Identification of Human Variability

Miguel Perez-Xochicale

Abstract—

Index Terms—Activity Recognition; On-Body Inertial Sensors

1 INTRODUCTION

HUMAN Activity Recognition (HAR) using body-worn sensors has increased during the last 20 years. This is due to three factors: (i) technology advances in sensors, (ii) longer battery lifetimes and (iii) different application-oriented scenarios. In contrast to speech recognition and computer vision frameworks, HAR offers different challenges based on the complexity and diversity of human activities (e.g. ambulation, transportation, phone usage, daily activities, exercise, military, upper body), the selection of different sensors to use (e.g. inertial, light, temperature or audio sensors) [1] and different bodily locations of sensors (e.g. chest, wrist, lower back, hip, thigh, foot) [2], [3].

According to Bulling *et al.* [4] the common challenges in HAR using body-worn sensors are: intraclass variability, interclass similarity, and the NULL class problem. For this PhD, identifying the variability of human activities is more challenging than identifying the action itself. Intraclass variability, therefore, occurs when an activity is performed differently either by a single person or several people. For instance, variability is presented in either dance features (e.g. fluency of motion, coordination, steadiness of the rhythm, adding erratic or additional movements [5], [6]) or biological features of dancers (e.g. gender, age, home country [5], [7]).

Hammerla *et al.* [8] have examined the effects of variability using artificial signals so as to create motion structures (strategy of the motion activity) and motion noise (the precision of the motion) of human activities. To quantify the variability of motion activities, Hammerla *et al.* [8] proposed the use of PCA to compute the area behind the cumulative energy curve which is used as a metric for motor skill assessment. The variability in human activities has therefore a relation with qualitative assessment of motion structures and motion noise of human activities. Velloso *et al.* [9], for example, assessed automatically the quality of weight-lifting activity. Similarly, Velloso *et al.* [10] quantify how *good* the repetition of weight-lifting activity is in terms of angles of each bone in relation to references planes.

Recently, concepts from non-linear analysis tools such as fractal dimensionality, the Lyapunov exponent or time-delay embedding has been applied to understand human activities. For instance, Yamamoto *et al.* [11], [12] used the fractal dimensionality of the attractors to model repeated forehand and backhand tennis strokes. Gouwanda *et al.* [13] showed that the variability in walking speed has a linear relationship with the Lyapunov exponent. This exponent is therefore suitable for analysing the temporal variation in gait stability. Time-delay embedding has been used as a feature for gait recognition [14] as well as the recognition of walking, lingering, running, up stairs and downstairs activities [15]. Additionally, Caballero *et al.* [16] reviewed another non-linear analysis tools (e.g. local dynamic stability, recurrent quantification analysis, entropy measurements, detrended fluctuation analysis) to measure the human movement variability. However, the questions to ask, as pointed out by Caballero *et al.* [16], are"...do this tools actually measure variability? and, what kind of variability?".

Given the case of study of the variability in dance activities, it is hypothesised that there are three possible reasons for this: (i) inherent noise in sensors, (ii) inherent properties of the activity itself and (iii) discrepancies of biological features of people. For this PhD, the following research questions will be addressed:

1) How can the time-delay embedding and PCA methods quantify the possible reasons of the variability of dance activities ?
2) Having known the limitations of the time-delay embedding and PCA methods, which other non-linear analysis tools would be suitable to explore the variability in dance activities and use them as a features for machine learning algorithms?

2 THE ACTIVITY RECOGNITION CHAIN

Bulling *et al.* [4] reviewed the state of the art of HAR using body-worn inertial sensors. Figure 1 illustrates the typical activity recognition chain (ARC) to identify activities with body-worn sensors.

The first stage of the ARC is the raw data collection from several sensors attached to different parts of the body. Sensors, s_i, provide multiple values, d^i, (e.g. 3-D acceleration referred to as x, y and z direction)

$$s_i = (\mathbf{d}^1, \mathbf{d}^2, \dots \mathbf{d}^t), \text{for } i = 1, \dots, k \tag{1}$$

where k denotes the number of sensors.

In the preprocessing of the ARC, stage raw multivariable times series are transformed into a pre-processed time series $D' = (d'_1, \dots, d'_n)^T$, where d'_i is one dimension data of the preprocessed time series, n is the number of total data dimension. The preprocessing tasks may involve synchronisation, calibration, unit conversion, normalisation, resampling, denoising or baseline drift removal of the raw data.

