

Triaxial Rehabilitative Data Analysis Incorporating Matching Pursuit

^{1,2}Tracey K. M. Lee, ³Kee-Hao Leo, ⁴Saeid Sanei, ⁵Effie Chew and ⁵Ling Zhao

¹School of Electrical and Electronic Engineering, Singapore Polytechnic, Singapore: tlee@sp.edu.sg

²School of Information Technology, Monash University, Malaysia

³School of Mechanical and Aeronautical Engineering, Singapore Polytechnic, Singapore: leokeehao@sp.edu.sg

⁴Faculty of Engineering and Physical Sciences, University of Surrey: s.sanei@surrey.ac.uk

⁵Division of Neurology, University Medicine Cluster, National University Hospital, Singapore: effie_chew@nuhs.edu.sg
ling_zhao@nuhs.edu.sg

ABSTRACT

The continuing drive for better rehabilitative healthcare hinges on the availability of sensor data which can be shared and analysed. This leverages on the widespread communications network to provide an integrated health management environment. For this paper, we delineate our current work in sensorizing rehabilitative tests of upper limb movements. Where previously we applied data driven analysis, we now employ time-frequency methods to provide a better analytical basis for our derivations. The use of Matching Pursuit algorithm in biological signals has concentrated on brain signals and much less on human motion. Thus we contribute to efficacy of the algorithm by employing it on rehabilitative data collected from widely available sensors. We describe how we obtained the parameters based on pre-analysing an available data set. By selecting the most useful signal constituents and applying this to signal denoising, we are able to better classify the condition of a patient automatically - which shows encouraging promise in the quest for integrative healthcare.

Index terms - Matching pursuit, rehabilitation, accelerometer, instrumented objects.

1. INTRODUCTION

The seemingly never-ending demand for connectivity has seen the proliferation of consumer devices like smartphones, tablets and laptops. Together with lifestyle monitoring devices like smartwatches and fitness trackers, these have seen the proliferation of networks and sensors and driven down their costs and sizes. Thus initiatives like Health 4.0 [1] seek to enhance traditional clinical practices which devolves the point of consultation and care away from traditional centralised facilities like clinics and hospitals. However there is still the need for proper clinical tests, assessments and treatment but these can then be administered remotely and can even be customized if necessary. Data can be collected, and analysed on a large scale to detect trends and anomalies. This being the case, we need to use established rehabilitative tests which have undergone rigorous ratification processes. In a networked environment, this is crucial to ensure that proper care is consistently dispensed to achieve consistent results for data analytics.

Although our focus here is on a particular test our approach can be *easily generalized* to other tests of a rehabilitative nature. As these tests pertain to human movements and the usage of objects employed for the activities of daily living (ADL), our work will have a wide range of applications and our findings are important for further development of devices which will be used for long-term monitoring and assessment of rehabilitation progress.

In a recent review [2], one of the most commonly used tests is the Action Research Arm Test (ARAT) devised by

Lyle [3]. It is a test of performance designed to measure the recovery of upper limb function subsequent to injury to the cerebral cortex. It can be used to evaluate treatment outcomes as well as monitor its progress. Furthermore it can be conducted quickly and is also dependable. The test prescribes the movements of several objects used in the ADL in specific ways. In doing so, the limb movements in grasp, grip and pinch actions will be assessed. In many cases, sensors are directly attached to the subject in an attempt to record signals from such fine movements. However a common difficulty among these methods are that they are intrusive and may impede motion. On the other hand, using video gives readings that are intrinsically noisy and the effects of variable lighting and occlusions have to be dealt with.

Our setup involves implanting sensors into the objects used in rehabilitation with the accruing benefits:

i) we have the ability to sense fine motion and the force applied by the subject and ii) the need to mount sensors on the person is done away with.

In Section 2 we describe the motivation for our approach as well as the background material. Section 3 outlines our physical setup followed by the signal analysis theory in Section 4. Our experimental results are presented in Section 5 and we summarize our discussion with conclusions in Section 6.

