary rhythmic oscillation movements of a body part [1]. As it can be caused by various neurological diseases [2], a correct diagnosis is quickly needed to choose the right therapy. Each of these diseases evokes a specific movement pattern which can be recognized visually by specialized neurologists. This visual diagnosis, however, is unreliable and common approaches for additional ex-post analysis of videos or measured sensor data are time-consuming and can not be easily integrated into daily clinical practice. A sonification has the advantage that it provides an auditory representation which is continuously following the spectral characteristics of the tremor and thus allows to keep track of the time-dependent spectral structures. Real-time sonification of tremor movement data could therefore become a promising extension to already available diagnostic tools. Sonification has recently been successfully used for for therapy of Parkinsonian tremor [3,4].

In a follow-up to our previous research [5], we developed two sonification methods for tremor analysis which are described in detail in [6]. These are intended to be used interchangeably dependent on tremor characteristics and personal preference. Both sonifications and the corresponding graphical user interface are implemented in Pure Data [1]. The interface is targeted towards real-time use as a supplementary medical tool in order to improve diagnostic quality. It is employed to extract relevant features of the tremor signal, which are exposed aurally by the developed sonification algorithms. The resulting feature space is high dimensional and therefore predestined for aural rendering in preference to visual representations. This tool, however, aims not at providing definite answers nor a final diagnosis of the disease.

In a previous pilot study (see [6]), test participants were asked to identify the tremor diseases of patients by using the interactive sonification interface on pre-recorded movement data of 30 patients who divided equally into three different groups of diseases (Parkinsonian, Essential, and Psychogenic tremor). Participants used headphones and obtained prior training with the same set of patients. On average, participants reached 61 % correct diagnoses, which is far above chance (1/3). According to participating neurologists, the proposed interface facilitates an insight in the movement pattern of an examined tremor without visual tools. An interactive switch between sonifications did not improve overall sensitivity. However, test participants positively welcomed the possibility to parametrize the sonifications to personal preference.

In retrospective, this pilot study suffered from unclear data and immature experimental design: It was based on a small sample size, clinical reference diagnoses were not perfectly reliable, and test participants were trained with the same set of patients' tremor data as used in the experiment. The results are therefore of limited significance. As a consequence, we carried out an extended study which is presented here. It includes more test participants and a larger dataset of patients with confirmed diagnosis.

This article is structured as follows. After this introduction, Sec. 2 gives an overview of the technical setup for data acquisition and the two different sonification methods. One used frequency analysis (Vocoder sonification, Sec. 2.1), while the other is based on Empirical Mode Decomposition analysis (EMD sonification, Sec. 2.2). These sonifications were evaluated with real patients' data in an identification test with the interactive audiovisual user interface as well as in a triangle discrimination test (Sec. 3).

[1] Pure Data (Pd): http://puredata.info/

Finally, in Sec. 4, we summarize our findings and give an outlook on future work.

Accompanying sound examples can be found on the project web page [7]. These include stereo recordings of both sonifications with two patients of each tremor type.

2. SONIFICATION APPROACH

Both sonification methods share the basic technical setup as well as some fundamental data conditioning steps.

Movement data is recorded by 3-axis accelerometers [2] attached to the patient's hands and sampled at 1 kHz. Signals are recorded with CED Spike2 [3] and pre-processed in Matlab. A DC removal and 70 Hz low pass filter is applied on the acceleration signal in order to cover the typical frequency range of pathological tremor (predominantly $3-15$ Hz). Both left and right arm sensors are individually sonified.

As strong amplitude variations can occur between different measurements, Automatic Gain Control (AGC) is applied to the input signal at different stages in both sonifications. The measurement data is further conditioned by a Principal Component Analysis (PCA) [8]. In the context of the presented sonifications, the first principal component is projected on the multichannel data to retrieve the monophonic input signal $x[t]$ (see [5] for a more detailed description).

We will briefly describe the basic sonification algorithms. For further information, please consult [6].

2.1 Vocoder Sonification

The first sonification method is similar to a vocoder.

By using a sliding window FFT, the input signal is divided into 5 frequency bands ($2-4$ Hz, $4-6$ Hz, $6-9$ Hz, $9-13$ Hz, and $13-20$ Hz). Center frequencies and bandwidths have been selected based on experience with the spectra of different tremor types.

Then, the normalized energies in the individual bands are used to control the amplitudes of 5 sinusoidal oscillators which are tuned harmonically to each other, i.e., following the harmonic series (f_0 , $2 \cdot f_0$, etc.). A Frequency Modulation (FM) with the smoothed half-wave rectified input signal $x[t]$ can be applied optionally. This results in a time-varying fundamental frequency of $f_i(t)$ instead of a constant f_0.

The sum of the five oscillator signals is finally amplitude modulated by the variably smoothed half-wave rectified input signal $x[t]$.

