Validation of computer models for evaluating the efficacy of cognitive stimulation therapy

Tuan D Pham

Linköping University Post Print

N.B.: When citing this work, cite the original article.

The original publication is available at www.springerlink.com:

Tuan D Pham , Validation of computer models for evaluating the efficacy of cognitive stimulation therapy, 2015, Wireless personal communications, (), , 1-14.

<http://dx.doi.org/10.1007/s11277-015-3017-7>

Copyright: Springer Verlag (Germany)

<http://www.springerlink.com/?MUD=MP>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-128597>

Validation of Computer Models for Evaluating the Efficacy of Cognitive Stimulation Therapy

Tuan D. Pham

Aizu Research Cluster for Medical Engineering and Informatics
Research Center for Advanced Information Science and Technology
The University of Aizu
Aizuwakamatsu, Fukushima, 965-8580, Japan
E-mail: tdpham@u-aizu.ac.jp

Abstract. The notion of using computational methods for evaluating cognitive stimulation therapy (CST) based on the synchronized recording of photoplethysmographic (PPG) signals of care-givers and participants offers an objective and cost-effective analysis in health care to improve the patient's quality of life. While computer models are promising as a useful tool for such a purpose, a question of interest is how the model reliability, which is the degree to which an assessment tool produces stable and consistent results, can be established. This paper addresses this issue with the application of dynamic-time warping and resampling to measure the performance of two PPG features known as the largest Lyapunov exponent and linear predictive coding, which have been applied for studying the efficacy of CST. The potential success of this computerized evaluation can be a precursor to the development of a personalized e-therapy system that operates on mobile devices.

Keywords: Cognitive stimulation therapy · Cognitive decline · Model performance assessment · Therapeutic communication · Dynamic-time warping · Photoplethysmograph · Largest Lyapunov exponent · Linear predictive coding.

1 Introduction

Computers, Internet technology and mobile devices are being progressively utilized for psychotherapeutic communication, and even a step beyond e-therapy that uses the Internet to allow simultaneous (synchronous) and time-delayed (asynchronous) communication between an individual and a therapist [1]. Given problems still existing for the implementation of telehomecare technology [2],

This manuscript is the revised and extended version of the following conference paper: Pham, T.D. (2015) Dynamic-time-warping analysis of feature-vector reliability. In *2015 Int. Conf. Information Science and Applications*, Springer LNEE 339, 235-241.

wireless-communication-based health systems that can remotely provide constant care and treatment to patients are being accepted as the way to do with future healthcare [3, 4].

Several therapeutic interventions have been developed to work directly with people with dementia on an individual or group basis, and also indirectly with family and social-care professionals to improve communication and quality of life for people suffering from the cognitive decline [5]. In fact, skills needed for effective communication with people with dementia have been explored, and include factors that influence the communication process and therapeutic relationships between nurses and patients [6]. Communication in the form of encouraging talking, body language, physical contact, and active listening is thought to be fundamental to the provision of good dementia care. Furthermore, symptoms of depression and anxiety are common in people with dementia and mild cognitive impairment. Although treatment of these symptoms is widely recommended in guidelines, the best way to carry out the treatment is still not clear. While drugs are thought to have limited effectiveness in this context and may involve the risk of side effects, psychological treatments can be applied as an alternative to improve the mental wellbeing and cognitive function of people with cognitive impairment [7].

While traditional cognitive training interventions are delivered by humans, a recent review concluded that computer-based cognitive interventions are comparable or better than paper-and-pencil cognitive training approaches [8]. This review suggests that the utilization of computerized technology offers an effective and labor-saving method for improving and maintaining the quality of life and confidence of the individual with age-related impairment in cognitive function. In general, cognitive stimulation therapy is found to be less expensive than usual care with respect to benefits in cognition and quality of life. There is also evidence showing that cognitive stimulation therapy (CST) can be more cost-saving than dementia medication (<http://www.cstdementia.com/page/cost-effectiveness>). Because the evaluation of the cost-effectiveness of psychosocial interventions in dementia is becoming increasingly important, assessing the efficacy of CST can even further contribute to the cost effectiveness.