2. ASSESSING LIMB FUNCTION AND SENSORS

In this section we outline the motivation for our work, and present the case for having bespoke instrumented objects be deployed in standardised rehabilitative tests.

2.1 Assessing limb function

When designing tests of limb function and movement, there needs to be a protocol for their administration so as to ensure repeatable, quantitative and objective measurements.

Presently the prevalence of these tests are scored visually, which interposes a degree of subjectivity and does not allow subtle motions to be noticed. Besides, the unvarying nature of these assessment activities bring on inattentiveness and human errors. All these motivate for automated test monitoring by electronic means, by instrumenting the objects used in these tests.

Yozbatiran et al. [4] further regulated the ARAT by explicitly stating the positions of the objects and the dimensions of the supporting furniture in the test. They also specified the scoring by recording the timing and quality of the movement performed. However the measure of quality was mostly descriptive in nature being based on visual observations.

In [5], Lee et al. reported on work done with the instrumented device outlined in this paper, deploying healthy patient simulators. For the sake of continuity in discussion, parts of their paper have been used here.

We seek to learn the factors that account for the assessment scores because these scores are awarded subjectively and should be incorporated into an electronic health system to provide a seamless flow of information. This can be done by supervised learning on features obtained from electronic signals generated during these tests.

These features can be derived by analysing biomedical signals, decomposing them into their constituent parts. However, signals derived from human movements are mainly nonstationary and thus fixed basis signal decomposition cannot give good results. Indeed, in [6] a chirp and impulse signal gave the same spectral diagram.

To overcome this, one approach is to use data driven decomposition [5], but the main drawback has been the inability to describe these signals analytically.

In time-frequency (TF) analyses for biomedical signals, Baumgartner et al. [7] suggest that with proper parameters, Matching Pursuit (MP) can give optimal results as well as an analytical TF description of the structure in a signal. However as discussed in [7] as well as in [8], much of prior work involving MP was for electroencephalograms (EEG). There are much less applications of MP for human movements, indicating there is scope of work in this area.

2.2 Sensors for motion sensing

Typical electronic consumer devices have a variety of motion sensors built into them. Smartphones have gyroscopes and accelerometers built into them as do smart watches, fitness trackers and gaming devices.

Using these low cost devices to measure movements seem an attractive option as the data interface is standardized and the results can be obtained quickly as can be seen in the literature. However consumer devices have the following problems: i) they have short life spans, where typically a new model appears the next year. ii) the software required to for data interfacing can change with upgrades in an operating system. iii) there is no control over the power source or product size and weight. Smartphones are getting larger while other devices are getting smaller and difficult to charge. A bespoke design is called for, which allows us to produce results that go beyond mere proofs-of-concept to actual products.

This provides the motivation to instrument an object used in the ARAT, by incorporating accelerometers and force sensors which are two types of sensors not often used together. To reiterate the advantages: these are the ability to sense fine motion and the force applied by the subject, the need to mount sensors on the person is done away and the online application is guaranteed through adjustable sampling frequencies. We describe our setup in the next section.

3. EXPERIMENTAL SETUP AND TRIAL CONDUCT

We now outline how Test4 of the ARAT Grasp Subtest is implemented and describe some initial results. This test requires grasping a cube (wooden) measuring 7.5 cm all around. This object which we will denote the Cube, is displaced from a given point, directly to another point. As shown in Fig. 1, the three main parts of our instrumented object system are a:

- i) Set of resistive force sensors used to measure forces imposed on the faces of the Cube.
- ii) Triaxial accelerometer to measure motion.
- iii) Microcontroller to convert force sensor and accelerometer readings, and transmitting the data to a workstation.

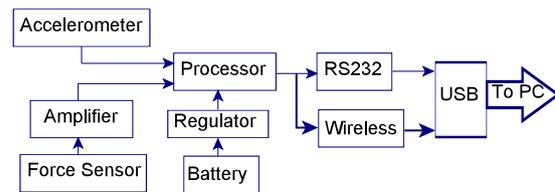


Fig. 1. Hardware block diagram of the embedded sensor system in the Cube. Dotted lines indicate optional portions.