The Vocoder sonification produces a harmonic complex, evoking a clear, optionally time-varying pitch percept (compare sound examples [7]). The time-varying timbral character resembles vocal formants whereas the overall amplitude modulation adds a rhythmic dimension.

2.2 EMD Sonification

The second sonification is based on Empirical Mode Decomposition analysis.

EMD was originally developed by [9] to analyze non-stationary and non-linear signals. The idea of EMD is that complex data sets can be decomposed into a finite (and often small) number of so-called Intrinsic Mode Functions (IMFs). Each IMF represents one mode of the signal. The higher the index of an IMF, the lower its frequency components.

In contrast to Fourier analysis where a signal is decomposed into a set of pre-defined base functions, the EMD obtains the base functions adaptively from the signal. A perfect reconstruction of the original signal is possible via summation of the contained IMFs and the resulting residual signal. The basic EMD algorithm is explained in [9–12].

For each individual IMF, the instantaneous phase, frequency, and amplitude can be obtained from the Hilbert transform. In conjunction with the EMD, this is called the Hilbert-Huang Transform (HHT) [9, 13]. The HHT has been proposed for tremor analysis in recent studies, e.g., [14–16].

For the sonification, only the first five IMFs of the input signal $x[t]$ are determined via EMD. Eventually, these IMFs are individually leveled by AGC.

Each IMF then controls the frequency and amplitude of an individual sinusoidal oscillator. Although both the instantaneous amplitude and the frequency can be computed at any time by using the Hilbert transform, we used the generalized zero-crossing method [17] for the frequency, as it was found to provide more stable results. The determined frequencies are then multiplied by a user-controlled constant factor to map the low tremor frequencies to the audible range. Each oscillator is then individually amplitude modulated by the smoothed half-wave rectified IMF signal itself, in order to display the original tremor frequency range as a superposition of rhythmic structures.

Finally, the output signal of the sonification is formed by the sum of these four signals.

Due to the specific time-varying characteristics of the tremor signals, the sonic result of the EMD-based sonification resembles the sound of singing birds (compare sound examples [7]). The register of each "bird" is dependent on the frequency and amplitude of the corresponding IMF.

3. EVALUATION STUDY

The presented sonification methods as well as the interactive interface were evaluated on the basis of pre-recorded movement data of 97 patients with confirmed diagnosis. The clinical tremor data were collected by the Medical University of Graz and UCL Institute of Neurology London between 2012 and 2013. Hand acceleration data were recorded for both hands simultaneously in rest (arms hanging) and posture (arms outstretched) condition. The patients divide unevenly into four tremor types: 24× Parkinsonian, 19× Essential, 36× Psychogenic, and 18× Dystonic tremor. To force an even distribution, we randomly se-

[2] Biometrics ACL300 (mass: 10 g, accuracy: ±2 % FS): `http://www.biometricsltd.com/accelerometer.htm` connected to CED 1401 interface.