Based on the motivation of the importance of the use of CST for mental health, photoplethysmograph (PPG) was applied for the pattern analysis of short-term effects on efficacy in a caregiver on the daily provision of therapeutic treatment to aging people with dementia using the feature vectors of the largest Lyapunov exponent (LLE) values and spectral distortion. PPG is an optically obtained plethysmogram as a volumetric measurement of an organ, often obtained by using a pulse oximeter which illuminates the skin and measures changes in light absorption [9]. A pulse oximeter monitors the perfusion of blood to the microvascular layer of the tissue of the skin into which infrared light is emitted [10]. PPG technology has been pervasive as a low-cost, non-invasive, and flexible tool for physiological analysis in medicine and health, such as cardiology [11, 12], paediatric intensive care [13], hypertension [14], and depression [15]. In particular, wearable devices of photoplethysmographic sensors have been pro-

gressively developed [16, 17], making the PPG technology more attractive for its applications in telemedicine and e-health. Recent reports indicate that the use of PPG is rapidly attracting attention from the biomedical research community and industry, because it can be practically utilized to measure cardiac activity with a color camera [18].

Our recent work [19] attempted to synchronize the measurements of PPG signals of the care-giver and participants with dementia during the CST session. To generate control samples for comparison, the PPG signals of the participants before and after the therapy were also recorded. The influence of the CST therapy over the participants were analyzed using the methods of spectral distortion measures and phylogenetic tree reconstruction. However, given the outcome obtained from the PPG-based pattern analysis, a question of interest is: How reliable are the analysis results since they are difficult to be validated by human response? This is the motivation of the present study that presents the application of the dynamic time-warping (DTW) and resampling for establishing the uncertainty of the computerized assessment of CST efficacy. With the availability of wearable wireless PPG sensors [20], the potential applications and validation of the PPG technology to CST are expected to offer a low-cost computer-aided healthcare and treatment to patients on a model that is tailored to the individual patient by selecting appropriate and optimal therapies based on the effective influence of the professional over the patient's mental content.

The rest of this paper is organized as follows. Section II presents the feature extraction of PPG signals, including the largest Lyapunov exponent and linear predictive coding. These features have been found useful for classification of mental subjects using finger PPG [15, 19]. Section III describes the framework of DTW, which can be applied to compare the similarity between the PPG-based feature vectors of the care-giver and participants to establish the model assessment. Experimental results are presented and discussed in Section IV. Finally, Section V is the conclusion of the research findings.

2 Feature Extraction of PPG Signals

2.1 Largest Lyapunov exponent

Given a time series or sequence of length N , the first step is to reconstruct the phase space of the dynamical system using the time-delay method [21]. Let m and L be the embedding dimension and time delay (lag). The reconstructed phase space can be expressed in matrix form as

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_M)^T \quad (1)$$

where $\mathbf{X}$ is matrix of size $M \times m$, $M = N - (m-1)L$, and $\mathbf{X}_i = (x_i, \dots, x_{i+(m-1)L})$ which is the state of the system at discrete time i .

$$d_j(0) = \min_{\mathbf{X}_{j^*}} \|\mathbf{X}_j - \mathbf{X}_{j^*}\| \quad (2)$$

where $|j - j^*| > MP$ where MP is the mean period which is the reciprocal of the mean frequency of the power spectrum.

The basic idea is that the LLE (λ_1) for a dynamical system can be defined as [22]

$$d(t) = c e^{\lambda_1 t} \quad (3)$$

where $d(t)$ is the average divergence of two randomly chosen initial conditions at time t , and c is a constant that normalizes the initial separation between neighboring points.

By the definition given in Eq. (3), the j pair of nearest neighbors can be assumed to diverge at a rate measured by λ_1 as follows:

$$d_j(i) \approx c_j e^{\lambda_1(i\Delta t)} \quad (4)$$

where $d_j(i)$ is the distance between the j pair of nearest neighbors after i discrete-time steps which is $i\Delta t$, Δt is the sampling period of the time series, and c_j is the initial separation between two neighboring points.