The sensor readings are sampled at 30 times per second so that a maximum frequency of 15 Hz can be recorded reliably. Pre-filtering is not done to prevent removing important information.

3.1 ARAT scoring and test subjects

In Fig. 2 we note the Cube being gripped, held vertically and moved. The ARAT scoring uses a four point scale, from 3 for satisfactory completion to 0 which is non-completion. A score of 3 indicates completion of the task within 5 seconds with appropriate hand, arm and posture movements detailed in [4].

A score of 2 is given when the subject completes the task “with great difficulty and/or takes abnormally long time”, from 5 to 60 seconds.

For a score of 1 which indicates partial completion, the timing would be greater than 60 seconds. Also being able to just grasp, hold and lift the Cube would be sufficient to warrant this score.

However a score of 0 indicates any of the following: i) inability to perform any part of the task within 60 seconds. ii) inability to grasp the Cube within the time period. iii) subject does not use the fingers to grasp the Cube or use another hand or mechanical support to manipulate the Cube.

In all, 34 patients who have had a history of stroke and undergone rehabilitation participated in the trial. This was conducted in a hospital over a period of 60 days.

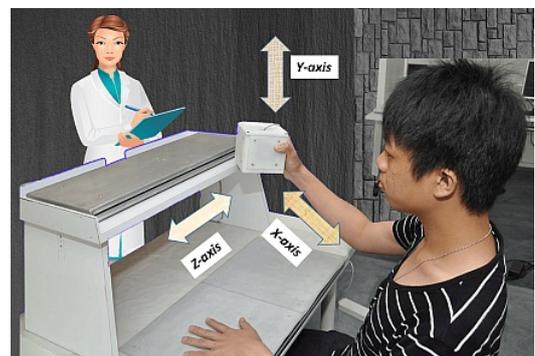


Fig. 2. ARAT Cube orientation with respect to observer, in background.

Each patient would perform a set of ARAT motions in one session, up to 3 iterations, if possible. For each session, the data is recorded continuously. It is then segmented manually into its component trials. A significant point to note is that the score is awarded on a *session* basis which implies some kind of averaging is done over the trials. Besides, the sessions were recorded over a period of time and different therapists performed the scoring so some variability is induced in the scores, even though briefings were conducted.

3.2 Qualitative results

In Fig. 3 we plot the signals from the *force* sensors recorded for two subjects with a score of 1 to show how the sensors can detect nuanced movements. These are lines with *magenta* ‘.’ and *blue* ‘◇’ markers. Initially they take on values that are near zero as they record the force exerted by the fingers on the Cube grasping surfaces.

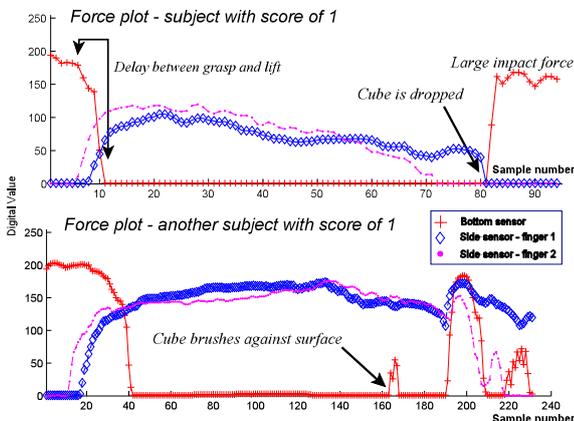


Fig. 3. Force sensor plots for score of 1 - top- Cube is dropped, not placed, bottom- Cube brushes against surface but is placed, not dropped. Marker with red + is bottom sensor plot, other markers are force exerted by fingers.