[3] CED Spike2: `http://ced.co.uk/products/spkovin`

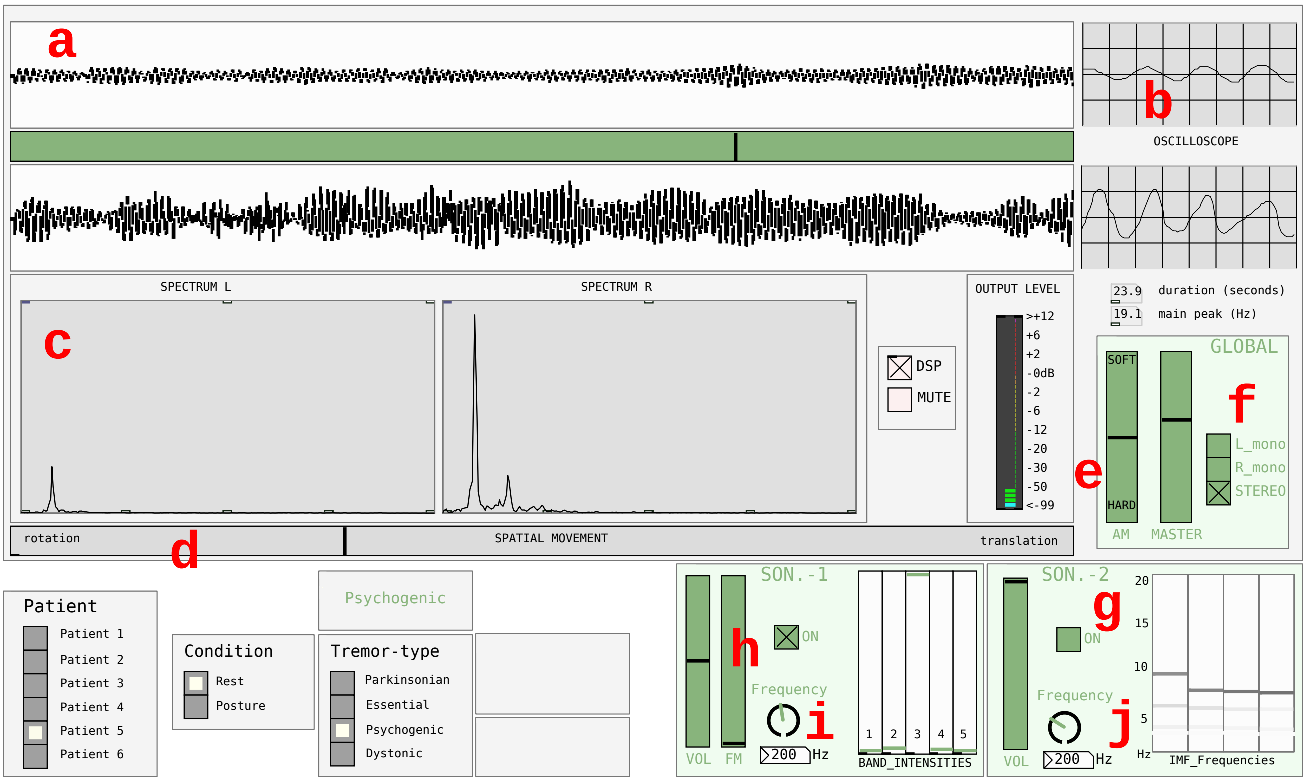

Figure 1. Graphical user interface (here in training mode). Annotations in red letters. Visual feedback: a) waveform, b) oscilloscope, c) FFT spectrum, each for left and right channel, d) ratio between rotational and translational movement. Global parameters: e) smoothing of the AM modulator, f) mono/stereo and left/right channel switch, g) sonification switch. Vocoder sonification: h) FM modulation, i) fundamental frequency. EMD sonification: j) frequency factor.

lected 18 patients of each type to form the reduced dataset of $18 \times 4 = 72$ patients.

Test participants for the extended study were recruited from an expert listening panel [18, 19], a group of musicians and sound engineers with experience in listening tests. Despite the target audience being neurologists, trained listeners have been chosen to ensure a best-case scenario and hence a more fair comparison of the results with the currently achieved diagnostic accuracy through visual and computer-aided ex-post analysis methods. Neurologists can acquire these abilities provided that they obtain appropriate ear training. The sonification was presented with closed headphones. Test participants were allowed to take handwritten notes throughout the experiments.

We carried out two experiments which are discussed below: a 4AFC identification task (Sec. 3.1) similar to the pilot study (cf., [6]) and a triangle test (also called 3AFC oddity task) which is described in Sec. 3.2.

3.1 Experiment 1: training and identification task

The basic design of Experiment 1 was similar to the pilot study [6]. It further addresses learning effects over several training sessions. An additional focus of this experiment was the interactive use of sonification parameters. The goal was to verify the results of the pilot study with a larger number of patients and tremor types.

3.1.1 *User Interface*

Apart from the sonic representation of the tremor data, a simple visualization is provided (see Fig. 1). It shows various visual information, such as waveform view, oscillogram, level meter, FFT spectrum, ratio between rotational and translational movement strength as well as band intensities and IMF frequencies for the individual sonifications.

Further, apart from standard controls such as volume, the interface features interactive access to a selection of sonification parameters. Globally for both sonifications, the smoothing of the AM modulator signal as well as stereo/mono playback can be controlled. In addition, each sonification allows individual access to oscillator frequency (base frequency for Vocoder sonification, multiplicative factor for EMD sonification) and dedicated gains for left and right arm sensors. Another slider provides control of the FM index for the Vocoder sonification.

3.1.2 *Procedure*

In Experiment 1, participants had to accomplish three sessions (S1–S3) of one hour on different days with approximately one week of pause in between. Each session started with approx. 30 minutes of free training where participants could freely listen to the sonifications of 8 patients per tremor type (24 patients in total) with disclosed diagnosis.

Participants were able to decide by themselves when they felt to be ready for the next part of the session, in which they had to identify the disease of 24 patients (6 per tre-

	Session 1	Session 2	Session 3	Overall (S1 and S3)
Percent correct answers	31.5 (10.2)	39.1 (7.3)	31.5 (10.0)	31.5 (9.9)
Confidence Interval CI95	26.9 − 36.4	34.2 − 44.1	26.9 − 36.4	28.2 − 34.9
Confidence (from 1 to 100)	38.8 (24.2)	43.3 (23.5)	39.9 (26.9)	39.4 (25.6)
Response time (seconds)	34.6 (25.2)	41.9 (27.1)	42.1 (33.9)	38.4 (30.1)

Table 1. Overview of the results from Experiment 1. Values in parentheses describe one standard deviation.

mor type) by using the interactive sonification interface. The task followed the same procedure as in the pilot study. Feedback was given after each trial.