Taking the logarithm of both sides of Eq. (4), giving

$$\ln[d_j(i)] \approx \lambda_1(i\Delta t) + \ln(c_j) \quad (5)$$

where $d_j(i) = \|\mathbf{X}_j(i) - \mathbf{X}_{j^*}(i)\|$. Eq. (5) gives a set of approximately parallel curves, one for each j ($j = 1, \dots, M$). If these curves are approximately linear, their slopes represent the LLE (λ_1). The LLE can be computed as the slope of a straight-line fit to the average logarithmic divergence curve defined by

$$s(i) = \frac{1}{i\Delta t} < \ln[d_j(i)] >_j \quad (6)$$

where $< \cdot >_j$ denotes the average over all values of j .

2.2 Linear Predictive Coding

Let $x(n)$ be a value of a sequence at time n . The estimate of $x(n)$, denoted by $\hat{x}(n)$, can be calculated as a linear combination of the past p samples of the sequence, which can be expressed as [23]

$$\hat{x}(n) = \sum_{k=1}^p a_k x(n-k) \quad (7)$$

where the terms $\{a_k\}$ are called the linear prediction coefficients.

The prediction error $e(n)$ between the observed value $x(n)$ and the predicted value $\hat{x}(n)$ can be defined as

$$e(n) = x(n) - \hat{x}(n) = x(n) - \sum_{k=1}^p a_k x(n-k) \quad (8)$$

The prediction coefficients $\{a_k\}$ can be optimally determined by minimizing the sum of squared errors

$$E = \sum_{n=1}^N e^2(n) = \sum_{n=1}^N \left[x(n) - \sum_{k=1}^p a_k x(n-k) \right]^2 \quad (9)$$

To solve (9) for the prediction coefficients, we differentiate E with respect to each a_k and equate the result to zero:

$$\frac{\partial E}{\partial a_k} = 0, \quad k = 1, \dots, p \quad (10)$$

The result is a set of p linear equations

$$\sum_{k=1}^p a_k r(m-k) = r(m), \quad m = 1, \dots, p \quad (11)$$

where $r(m)$ is the autocorrelation of $x(n)$, that is

$$r(m) = \sum_{n=1}^N x(n) x(n+m) \quad (12)$$

Equation (11) can be expressed in matrix form as

$$\mathbf{R} \mathbf{a} = \mathbf{r} \quad (13)$$

where $\mathbf{R}$ is a $p \times p$ autocorrelation matrix, $\mathbf{r}$ is a $p \times 1$ autocorrelation vector, and $\mathbf{a}$ is a $p \times 1$ vector of prediction coefficients. Hence

$$\mathbf{a} = \mathbf{R}^{-1} \mathbf{r}. \quad (14)$$

The cepstral coefficients can be directly derived from the LPC parameters using the following recursive procedure [23]

$$c_0 = \ln(G), \quad (15)$$

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m} \right) c_k a_{m-k}, \quad 1 \leq m \leq p, \quad (16)$$

$$c_m = \sum_{k=1}^{m-1} \left(\frac{k}{m} \right) c_k a_{m-k}, \quad m > p, \quad (17)$$

where G is the LPC gain, whose squared term is given as [24]

$$G^2 = r(0) - \sum_{k=1}^p a_k r(k). \quad (18)$$

Spectral distortion measures are designed to compute the dissimilarity or distance between two (power) spectra [25] (the power spectrum of a signal describes how the variance of the data is distributed over the frequency components into which the signal may be decomposed, and the most common way of generating a power spectrum is by using a discrete Fourier transform) of the two feature vectors, originally developed for comparison of speech patterns [23]. Two methods of spectral-distortion measures were used in this study, based on their popular applications in signal processing: Itakura distortion (ID), and log spectral distortion (LSD) [23].

Consider two signals S and S' , and their two spectral representations $S(\omega)$ and $S'(\omega)$, respectively, where ω is normalized frequency ranging from $-\pi$ to π .