The line with the *red* ‘+’ marker indicates the force imposed on the bottom sensor by the weight of the Cube when it is at rest. As the Cube is lifted, the value goes to zero and this serves to indicate the *beginning* and *end* of a move. This timing cue permits automatic segmentation of the signals so as to obtain an accurate measurement of the period of the movement. For this signal in Fig. 3, another observation is that the subject may *drop* rather than *place* the Cube on the platform. Also the bottom of the Cube may briefly touch another object during the move.

4. METHODOLOGY

The MP theory will be described in this section, along with the identification of the important parameters used.

4.1 Matching Pursuit Algorithm

We cover the MP algorithm, based on the work by Mallat and Zhang [9]. It decomposes a signal in a suboptimal, nonlinear and iterative manner. The result is a sum of basis functions (or atoms) selected from a highly redundant set of functions, termed a dictionary. Each iteration process selects the atom that gives the largest dot product with the signal or its residual. Subtracting the signal from the atom gives a residual, to which the same process is applied with another atom until a stopping criterion is reached.

In our application, we use Gabor atoms, which are Gaussians modulated by a sinusoid and described by:

$$g_\gamma(t) = K(\gamma)e^{-\pi((t-u)/s)^2} \cos(\zeta t + \varphi) \quad (1)$$

where s, u, ζ, φ are the scale, time translation, frequency of modulation and phase. K is a normalizing constant so that $\|g_\gamma\| = 1$. The scale s determines the extent of the atom. For a signal f with N samples, the decomposition at iteration m can

be expressed as:

$$(2)$$

$$f = \sum_{n=0}^m C_n g_n + R^m f$$

where C_n is $\langle R^n f, g_n \rangle$ the dot product of the residue $R^n f$ at step n with the atom g_n selected at that step and $R^0 f = f$. This atom is chosen from a dictionary:

$$D = \{g_1, g_2, \dots, g_G\}$$

which is an $N \times G$ matrix and the $N \times 1$ vectors g are made up of values of the atom from (1). The derivation of number of atoms G in the dictionary is explained in Section 4.2.

This dictionary is overcomplete in that it has more atoms than is needed to explain the signal structure, so an optimization process is needed to select those the relevant atoms. For this version of MP we use, the atom is selected in a greedy manner according to:

$$g_n = \underset{g_\gamma \in D}{\operatorname{argmax}} |\langle R^n f, g_\gamma \rangle|$$

where at every iteration, the selected atom is the one which gives the largest absolute value, of the dot products between the atoms in the dictionary and the signal residue.

After decomposition, as explained in [8], the Wigner-Ville distribution (WVD) can be used to display the energy distribution of the signal across the TF domain. Because of the manner of decomposition, the cross terms effect of the WVD for other types of signals is absent.

We now describe how initial data analyses are used to set the parameters s, u, ζ and for the MP decomposition

4.2 Parameter settings for Matching Pursuit

We also know from our previous work, the movement of interest lies in the frequency range of 2 Hz to 10 Hz. We develop a script to generate the frequencies of interest automatically and here we use 50 frequencies, of which 80% lie in the 2-10 Hz range. The frequencies are linearly distributed with two breakpoints, to which a 5th order polynomial was fitted to generate the final frequencies.

The plot of the frequencies generated with a typical Gabor waveform is shown in Fig. 4.

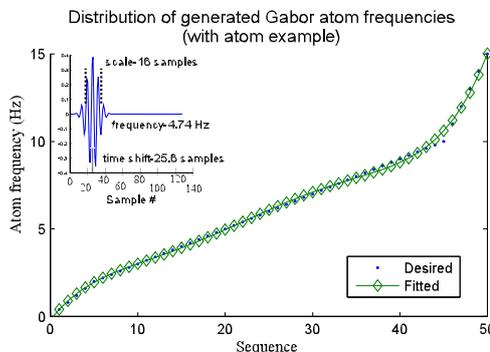


Fig. 4. Plot of generated Gabor atom frequencies, desired and fitted with 5th order curve. Inset shows a typical Gabor waveform.