For each participant individually, the dataset was split randomly into 3 subsets of 24 patients (6 per tremor type). These 3 subsets were presented in the training and testing parts of the 3 sessions. The presentation order of patients within each session and within each mode (training/ testing) was randomized across participants. In the first and in the last session, training was performed with subset A, while a completely new subset (B/C) was introduced for the testing part. In the second session, both training and testing were performed with the same subset (B):

	S1	S2	S3
Training	A	B	A
Testing	B	B	C

S1 and S3 are seen as a realistic scenario where known patients' characteristics were applied to unknown patients. In S2, the ability of memorizing specific patients' characteristics was tested, comparable to the procedure in the pilot study.

During the testing part of each session, participants were asked to submit a diagnosis (4AFC) and judge their confidence (values from 1 to 100) for each of the 24 patients. Feedback (reference diagnosis) was given after each trial. Sonification-specific parameters as well as the condition (rest/posture) could be additionally controlled by a motorized MIDI controller (Behringer BCF-2000).

3.1.3 Results

Overall results of the identification task are shown in Tab. 1. On average, the test participants achieved 31.5% correct answers for S1 and S3, while S2 led to 39.1 % correct answers. Confidence intervals in Tab. 1 are based on the inverse Beta CDF. The total percent correct answers lead to a discrimination index d' of 0.26 for S1/S3 and 0.49 for S2 [20].

The primary test results (percent correct answers) were analyzed by using a binomial test with one variable "disease" (4 levels), assuming a constant hit rate of $1/4$, sample size of 384 (24 patients $\times$ 16 participants), and significance level of p=0.05. For all three sessions, the achieved percent correct answers are significantly higher than chance (p=0.002 for S1/S3, p<0.001 for S2). The overall results of the first and third session are identical and therefore considered equivalent.

For further analysis of the results, contingency tables were created for the individual sessions (Tab. 2a–2c) as well as for the pooled results of S1 and S3 (Tab. 2d). The tremor types Parkinsonian, Essential, Psychogenic, and Dystonic, are referred to as Par, Ess, Psy, and Dys.

The higher percent correct responses in S2 compared to S1 and S3 is further supported by the higher contingency coefficient [4] (0.32 vs. 0.22 and 0.25), which means that there is a stronger relation between reference diagnosis and submitted diagnosis for S2 compared to both other sessions.

The pooled results of S1 and S3 (Tab. 2d) reveal a high confusion between Parkisonian and Psychogenic tremor (31.8 % Par falsely assigned to Psy, 28.1 % Psy falsely assigned to Par). The lowest false positive rate appeared between Parkinsonian and Dystonic tremor (16.1 % Par falsely assigned to Dys, 18.8 % Dys falsely assigned to Par).

A binomial test as above was performed on the results per tremor type. When looking at the pooled results of S1 and S3, only the percent correct identifications of Parkinsonian, Psychogenic, and Dystonic tremor were significantly higher than chance (Par: p=0.014, Psy: p=0.042, Dys: p=0.004), while Essential tremor could not be identified (p=0.107). However, the differences in sensitivity (Par: 0.32, Ess: 0.29, Psy: 0.31, Dys: 0.34) and F-measure (Par: 0.38, Ess: 0.36, Psy: 0.33, Dys: 0.42) are only small. The difference in percent correct responses between the individual tremor types therefore seems marginal.

We did not find any significant clusterization of test participants based on their individual contingency tables. Also, the results of those participants who primarily used the vocoder-based sonification did not significantly differ from the results of those wo preferred the EMD-based sonification (9 participants vs. 7 participants, 32.3 % vs. 30.4 % correct answers as average of S1 and S3). However, there were five top performers who achieved at least 33.3 % correct answers in every session (Average of S1 and S3: 36.6 % correct answers, ranging from 33.3 % to 47.7 %).

On average (pooled over S1 and S3), participants took 38.4 s (SD=30.1 s) to complete one trial and reported a confidence of 39.4 (SD=25.6).

3.1.4 Interactive use of sonification parameters

For the slider- or rotary-based controls (AM smoothing, Vocoder sonification base frequency and FM index, EMD sonification frequency factor), we observed that participants found their preferred values already within the first training session. These parameters were rarely changed during diagnosis.

The sonification switch also exhibited convergent behavior for all but one participant who was continuously switch-

[4] Pearsons's contingency coefficient: $K = \sqrt{\frac{Chi^2}{N+Chi^2}}$, where N is the sample size.