The Itakura-Saito distortion (ISD) between S and S' is defined as [26]

$$ISD(S, S') = \int_{-\pi}^{\pi} \left[\frac{|S(\omega)|^2}{|S'(\omega)|^2} + \log \frac{|S'(\omega)|^2}{|S(\omega)|^2} - 1 \right] \frac{d\omega}{2\pi}, \quad (19)$$

where

$$|S(\omega)|^2 = \frac{\sigma^2}{|1 + a_1 e^{-j\omega} + a_2 e^{-j2\omega} + \dots + a_p e^{-jp\omega}|^2}, \quad (20)$$

where σ and a_i , $i = 1, \dots, p$, are the gain and i th linear-predictive-coding coefficients of the p th-order LPC model [23], respectively (in digital signal processing, linear prediction is often called linear predictive coding (LPC) that estimates future values of a discrete-time signal as a linear function of previous samples, used for representing the spectral envelope of a digital signal in a compressed form).

$$ID(S, S') = \min_{\sigma' > 0} ISD \left(\frac{\sigma^2}{|A(\omega)|^2}, \frac{\sigma'^2}{|A'(\omega)|^2} \right) = \log \int_{-\pi}^{\pi} \frac{|1 + a'_1 e^{-j\omega} + a'_2 e^{-j2\omega} + \dots + a'_p e^{-jp\omega}|^2}{|1 + a_1 e^{-j\omega} + a_2 e^{-j2\omega} + \dots + a_p e^{-jp\omega}|^2} \frac{d\omega}{2\pi}, \quad (21)$$

where $\frac{\sigma^2}{|A(\omega)|^2}$ and $\frac{\sigma'^2}{|A'(\omega)|^2}$ are two LPC spectra of two given autoregressive models of $S(\omega)$ and $S'(\omega)$, respectively.

The distortion defined in Eq. (21) is known as the Itakura distortion. It is also known as the log-likelihood ratio distortion or the gained-optimized Itakura-Saito distortion that can be derived as follows [27, 28]:

$$ID(S, S') = \log \frac{\mathbf{a}^T \mathbf{R}' \mathbf{a}}{\sigma'^2}, \quad (22)$$

where σ'^2 is the prediction error of S' produced by the linear predictive coding (LPC) [23], $\mathbf{a}$ is the vector of LPC coefficients of S , $\mathbf{R}'$ the LPC autocorrelation matrix of S' . It is shown that $ID(S, S') \neq ID(S', S)$, hence to make the measure symmetrical, a natural expression of its symmetrized version, denoted as $ID_s(S, S')$, is

$$ID_s(S, S') = \frac{ID(S, S') + ID(S', S)}{2}. \quad (23)$$

The log spectral distortion distance (LSD) between two signals S and S' is defined as [23]

$$LSD(S, S') = \int_{-\pi}^{\pi} |V(\omega)|^m \frac{d\omega}{2\pi}, \quad (24)$$

where $m=1$ gives the mean absolute log spectral distortion, $m=2$ defines the root-mean-square log spectral distortion that has been widely applied in speech signal processing and also used in this study, when m approaches ∞ Eq. (24) reduces to the peak log spectral distortion, and $V(\omega)$ is the difference between the two spectra $S(\omega)$ and $S'(\omega)$ on a log magnitude versus frequency scale and defined by

$$V(\omega) = \log S(\omega) - \log S'(\omega). \quad (25)$$

3 Dynamic-Time Warping Analysis of PPG-based Features

Dynamic-time warping (DTW) is often used as a method for pattern comparison in time series [23, 29, 30]. Based on the principle of dynamic programming, DTW attempts to optimally align time series by stretching or compressing the reference and test patterns so that the accumulative distance of the DTW path is minimized. More specifically, let $\mathcal{F}$ be a feature space, and also let $R = (r_1, r_2, \dots, r_n, \dots, r_N)$, and $T = (t_1, t_2, \dots, t_m, \dots, t_M)$ be the real PPG-based feature vectors of reference and test patterns, respectively. A warping path is a sequence of grid points: $W = (w_1, w_2, \dots, w_k, \dots, w_K)$, where $w_k = (n_k, m_k) \in [1, N] \times [1, M]$, a grid point along axes R and T . A legal warping path W needs to satisfy the following three criteria:

1. Boundary condition: $w_1 = (1, 1)$, and $w_K = (N, M)$ (endpoint constraints).
2. Monotonicity condition: $n_1 \leq n_2 \leq \dots \leq n_K$, and $m_1 \leq m_2 \leq \dots \leq m_K$ (monotonic order with respect to time).
3. Continuity condition: $n_{k+1} - n_k \leq 1$ and $m_{k+1} - m_k \leq 1$, $k \in [1, K - 1]$ (step-size constraints).