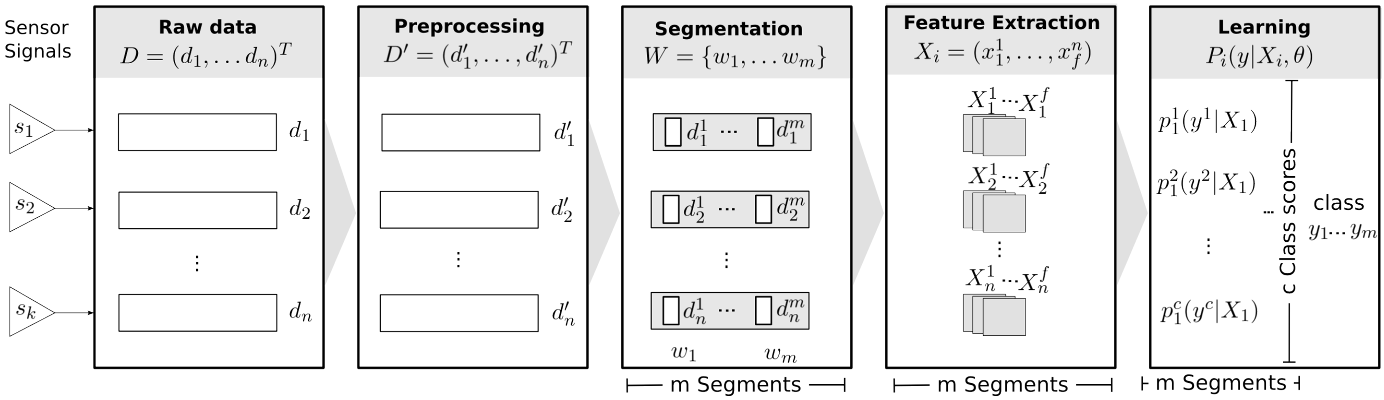

Fig. 1. Typical activity recognition chain (ARC) to identify activities or gestures from body-worn inertial sensors. Diagram is replicated from the work of Bulling *et al.* [4].

The stage of the data segmentation identifies segments that are likely to have information about activities. The segmentation stage creates a set of segments w_m containing a possible activity y

$$W = \{w_1, \ldots, w_m\} \tag{2}$$

where m correspond to the number of segments. Since the segmentation of the data is a difficult problem, there are various methods in the literature to tackle this problem: sliding window, energy-based segmentation, rest-position segmentation, additional sensors and external context sources.

In the feature extraction stage, a feature extraction function F reduces the signals D' into segmented signals W. The total number of features X_i is the future space.

$$X_i = F(D', w_i) \tag{3}$$

In the literature on activity recognition different methods for feature extraction can be found including signal-based features, body model features, event-based features, multilevel features or automatic feature ranking and selection.

Machine learning tools have been used in HAR over the last 15 years so as to describe, analyse and predict human activities [4]. However, the chosen approach is subject to computational complexity, recognition performance or latency. Generally for the learning stage, a training data $T = \{X_i, y_i\}_{i=1}^N$ is computed prior to the classification with N pairs of feature vectors X_i and ground truth labels y_i. For this stage, model parameters θ can be learned to decrease the classification error on T. Then, with the trained model T, each feature vector X_i is mapped to a set of class labels $Y = \{y^1, \ldots, y^c\}$ with scores $P_i = \{p_i^1, \ldots, p_i^c\}$:

$$p_i(y \mid X_i, \theta) = I(X_i, \theta), \text{for } y \in Y \tag{4}$$

and inference method I. Finally, the classification output y_i is computed with the maximum score P_i

$$y_i = \underset{y \in Y, p \in P_i}{\operatorname{argmax}} \, p(y|X_i, \theta) \tag{5}$$

The most common classification algorithms are: decision trees, Bayesian, instance base, domain transform, fuzzy logic, Markov models support vector machines, artificial neural networks and ensembles of classifiers [1].

Similarly, when the recognition of activities can miss, confuse or falsely recognise activities that did occur, several metrics can be used to optimise the classification. Some of the metrics are confusion matrices, accuracy, precision, recall, and F-scores, decision-independent Precision-Recall or receiver operating characteristic curves [4].

3 Recognising Dexterity In Dance

As Miura et al. [17] point out "... how the human motor system produces dance movements is still poorly understood." A key issue concerns the manner in which experienced dancers solve the 'degrees of freedom' problem in the face of changing contextual demands. Miura et al. [18] measured muscle activation using electromyographic (EMG) data collected from muscles in the lower limb, for a task requiring participants to bounce up and down in time to a metronome beat. They demonstrated that experienced dancers show much better precision in synchronizing movements to beat than non-dancers, i.e., dancers maintained much lower standard deviation in temporal deviation against the beat than non-dancers. This result is consistent with work which shows that, compared with inexperienced- or non-dancers, trained ballet dancers exhibit superior postural stability [19], and show superior ability in position matching of upper limbs [20].