R \ D	Par	Ess	Psy	Dys
Par	**32.3**	20.8	31.3	15.6
Ess	28.1	**22.9**	25.0	24.0
Psy	28.1	20.8	**34.4**	16.7
Dys	24.0	22.9	16.7	**36.5**
Sum	112.5	87.5	107.3	92.7

(a) Session 1. $K = 0.22$.

R \ D	Par	Ess	Psy	Dys
Par	**36.6**	16.7	21.9	21.9
Ess	21.9	**38.5**	19.8	19.8
Psy	18.8	18.8	**40.6**	21.9
Dys	13.5	21.9	27.1	**37.5**
Sum	93.8	95.8	109.4	101.0

(b) Session 2. $K = 0.32$.

R \ D	Par	Ess	Psy	Dys
Par	**32.3**	18.8	32.3	16.7
Ess	14.6	**35.4**	30.2	19.8
Psy	28.1	18.8	**27.1**	26.0
Dys	13.5	21.9	33.3	**31.3**
Sum	88.5	94.8	122.9	93.8

(c) Session 3. $K = 0.25$.

R \ D	Par	Ess	Psy	Dys
Par	**32.3**	19.8	31.8	16.1
Ess	21.4	**29.2**	27.6	21.9
Psy	28.1	19.8	**30.7**	21.4
Dys	18.8	22.4	25.0	**33.9**
Sum	100.5	91.1	115.1	93.2

(d) Sessions 1 & 3. $K = 0.19$.

Table 2. Contingency tables for the three sessions and the pooled results of S1 and S3. Values describe $\%$ of submitted diagnoses D. R is the reference diagnosis. Correct answers (main diagonal) are highlighted. K is Pearson's contingency coefficient.

ing between sonifications throughout the whole experiment. The average amount of sonification switches per trial decreased from 0.57 switches in S1 to 0.27 in S2 and 0.08 in S3. Both other toggles (mono/stereo and rest/posture) never converged during the experiment (1.22 and 1.46 switches per trial in S1 vs. 1.42 and 2.21 in S3, respectively). Even in S3, 7 participants regularly switched to mono, and 14 participants regularly switched between rest and posture position.

3.1.5 Discussion

The quick convergence of the slider- and rotary-based parameters (already during the first training session) indicates that they did not contribute to the test participants' judgments of tremor diseases. During informal interviews, test participants mentioned that the base frequency for the Vocoder sonification and the frequency factor for the EMD sonification had only minor effect on the information conveyance and were only changed in respect to personal preference. Also frequency modulation for the Vocoder sonification did not add additional information but was often applied for diversification of the otherwise relatively monotonic sounds. For the smoothing factor of the amplitude modulation, however, participants had to find a compromise: while less smoothing leads to higher rhythmic resolution, the stronger transients implicate bandwidth expansion and therefore deteriorated perception of frequency changes. While fixed default values for the mentioned parameters might work for most users, an optional parametrization for expert users is assumed to be advantageous. The interactive switch between sonifications has been used regularly by at least one test participant, which further emphasizes its necessity.

The stereo/mono (connected with left/right switch) and rest/posture switch were used excessively by most test participants. During informal interviews, many test participants mentioned that a comparison between left and right arm sensor as well as between rest and posture condition revealed critical information for tremor type identification. These parameters are therefore regarded as essential for acoustic tremor analysis.

The discrimination between 4 tremor types by using the interactive sonification interface seems to be excessively demanding (31.5 $\%$ correct diagnoses). According to their hand-written notes as well as informal interviews, most of the test participants were focused on finding specific sound characteristics that are unique to a specific tremor disease. However, there were outlying patients in all tremor categories which led to confusion. This observation is also reflected in the submitted confidence values which did not correlate with the correctness of the submitted diagnoses.

The better performance in the second session shows that the test participants were able to remember specific patients' sound characteristics and corresponding reference diagnosis (39.1 $\%$ correct responses in S2 vs. 31.5 $\%$ in S1/S3). This also explains the promising results of the pilot study: in both pilot study and S2 of Experiment 1, the training was performed with the same set of patients. In real life, however, a diagnosis is only needed for unknown patients. In the more realistic conditions of S1 and S3, the test participants could not rely on patient-specific sound characteristics, but only on more general tremor-specific structures.

There was no improvement after training (S1 vs. S3). We assume that the training parts of the second and third sessions were needed to recapitulate the individual tremor types' sound characteristics. Further, we noticed that the new set of patients in S2 did not always match the participants' notes, which forced them to start again from scratch. We suppose that training time was not sufficient for finding the possibly very subtle differences between different tremori.

Both sonifications were used equally often and provided similar results. This finding is consistent with the pilot study and shows once more that an optional choice could be important in order to adapt to the users' individual preferences.