To compare the similarity between two feature vectors R and T , it is necessary to define a local cost function or local distance measure $\delta(r_n, t_m)$, such that $\delta : \mathcal{F} \times \mathcal{F} \rightarrow \mathcal{R} \geq 0$. The smaller $\delta(r_n, t_m)$ is, the more similar r_n and t_m are to each other. DTW searches for a similarity between the reference and test feature vectors to identify the optimal warping path, denoted as $\mathbf{W}^*$, that has the minimal total distance (cost) over all possible warping paths:

$$W^* = \arg \min_W [DTW(R, T)] = \arg \min_W [\delta_W(R, T)] \quad (26)$$

In dynamic programming, an optimization that can be used for solving certain problems requiring sequential decisions that must be made at various stages. The formulation of a dynamic programming problem needs a *stage* variable, *state* variable, and *decision* variables [31]. The basic procedure of a dynamic programming approach for solving an optimization problem includes the following steps [31]:

1. Presenting the problem as a system that consists of a number of stages with associated states, and a decision that needs to be made at each stage.
2. The decision made at one stage affects the state of the system at the next stage.
3. Starting at the last stage, a one-stage subproblem can be solved giving the optimal decisions for each state in the final stage.
4. Finally, using recursive relations that allow the solution of the one-stage subproblem to be used to find solutions to larger and larger subproblems with the final subproblem being the original problem.

To reduce the search space involving the numbers of stages and states, some restrictions are necessary to identify possible warping paths in an efficient way. These restrictions are outlined as follows.

1. Warping window: $|n_k - m_k| \leq \omega$, where ω is a positive integer that represents the window bandwidth (only grid points within a warping-path window are considered).
2. Slope constraint: Slope-weighting vector (w_H, w_V, w_D) , where w_H , w_V , and w_D are the weights for the horizontal, vertical, and diagonal directions, respectively (avoiding warping path that is either too steep or too shallow, and preventing very short segments from being matched with very long ones).

Based on the dynamic programming formulation, the cumulative distance for each point of a warping path, denoted as $\gamma(n_k, m_k)$, is computed using the following backward recursive relation:

$$\gamma(n_k, m_k) = \delta(n_k, m_k) + \min[\gamma(n_k - 1, m_k - 1), \gamma(n_k - 1, m_k), \gamma(n_k, m_k - 1)]. \quad (27)$$

Similarly, the optimal warping path W^* can be found by tracing backward the index order, selecting the point with the lowest accumulative distance. Starting with the endpoint $w_K^* = (N, M)$ of which accumulative distance is minimum, the back-tracking algorithm for finding the optimal warping path is described as follows.

$$w_{k-1}^* = \begin{cases} (1, m_k - 1), & \text{if } n_k = 1 \\ (n_k - 1, 1), & \text{if } m_k = 1 \\ \arg \min[DTW(n_k - 1, m_k - 1), \\ DTW(n_k - 1, m_k), \\ DTW(n_k, m_k - 1)], & \text{otherwise.} \end{cases} \quad (28)$$

4 Experiments

The PPG data of 18 elderly participants were used in this study, and originally described in [19]. These participants were clinically diagnosed with dementia. The qualified middle-age care-giver was familiar with the daily living activities of the participants.

The participants’ pulses of the index finger of the left hand were measured with a PPG sensor connected to a personal computer for 3 minutes before the therapeutic session, 3 minutes during the one-on-one session with the care-giver whose left-finger pulse waves were also recorded at the same time as each of the participants, and 3 minutes after the session. The PPG measurements of the participants before and after the session were designed to be used as the control signals to study the influence of the care-giver over the participants during the cognitive stimulation therapy. The therapeutic conversation mainly includes questions about the participants’ feelings, physical condition, what they had done before the measurement, hobbies, family, friends, and memories of the past. The cognitive stimulation therapy between the care-giver and elderly participants involved conversation only (without touching) by engaging the participants in talking about topics of relevance to them.