As we are working with fixed length dictionaries, this length has to be determined as the trials capture varying length signals. Of the 78 trials, a statistical analysis shows that the mode of signal length is 72 samples. Since the MP works optimally for powers of two, we use a sample length N of 128 which corresponds to a time of 4.27 seconds, which is more

than enough to complete a movement. In our application, the signals do not have sudden changes and so we ignore the phase ϕ . Signals shorter than this are zero-padded while longer ones are truncated.

The scales $s = 2^j, 1 \leq j \leq \log_2 N$. We set the translation u to one-tenth of 128 samples, bearing in mind that in terms of time, we are sampling the signal at 0.427 (128/30Hz)/10 seconds. With these, our dictionary consists of 3958 atoms.

We use 40 iterations to decompose a signal and consider that the residual $R^m f$ after these iterations as noise. The considerations for deciding on the threshold for noise is complex, but here we let the denoised signal for $m = 40$ be:

$$\hat{f} = \sum_{n=0}^{40} C_n g_n \quad (3)$$

Another benefit in the use of Gabor atoms is that the signal can be reconstituted to its original sample length by just removing the extraneous atoms.

5. INITIAL ANALYSIS AND RESULTS

In this section we present some initial results followed by the results of denoising in conjunction with classifying a move.

5.1 Time Frequency plot

A typical Wigner-Ville TF plot is shown in Fig. 5. The darker areas of the plot denote regions where the energy distribution is higher. So for example, the x -axis plot shows a burst of energy at approximately the 11 sample mark across frequencies from 0 to about 10 Hz. There is also a relatively prominent wave of about 9 Hz lasting from sample 7 to 36.

The z -axis on the other hand, shows bursts of low frequency energy at equal intervals at approximately samples 6, 11, 15 and 21 which last for approximately 2 sampling periods.

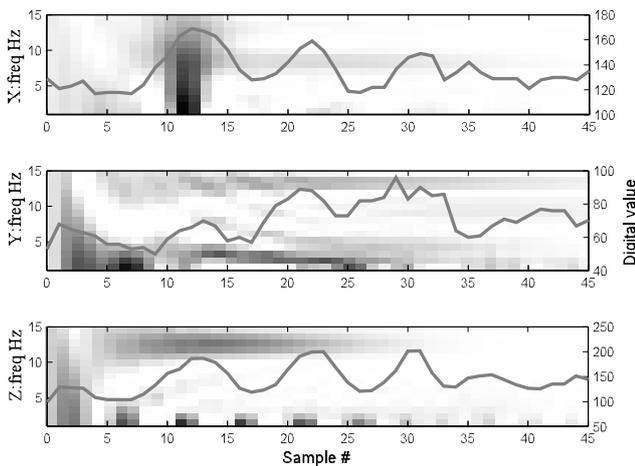


Fig. 5. Time frequency plot with the signal superimposed. Darker areas denote higher energy regions. From top, signals are x, y, z axes. Left y-axes denote the frequencies in Hz, right y-axes the amplitude of the signal superimposed, in 8 bit digital values.

Thus the triaxial signal structure is clearly mapped out in the plots and gives much insight into the evolution of the signal.

5.2 Significant axis of movement

In our earlier work [10] we employed features based on a single signal to train a classifier which automatically assigned the scores based on pre-computed thresholds. We now use the

the Waikato Environment for Knowledge Analysis (WEKA) [11] tool to learn a hierarchical tree based classifier.

Of the three axes of movement, the z -axis data which corresponds to the side-to-side movement of the Cube (as shown in Fig. 2), showed results more correlated with the score. On particular, we observed that the signal's coefficient of variation, which is the standard deviation divided by the mean value gave useful results. We proceed from there by using three different z -axis signal features. First, by taking away the signal mean MEANS, we compute the root mean squared (RMS) value of the AC component of our signals. This quantity gives an indication of the energy spent in the movement. We separate the RMS value of the signal by denoising it as in (3) to obtain the component explained by the signal, RMSS and that of noise, RMSN.