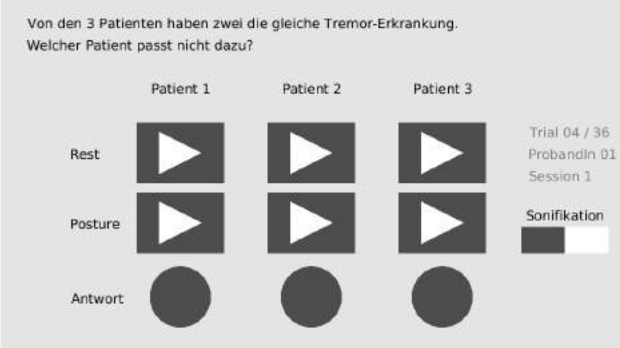

Figure 2. Graphical user interface for Experiment 2. Colors are inverted and converted to grayscale for this article.

A further interpretation of participants' confidence and response time is not possible due to the high spread.

The results of Experiment 1 suggest that an absolute identification of tremor diseases with the proposed interface might be too difficult. However, the contingency table in Tab. 2d shows that some pairs of tremori were less often confused with each other than other pairs. In order to verify these pairings, we performed a second experiment.

3.2 Experiment 2: triangle test

After having accomplished the training and absolute diagnosis task of Experiment 1, the same 16 participants performed a discrimination task comprising of 72 trials which were equally split in two sessions of approx. one hour each, performed on different days.

3.2.1 Procedure

Before the actual test, all participants performed a mandatory training with the interactive sonification interface in order to recapitulate the different tremor characteristics.

In a triangle test setup (3AFC oddity task), only a simple graphical user interface without visual tools was shown (see Fig. 2). Sonifications of three different patients were presented: two of the same tremor type and one of a different type. For each triple of patients, participants had to indicate the odd, i.e., the patient whose tremor type differed from the others'. After each trial, participants obtained feedback consisting of the reference diagnosis for all three patients, with the correct answer highlighted.

For each participant and session individually, the patients of the trials were chosen randomly so that every patient appeared once as the odd stimulus and twice inside the pair of similar stimuli. The three patients per trial as well as the 36 trials per session were presented in random order.

Participants could freely listen to the sonifications of all three patients in rest as well as in posture condition. There was no restriction in time. Switching between Vocoder and EMD sonification was still possible. Sound was always presented in stereo in order to limit the possibilities for interaction to a minimum.

3.2.2 Results

The overall results of Experiment 2 are shown in Tab. 3 for the 6 different pairs of tremor types individually. On average, test participants achieved $41.2\,\%$ correct judgments, yielding a sensitivity index d' of 0.95 (based on [21]). Test participants achieved above-average results only when discriminating Dystonic tremor from Parkinsonian or Psychogenic tremor (45.8 and $49.5\,\%$ correct, respectively).

Type pair	% correct	d'	response time
Par vs. Ess	39.6	0.88	59.7
Par vs. Psy	38.5	0.81	51.5
Par vs. Dys	45.8	1.25	58.4
Ess vs. Psy	38.0	0.73	50.3
Ess vs. Dys	35.9	0.55	54.5
Psy vs. Dys	49.5	1.41	52.7
Overall	41.2	0.95	54.5

Table 3. Overview of the results per pair of tremor types in Experiment 2: average percent correct responses, sensitivity index d' [21], and response time in seconds.

For statistical analysis of the results, a one-tailed binomial test was performed for each pair of tremor types (sample size n=192 (pooled over all subjects), chance level p=1/3, significance level: p=0.05). According to the binomial test there was a highly significant perceptual difference between Parkinsonian and Dystonic as well as between Psychogenic and Dystonic tremor (p<0.001 for both pairs). A marginally significant difference between Parkinsonian and Essential tremor was found (p=0.040). Between the other three pairs of tremor types no significant perceptual difference could be found (p=0.074 for Par/Psy, p=0.097 for Ess/Psy, and p=0.244 for Ess/Dys).

We observed that participants usually chose their preferred sonification at the beginning and did not change it anymore. User interaction is therefore not further investigated.

The average response time (time to accomplish one trial) was 54.5 seconds (SD=35.0).

In order to examine similarities between individual test participants, we performed a hierarchical cluster analysis. Individually for each participant, the total percent correct responses for the pairs of tremor types form a 6-dimensional vector. Based on these vectors and Ward's method of minimum inner squared euclidean distance within clusters [22], test participants divide equally into two groups of 8. Results are shown in Tab. 4. While Cluster 1 inhibits only small differences in percent correct responses between tremor pairs, the participants of Cluster 2 were exceptionally good at discriminating Dystonic tremor from Parkinsonian and Psychogenic tremor ($62.5\,\%$ and $54.2\,\%$, respectively). Interestingly, in Cluster 1, all participants except two preferred the EMD sonification, while in Cluster 2 all participants except one preferred the Vocoder sonification.

Differences in overall percent correct responses between participants grouped by cluster or sonification preference, however, were not significant.