The PPG signals were detrended and smoothed by using the Savitzky-Golay filter [32]. The preprocessed PPG data of the care-giver and participants with dementia, which were synchronously recorded, were used to calculate the corresponding LLE-based and LPC-based features before, during and after the therapeutic session. Figure 1 shows segments of typical pre-processed PPG signals of the care-giver and a participant. The dissimilarity matrices of the PPG data between the care-giver and the participants obtained from the LLE-based feature using the Euclidean distance, and LPC-based using the distortion measures were then used to construct the “phylogenetic” trees with the UPGMA algorithm [33]. Figures 2 and 3 show two typical trees of the synchronized PPG data of two participants and the care-giver, in which the terms *Care-giver*, *Before care*, *During care*, and *After care* in the tree nodes denote the care-giver, the participated individual before, during, and after the therapeutic session, respectively. The topologies of the trees shown in Figures 2 and 3 suggest effective results of the CST because the care-giver and the participant during care are in the same node. Figures 4 and 4 show two typical trees of the synchronized PPG data of other two participants and the care-giver, suggesting ineffective CST sessions because the care-giver and the participant during care are not grouped in the same node.

The DTW was used to evaluate the reliability of the trees obtained from the ID and LLE based measures of similarity. The global constraint region (warping window) adopted in this study is the Itakura parallelogram [27]. For the (local) sloping constraint, the equally weighted conditions $(w_H, w_V, w_D) = (1, 1, 1)$ are used, which reduces to classical DTW in which alignment in the diagonal direction (cost of one) is preferred to the horizontal and vertical directions (cost of two). The end points of the feature vectors were considered given a priori, and the temporal variations of the path were located within the range defined by

Fig. 1. First 10000 samples of synchronized PPG signals of (a) elderly participant with dementia, and (b) care-giver.

Fig. 2. Synchronized cognitive stimulation communication, which is considered to be effective as the care-giver and participant “during care” are from the same node (sister groups), while “Before care” and “After care” are the outgroups to the care-giver and “during care”.

Fig. 3. Synchronized cognitive stimulation communication, which is considered to be effective as the care-giver and participant “during care” are from the same node (sister groups), while “Before care” and “After care” are the sister groups and outgroups to the care-giver and “during care”.

Fig. 4. Synchronized cognitive stimulation communication, which is considered to be not effective as the care-giver and participant “Before care” are from the same node (sister groups), while the participant “During care” is an outgroup to the care-giver.

Fig. 5. Synchronized cognitive stimulation communication, which is considered to be not effective as the care-giver and participant “After care” are from the same node (sister groups), while the participant “During care” is an outgroup to the care-giver.

the end points. The Euclidean distance was used to calculate the cost for the transitions where no cost is imposed on the transitions to a specific node. This Euclidean-based cost function fully depends on the feature vectors corresponding to the respective node. Detailed descriptions of these constraints can be found in [23, 34].

To assess the robustness of the LLE and spectral features, the vectors of LPC coefficients (used for calculating ID and LSD), and vectors of LLEs were used for DTW-based template matching, respectively. The lengths of the LPC coefficients are 16, 20, 24, and 28; and the length of the LLE vectors are 103. The idea is to compare a randomly generated vector that is of the same length of the feature vector of the care-giver with the feature vectors of the care-giver and during care. Using the DTW, the calculated similarity of the feature vectors of the care-giver and during care must be smallest among other pair-wise similarity measures associated with the random vector. To establish statistics for the uncertainty of the quantity of the reliability evaluation of the feature vectors, the bootstrap resampling [35] was employed. Bootstrap methods are often used as a general tool for assessing statistical accuracy of the performance of a learning model, and can provide a measure of the quality of the selected model [36]. The bootstrap procedure involves choosing random feature vectors with replacement from a sample feature vector and analyzing each sample the same way. Sampling with replacement means that each observation is selected separately at random from the original feature space. So a particular data point from the original feature space could appear multiple times in a given bootstrap vector. The number of

elements in each bootstrap vector equals the number of elements in the original feature vector of the care-giver.

Table 1. Reliability of LLE and LPC features for computerized assessment of cognitive stimulation therapy.