In our data set, an identification (ID) code is assigned to each trial and has the form SCC_MM_T where S is P for our subjects who were patients, CC the subject code, MM the movement type, which has a value of TS for our subjects and T being the trial number, 1 to 3. Some results are shown in Table 1.

Table 1 RMS signal/noise value using MP denoising of accelerometer readings for patients with score 1 and 2. Only the 3rd or z -axis shown.

Subject	Axis	RMS signal/noise	score	
P27_TS_1	3	14.45	5.47	1
P27_TS_1	3	14.73	6.25	1
P27_TS_1	3	12.78	5.12	1
P28_TS_2	3	8.85	3.92	2
P28_TS_2	3	9.11	5.23	2
P28_TS_2	3	11.46	5.54	2

Next we present the results of our automated scoring. Because there were not enough volunteers, only two score 1 sessions were recorded. One of them got this score because the Cube was improper handled. It should be noted that out of 78 trials, 31 trials were scored at 3, 38 scored at 2, 6 scored at 1 and 3 scored at 0, so that there is considerable statistical bias in results. This is depicted in the results shown in the last row of the confusion matrices in Tables 2 and 3.

Table 2 Confusion matrix on per-trial scoring using the RMSS attribute. Values are original (underlined> then denoised. Bottom row shows the number of trials receiving the score. used.

Actual \ Predict	3	2	1	0
3	<u>22</u> / 24	<u>6</u> / 8	<u>2</u> / 2	<u>1</u> / 1
2	<u>7</u> / 7	<u>32</u> / 29	<u>1</u> / 1	<u>2</u> / 0
1	<u>2</u> / 0	<u>0</u> / 0	<u>3</u> / 3	<u>0</u> / 1
0	<u>0</u> / 0	<u>0</u> / 1	<u>0</u> / 0	<u>0</u> / 1
# actual scored	31	38	6	3

From this, we note that the accuracy for trial scoring which uses unfiltered signals is (total correct/total trials) = $(22+32+3)/78 = 73\%$. For MP denoised signals, we have $(24+29+3+1)/78 = 71\%$ so a slight improvement comes from denoising, but not statistically significant. The resulting decision tree (which is not shown) has 4 binary levels and 5 leaves.

Next, using the RMSS, RMSN and MEANS features in Table 3, for the unfiltered and MP denoised signals

respectively, we have $(24+35+2)/78 = 78\%$ and $(28+35+2)/78 = 83\%$ so now the MP denoised signal gives a better classification rate, significantly better than using one feature alone. The classification tree used, in Table 4 has 5 binary levels and 8 leaf nodes.

Table 3 Confusion matrix on per-trial scoring using RMSS/RMSN/MEANS attributes. Values are original (underlined) then denoised. Bottom row shows the number of trials receiving the score.

Actual \ Predict	3	2	1	0
3	<u>24</u> / 28	<u>3</u> / 3	<u>3</u> / 2	<u>3</u> / 1
2	<u>7</u> / 2	<u>35</u> / 35	<u>1</u> / 4	<u>0</u> / 0
1	<u>0</u> / 0	<u>0</u> / 0	<u>2</u> / 0	<u>0</u> / 0
0	<u>0</u> / 1	<u>0</u> / 0	<u>0</u> / 0	<u>0</u> / 2
# actual scored	31	38	6	3

In summary, the steps to automatically score a test using MP are:

- i) Preprocess the accelerometer signals using MP.
- ii) Compute the MEANS, RMSS and RMSN values.
- iii) Classify the signal according to the following set of rules in Table 4 based on the tree obtained.