Five of the 16 participants (see "Top 5" in Tab. 4) reached a level of above $45\,\%$ correct responses (on average: $47.8\,\%$, ranging from $45.8\,\%$ to $50.0\,\%$). Three of those used the Vocoder Sonification while two worked with the EMD Sonification.

Type pair	Cluster 1	Cluster 2	Top 5
Par vs. Ess	44.8	34.4	35.0
Par vs. Psy	38.5	38.5	50.0
Par vs. Dys	29.2	62.5	50.0
Ess vs. Psy	39.6	36.5	48.3
Ess vs. Dys	42.7	29.2	43.3
Psy vs. Dys	44.8	54.2	60.0
Average	39.9	42.5	47.8

Table 4. Average results per pair of tremor types in Experiment 2 for the two main clusters and the 5 best performing participants. Values describe $\%$ correct answers.

3.2.3 *Discussion*

The overall discrimination index d'=0.95 in Experiment 2 shows that the oddity task was indeed easier than the identification task of Experiment 1 (d'=0.26). However, overall percent correct responses are still too low for direct applications in the medical context.

A cluster analysis on the basis of the test participants' results leads to an almost perfect division into the users of the different sonifications. While the Vocoder sonification seems to provide similar discrimination performance for all tremor pairs, users of the EMD sonification achieved extraordinary results for the pairs Par/Dys and Psy/Dys (at the expense of lower percent correct responses for the other pairs). Each sonification might emphasize different aspects of the observed tremori. If those aspects were known, participants might be able to combine both sonifications in order to improve their sensitivity. The interactive change between both sonifications therefore seems reasonable and might facilitate the construction of a coherent picture based on the observed phenomena in order to make informed decisions.

The five top performers of Experiment 2 (Tab. 4) also achieved high identification rates in Experiment 1 (on average: $33.8\,\%$ correct, ranging from $31.3\,\%$ to $37.5\,\%$, two of them among the Top 5). This leads to the assumption that at least some test participants have found consistent tremor-specific sound characteristics and were able to apply these complex models in an identification task as well as in a discrimination task.

4. CONCLUSIONS AND OUTLOOK

In the presented study, we evaluated tremor disease identification and discrimination by means of interactive sonification.

In Experiment 1, test participants were asked to indicate the tremor type of unknown patients by analyzing pre-recorded movement data with the interactive sonification interface. Average percent correct diagnoses were significantly above chance for patients with Parkinsonian, Psychogenic, and Dystonic tremor, but not for Essential tremor. Overall sensitivity index d' was 0.26.

In Experiment 2, participants were able to distinguish Parkinsonian tremor from Essential and Dystonic tremor with sensitivity significantly above guessing rate. A significant perceptual difference was also shown between Dystonic and Psychogenic tremor. Overall d' was 0.95.

During the first experiment, participants were allowed to manipulate critical sonification parameters. All test participants found their preferred sonification and corresponding parameters quickly. Still, the permanent comparison between left and right arm sensor as well as between rest and posture condition provided useful information. The interactive change of these parameters facilitates the construction of a coherent image of the observed tremor and allows informed decision making.

The proposed sonification interface is not meant to replace currently available diagnostic tools, but rather to complement them. The fact that both sonifications led to different results in the triangle test lets us conjecture that they could supplement each other in a beneficial way, especially if additional knowledge from other diagnostic tools is available. A combination and cross-checking of different sources of information is essential for an efficient and correct diagnosis.

It is obvious that both sonifications as well as the corresponding user interface need to be substantially revised. This revision could be based on the test participants' qualitative feedback which we collected. Informal interviews with test participants revealed that the offered sonification parameters might not be sufficient. One of the top performing participants complained that the first IMF can become obtrusive and often masks the important fine structure of the higher IMFs; however, at the same time, it also contains critical information and can therefore not be omitted. Individual volume control for the IMFs may therefore an option for future improvements.

Further, we observed that some test participants successfully used the spatial movement visualization for decision making: while Parkinsonian and Dystonic tremor showed almost only rotational movement, Essential and Psychogenic tremor sometimes caused sudden switches towards a translational movement. At this stage, however, the ratio between rotational and translational movement is not conveyed by the sonification. A prior approach mapping this movement quality to a chorus effect (see [6]) did not work due to the parameter's transient behavior. A promising option might be a mapping of the translation/rotation-ratio to relative pitch.

Finally, we think that the collected strategies of the test participants could be brought in correspondence with Music Information Retrieval (MIR) analysis descriptors applied to sonifications and raw movement data. Thus, the sonification qualities which have proven to be perceptually relevant in tremor segregation could be enhanced in the next development steps. In general, the re-design of the developed sonifications could be informed by the findings of this combined qualitative and quantitative analysis.