Method	Average Reliability (%)
LLE-based distance	18
Itakura distortion (LPC-based)	100
Log spectral distortion (LPC-based)	100

The bootstrap was used to resample the LPC and LLE vectors 100 times. Through the resampling analysis, the average reliability of using the LPC for computing the ID, LSD, and LLE for calculating the Euclidean distance for the construction of the phylogenetic trees was found to be 100%, 100%, and 18%, respectively. The results suggest that the use of the spectral distortion for CST assessment is obviously robust and preferred to the LLE-based Euclidean distance. Some reasons for the robust performance of the LPC-based measures of signal similarity are that, firstly the LPC coefficients are derived by taking an advantage of a high correlation between adjacent samples, such as the PPG signals studied herein; secondly LPC coefficients are mathematically presented in sequential order, which is sensitive to random vectors. The LLE values are computed as scalars. These two categories can be distinguished from one another by their distinct definitions: scalars are quantities that are described by a magnitude (or numerical value) alone, while vectors are quantities that are expressed by both a magnitude and a direction.

Another important finding of the assessment of computer models for cognitive stimulation therapy is that as a single computational method for feature extraction of PPG signals rarely exists, the establishment of the reliability of different learning models can equivalently constitute to the construction of the degrees of importance of the models. These are essential parameters for model or information fusion for solving problems in many applications, by which results can be improved [19, 38, 39].

5 Conclusion

The application of dynamic-time warping and bootstrapping for assessing the performance of the LLE and LPC features for the evaluation of cognitive stimulation therapy for people with dementia have been presented and discussed. The performance the LPC feature is more robust than the LLE, and shown to be consistent with the use of two different distortion measures. This pilot research is useful for establishing the reliability of computerized methods in health informatics, and assisting researchers in cognitive stimulation therapy to quickly

validate hypotheses that would benefit the quality of life of people with cognitive disorders [40].

Acknowledgments: PPG and LLE data were provided by Mayumi Higa of Chaos Technology Research Lab, Shiga, Japan. Satoshi Haga assisted the author in carrying out the computer experiments.

References

1. Grohol, J.M. (2004). What to look for in quality online support groups. http://psychcentral.com/archives/support_groups.htm. Accessed 01 May 2015.
2. Dinesen, B., Toft, E. (2009). Telehomecare challenge collaboration among health-care professionals. *Wireless Personal Communications*, 51, 711-724.
3. Delmastro, F. (2012). Pervasive communications in healthcare. *Computer Communications*, 35, 1284-1295.
4. Algaet, M.A., Noh, Z.A., Shibghatullah, A.S., Milad, A.A., Mustapha, A. (2014). Provisioning quality of service of wireless telemedicine for e-health services: a review. *Wireless Personal Communications*, 78, 375-406.
5. Clark, L.W. (1995). Interventions for persons with Alzheimer's disease: Strategies for maintaining and enhancing communicative success. *Topics in Language Disorders*, 15, 47-65.
6. Jootun, D., McGhee, G. (2011). Effective communication with people who have dementia. *Nursing Standard*, 25, 40-46.
7. Orgeta, V., Qazi, A., Spector, A.E., Orrell, M. (2014). Psychological treatments for depression and anxiety in dementia and mild cognitive impairment. *Cochrane Database of Systematic Reviews*, 1, CD009125. DOI: 10.1002/14651858.CD009125.pub2.
8. Kueider, A.M., Parisi, J.M., Gross, A.L., Rebok, G.W. (2012). Computerized cognitive training with older adults: a systematic review. *PLoS ONE*, 7, e40588. DOI:10.1371/journal.pone.0040588.
9. Shelley, K., Shelley, S. (2001). Pulse oximeter waveform: photoelectric plethysmography. In *Clinical Monitoring*, 420-428.
10. Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28, R1-R39.
11. Asada, H.H., Shaltis, P., Reisner, A., Rhee, S., Hutchinson, R.C. (2003). Mobile monitoring with wearable photoplethysmographic biosensors. *IEEE Engineering in Medicine and Biology Magazine*, 22, 28-40.
12. Reisner, A.T., Shaltis, P.A., McCombie, D., Asada, H.H. (2008). Utility of the photoplethysmogram in circulatory monitoring. *Anesthesiology*, 108, 950-958.
13. Frey, B., Waldvogel, K., Balmer, C. (2008). Clinical applications of photoplethysmography in paediatric intensive care. *Intensive Care Medicine*, 34, 578-582.
14. Monte-Moreno, E. (2011). Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques. *Artificial Intelligence in Medicine*, 53, 127-138.
15. Pham, T.D., Truong, C.T., Oyama-Higa, M., Sugiyama, M. (2013). Mental-disorder detection using chaos and nonlinear dynamical analysis of photoplethysmographic signals. *Chaos, Solitons & Fractals*, 51, 64-74.