Table 4 Classification tree of automated scoring using the RMSS/RMSN/MEANS features.

```

MEANS <= 127.67
| RMSS <= 7.17: score: 3
| RMSS > 7.17 score: 2
MEANS > 127.67
| RMSN <= 2.48 score: 3
| RMSN > 2.48
| | RMSS <= 9.56
| | | MEANS <= 130.65 score: 2
| | | MEANS > 130.65
| | | | MEANS <= 134.12 score: 3
| | | | MEANS > 134.12 score: 2
| | | RMSS > 9.56
| | | | RMSN <= 7.7 score: 0
| | | | RMSN > 7.7 score: 3
    
```

6. CONCLUSIONS

In summary, we attempted to automatically award a score to a subtest from Test4 of the ARAT by using data from a triaxial accelerometer. The z-axis, which corresponds to the side by side movement of the Cube, appears to be visually more prominent to the assessor. We compared the effectiveness of signal denoising using MP on the following features of the z-axis accelerometer signal: the root mean squared value signal, its noise and the mean signal value, singly and in combination. The denoising operation here differs from traditional filtering which is a convolution across all frequencies. Here it can be more selective and is an area to explore.

It is worth recalling from Section 3.1 that we are undertaking to objectify what is basically a subjective rating, awarded by different scorers without a way of normalizing results. This shows the possibility of obtaining features from a subjectively scored rehabilitative test to derive an objective assessment. This data can then be aggregated, used as input

for other kinds of analytics, as part of an integrated healthcare scheme.

Subsequent work would require the analyses of other accelerometer signals, as well as other types of time frequency features for more discriminative power. There is also the need to gather more subjects with a more even distribution of scores for better validation of the results. A more extensive briefing would be needed on the scoring procedures for clinicians serving as scorers.

ACKNOWLEDGEMENT

This work was funded by the Ministry of Education of Singapore under grant number 2010MOE-IF-005. We thank the authors of Wavelab 8.50 for making their software available. We also thank the encouraging and useful comments from the anonymous reviewers which have greatly improved the quality of the paper.

REFERENCES

1. C. Thuemmler and C. Bai, *Health 4.0: How virtualization and big data are revolutionizing healthcare*, Springer International Publishing, 2017.
2. L. Santisteban et al., "Upper limb outcome measures used in stroke rehabilitation studies: A systematic literature review," *PLoS ONE*, 11(5) e0154792, 2016.
3. R. C. Lyle, "A performance test for assessment of upper limb function in physical rehabilitation treatment and research," *Int. J. Of Rehab. Research*, vol.4, no.4, pp.483-492, 1981.
4. N. Yozbatiran, L. Der-Yeghiaian and S. C. Cramer, "A standardized approach to performing the Action Research ArmTest," *Neurorehabilitation and Neural Repair*, vol.22, no.1, pp.78-90, 2008.
5. T. K. M. Lee, J. G. Lim, S. Sanei, and S. S. W. Gan, "Advances on singular spectrum analysis of rehabilitative assessment data", *J. Med. Imaging Health Inf.*, vol. 5, pp. 350-358, 2015.
6. M. Sandsten, "Time-frequency analysis of time-varying signals and non-stationary processes", Lund University, Centre for Mathematical Sciences, available online at: http://www.maths.lth.se/matstat/kurser/masm26/2016/TIME_FREQkompndie.pdf, 2016.
7. C. Baumgartner et al., "Discussion of time-frequency techniques in biomedical signal analysis: A tutorial review of similarities and differences", *Methods of information in medicine*, vol. 52, no. 4, pp. 297-307, 2013.
8. S. Sanei, *Adaptive Processing of Brain Signals*, Wiley, 2013.
9. S. G. Mallat, Z. Zhang, "Matching pursuits with time-frequency dictionaries", *IEEE Transactions on Signal Processing*, vol. 41, Issue 12, pp. 3397-3415, 1993.
10. T. Lee, S. Sanei, K. H. Leo, Y. H. Wong, E. Chew and Z. Ling, "Automated Scoring of Rehabilitative Tests with Singular Spectrum Analysis", *EUSIPCO 2015*, Sep. 2015.
11. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, "The WEKA Data Mining Software: An Update", *SIGKDD Explorations*, vol.11, no.1, 2009.