Capturing dance activity through sensors has tended to rely on motion capture [21] or sensors mounted on the person [22] or in their shoes [23] or from their smartphones [24]. Much of this work has been concerned with using the dancers motion to work with multimedia presentations that augment and complement the dance [25], [26] or as interfacing to a game [27] or commercial games, such as Dance Dance Revolution. While the range of sensing technology used in these papers is diverse and the result of the activity recognition is varied, it is fair to say that few of the papers have considered variability or dexterity in how a dance is performed. In their work, Aristidou et al. [6] have considered the manner in which dance steps conform to a set of defined templates that describe steps in terms of three-dimensional rotation (described using quaternions). The implication is that a goodness-of-fit can be ascertained to determine how well a dancer performs a step, and how any deviation from good can be modified through practice.

For this report, we are interested in the question of how time-delay embedding techniques can provide insight into the variability and dexterity of dancers. To this end, we consider the performance of a set of steps from Salsa dance and compare untrained, inexperienced or non-dancers in one cohort with experienced dancers in another. Before explaining how the data is collected, the next section outlines the approach to time-series time-delay embedding and the resulting phase space representation used in this report. It should be noted that dynamical systems research offers a range of techniques for the study of human activity (see [28] for an overview of alternative techniques).

4 Time-delay Embedding

The aim of time-delay embedding, also known as Takens's Theorem, is to reconstruct a $k-$dimensional manifold M of an unknown dynamical system $s(t)$ from a time series $x(t)$. Time-delay embedding assumes that the time series is a sequence $x(t) = h[s(t)]$, where $h : M \rightarrow \mathbb{R}$ is a measurement function in the unknown dynamical system, being $x(t)$ measurable.

Thus, the time delay reconstruction is defined as: $\overline{x}(t) = (x(t), x(t-\tau), ..., x(t-(m-1)\tau))$ where m is the embedding dimension and τ is the embedding time-delay. $\overline{x}(t)$ defines a map $\Phi : M \rightarrow \mathbb{R}^m$ such that $\overline{x}(t) = \Phi(s)$. Similarly, $y(t) = \Psi[\overline{x}(t)]$ is a n-dimensional vector where $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a further transformation (e.g., PCA [29], Nonlinear PCA [30], Locally Linear Embedding [31]). Figure 2 illustrate the time delay reconstruction process. For details, see the work of Uzal *et al.* [32].

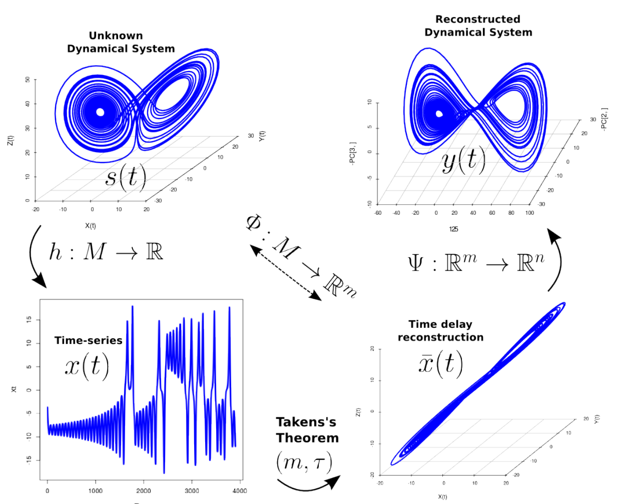

Fig. 2. The reconstruction problem. The figure is based on the work of Uzal *et al.* [32].

4.1 Embedding Parameters m and τ

Given any time series $x(t)$, the time delay reconstruction system, $\overline{x}(t)$, is easy to implement. For this work, Cao's method [33], a modification of the False Nearest Neighbours (FNN) algorithm, and mutual information algorithm by Fraser *et al.* [34] have been used to calculate minimum embedding parameters (m_{min}, τ_{min}).

4.1.1 Minimum Embedding Dimension m_{min}

Cao's method [33] for computing the minimal embedding dimension is based on the mean values $E1(d)$ and $E2(d)$ in which d is the range of evaluation of the embedding dimension.

$E1(d)$ is used to obtain the minimal dimension m_{min} and stops changing when the time series comes from an attractor (Figure 3 B). We computed $E1(d)$ values for $1 \leq \tau \leq 10$ to exemplify the dependency of τ given periodic, chaotic and random time series (Figures 3 (A,B,C)).