5. ACKNOWLEDGMENTS

We would like to thank Petra Schwingenschuh and the Medical University of Graz, who kindly provided us with the clinical tremor data and thus made this research possible.

6. REFERENCES

[1] K. T. Wyne, “A comprehensive review of tremor,” *JAAPA*, vol. 18, no. 12, pp. 43–50, 12 2005.

[2] G. Deuschl, P. Bain, and M. Brin, “Consensus statement of the Movement Disorder Society on Tremor. Ad Hoc Scientific Committee,” *Mov. Disord.*, vol. 13 Suppl 3, pp. 2–23, 1998.

[3] M. Schedel, D. Weymouth, T. Pinkhasov, J. Loomis, I. B. Morris, E. Vasudevan, and L. Muratori, “Interactive sonification of gait: Realtime biofeedback for people with parkinson’s disease,” in *5th Interactive Sonification Workshop (ISon)*, 12 2016.

[4] A.-M. Gorgas, L. Schön, R. Dlapka, J. Doppler, M. Iber, C. Gradl, A. Kiselka, T. Siragy, and B. Horsak, *Short-Term Effects of Real-Time Auditory Display (Sonification) on Gait Parameters in People with Parkinsons’ Disease – A Pilot Study*, 2017, pp. 855–859.

[5] D. Pirrò, A. Wankhammer, P. Schwingenschuh, R. Höldrich, and A. Sontacchi, “Acoustic interface for tremor analysis,” in *Proc. International Conference on Auditory Display (ICAD)*, 02 2012.

[6] M. Weger, D. Pirrò, A. Wankhammer, and R. Höldrich, “Discrimination of tremor diseases by interactive sonification,” in *5th Interactive Sonification Workshop (ISon)*, 12 2016.

[7] IEM. (2017) Acoustic interface for tremor analysis. [Online]. Available: http://tremor.iem.at

[8] J. Shlens, “A tutorial on principal component analysis,” in *Systems Neurobiology Laboratory, Salk Institute for Biological Studies*, 2005.

[9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[10] A. Zeiler, R. Faltermeier, I. R. Keck, A. M. Tomé, C. G. Puntonet, and E. W. Lang, “Empirical mode decomposition - an introduction,” in *International Joint Conference on Neural Networks (IJCNN)*, 7 2010, pp. 1–8.

[11] M. C. Peel, G. E. Amirthanathan, G. G. S. Pegram, T. A. Mcmahon, and F. H. S. Chiew, “Issues with the Application of Empirical Mode Decomposition Analysis,” in *International Congress on Modelling and Simulation (MODSIM)*, 12 2005, pp. 170–176.

[12] R. T. Rato, M. D. Ortigueira, and A. G. Batista, “On the HHT, its problems, and some solutions,” *Mechanical Systems and Signal Processing*, vol. 22, pp. 1374–1394, 8 2008.

[13] N. E. Huang and S. S. Shen, Eds., *Hilbert-Huang Transform And Its Applications*. World Scientific, 2005.

[14] S. Ayache, T. Al-ani, and J.-P. Lefaucheur, “Distinction between essential and physiological tremor using hilbert-huang transform,” *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 44, no. 2, pp. 203–212, 2014.

[15] S. S. Ayache, T. Al-ani, W. H. Farhat, H. G. Zouari, A. Creange, and J. P. Lefaucheur, “Analysis of tremor in multiple sclerosis using Hilbert-Huang Transform,” *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 45, no. 6, pp. 475–484, 12 2015.

[16] I. Carpinella, D. Cattaneo, and M. Ferrarin, “Hilbert-Huang transform based instrumental assessment of intention tremor in multiple sclerosis,” *J Neural Eng*, vol. 12, no. 4, 8 2015.

[17] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, “On instantaneous frequency,” *Advances in Adaptive Data Analysis*, vol. 1, no. 2, pp. 177–229, 2009.

[18] R. Höldrich, H. Pomberger, and A. Sontacchi, “Recruiting and evaluation process of an expert listening panel,” in *Fortschritte der Akustik (NAG/DAGA 2009)*, B. Rinus, Ed., Rotterdam (Niederlande), 03 2009.

[19] M. Frank and A. Sontacchi, “Performance review of an expert listening panel,” in *Fortschritte der Akustik (DAGA 2012)*, H. Hanselka, Ed., Berlin (Deutschland), 09 2012.

[20] M. J. Hacker and R. Ratcliff, “A revised table of d’ for m-alternative forced choice,” *Perception & Psychophysics*, vol. 26, no. 2, pp. 168–170, 1979.

[21] B. J. Craven, “A table of d’ for m-alternative odd-man-out forced-choice procedures,” *Perception & Psychophysics*, vol. 51, no. 4, pp. 379–385, 1992.

[22] J. H. W. Jr., “Hierarchical grouping to optimize an objective function,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.