16. Sonoda, K., Kishida, Y., Tanaka, T., Kanda, K., Fujita, T., Higuchi, K., Maenaka, K. (2013). Wearable photoplethysmographic sensor system with PSoC microcontroller. *Int J Intelligent Computing in Medical Sciences & Image Processing*, 5, 45-55.
17. Yousefi, R., Nourani, M., Ostadabbas, S., Panahi, I. (2014). A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors. *IEEE J Biomedical and Health Informatics*, 18, 670-681.
18. Wang, W., Stuijk, S., de Haan, G. (2015). Exploiting spatial-redundancy of image sensor for motion robust rPPG. *IEEE Trans Biomedical Engineering*, 62, 415-425.
19. Pham, T.D., Oyama-Higa, M., Truong, C.T., Okamoto, K., Futaba, T., Kanemoto, S., Sugiyama, M., Lampe, L. (2015). Computerized assessment of communication for cognitive stimulation for people with cognitive decline using spectral-distortion measures and phylogenetic inference. *PLoS ONE*, 10, e0118739. DOI:10.1371/journal.pone.0118739.
20. Spigulis, J., Erts, R., Nikiforovs, V., Kviesis-Kipge, E. (2008). Wearable wireless photoplethysmography sensors. In *Proc. SPIE Biophotonics: Photonic Solutions for Better Health Care*, 6691, 7 pages. DOI: 10.1117/12.801966.
21. Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, 898, 366-381.
22. Rosenstein, M.T., Collins, J.J., DeLuca, C.J. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65, 117-134.
23. Rabiner, L., Juang, B.H. (1993). *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall.
24. Ingle, V.K., Proakis, J.G. (1997). *Digital Signal Processing Using Matlab V.4*. Boston: PWS Publishing.
25. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge: Cambridge University Press.
26. Itakura F., Saito S. (1968). An analysis-synthesis telephony based on maximum likelihood method. In *Reports of 6th International Congress on Acoustics*, C:5-5, C:17-20.
27. Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoust Speech Signal Process*, 23, 67-72.
28. Soong, F., Sondhi, M.M. (1987). A frequency-weighted Itakura spectral distortion measure and its application to speech recognition in noise. In *Proc IEEE Int Conf Acoustics, Speech, and Signal Processing*, 625-628.
29. Berndt, D.J., Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proc. AAAI-94 Workshop on Knowledge Discovery in Databases*, 359-370.
30. Muller, M. (2007). Dynamic time warping. In *Information Retrieval for Music and Motion*, Berlin: Springer, 69-84.
31. Ecker, J.G., Kupferschmid, M. (1988). *Introduction to Operations Research*. New York: John Wiley & Sons.
32. Orfanidis, S.J. (1996). *Introduction to Signal Processing*. Englewood Cliffs: Prentice-Hall.
33. Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum.
34. Theodoridis, S., Koutroumbas, K. (2009). *Pattern Recognition*. London: Academic Press.

35. Efron, B., Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Boca Raton: Chapman & Hall/CRC.
36. Hastie, T., Tibshirani, R., Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
37. Pham, T.D. (2011). Fuzzy posterior-probabilistic fusion. *Pattern Recognition*, 44, 1023-1030.
38. Mirhosseini, A.R., Yan, H., Lam, K.M., Pham, T. (1998). Human face image recognition: An evidence aggregation approach. *Computer Vision & Image Understanding*, 71, 213-230.
39. Pham, T.D., Yan, H. (1997). "Fusion of handwritten numeral classifiers based on fuzzy and genetic algorithms. In *Proc. North America Fuzzy Information Processing Society*, 257-262.
40. Spector, A., Thorgrimsen, L., Woods, B., Royan, L., Davies, S., Butterworth, M. Orrell M. (2003). Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial. *Br J Psychiatry*, 183, 248-254.