The second of these values, $E2(d)$, is used to distinguish deterministic signals from random signals in which case the $E2(d)$ values will be approximately equal to 1 for any d (Figure 3 F). Similarly, we computed $E2(d)$ values for periodic, chaotic and random time series, to exemplify the dependency of $1 \leq \tau \leq 10$ (Figures 3 (D,E,F)).

Cao's method is a modified version of the FNN method, and $E1(d)$ and $E2(d)$ values are only dependant on m and τ [33].

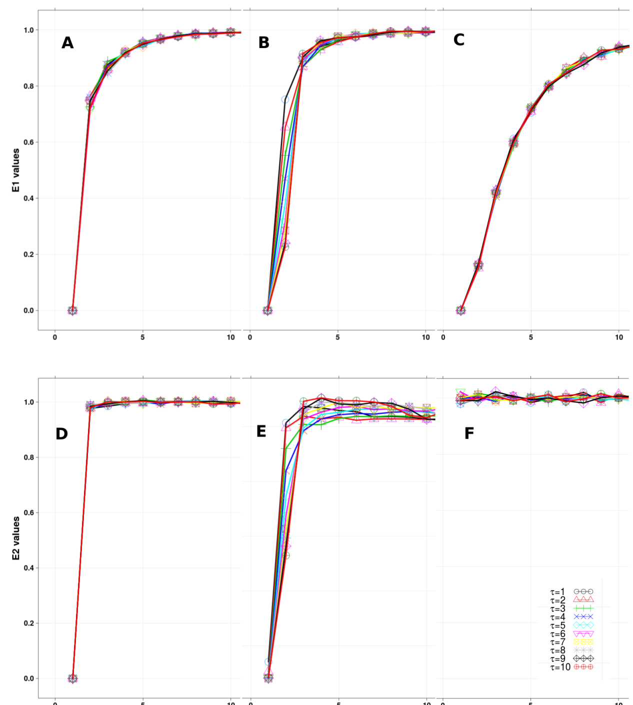

Fig. 3. The values of $E1(d)$ and $E2(d)$ with different time delay embedding parameters from periodic (A,D), chaotic (B,E) and random (C,F) time series.

4.1.2 Minimum Time-delay Embedding τ_{min}

The method of choosing the minimum Time-delay embedding τ_{min} was proposed by Fraser *et al.* [34] in which the first minimum of the mutual information graph is chosen to estimate the minimal time-delay embedding. For instance, Figure 4 illustrates the mutual information from periodic, chaotic and random time series. The local minimum for the Chaotic series in Figure 4 is $\tau_{min} = 18$. On the other hand, for the periodic and random time series the mutual information plots have no local minimum and values are monotonically decreasing which means that $\tau_{min} = 1$ (Figure 4) [34].

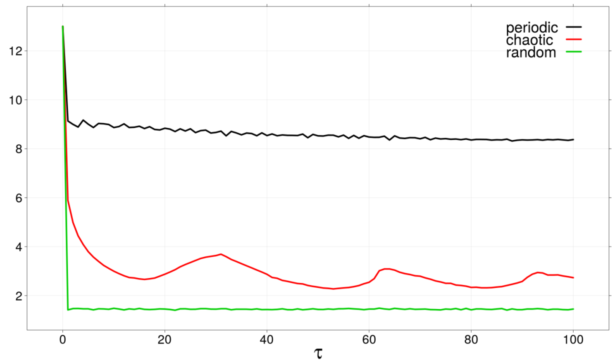

Fig. 4. Mutual information plots from periodic, chaotic and random time series.

4.1.3 Embedding Parameters Setbacks

Although the time-delay embedding method using inertial sensors has been used extensively in gait recognition [14], gait stability [13] and walking, running and cycling activities [15], some problems with the minimal embedding parameter estimation (m_{min} and τ_{min}) still remain to be solved.

Sama *et al.* [14] and Gouwanda *et al.* [13] estimated the minimal embedded dimension (m_{min}) with the False Nearest Neighbours (FNN) method. However, Cao [33] pointed out that the FNN algorithm introduces new parameters (R_{tol} and A_{tol}) that lead to different results and cannot differentiate random series from deterministic series. Frank *et al.* [15] proposed a grid search method to find the minimal embedded parameters, but there are no details about their approach.

In the case of the minimal time delay embedding value, τ_{min}, Fojt *et al.* [35] mentioned a method in which the chosen τ is made in function of filling optimally the reconstructed state space; however, Fojt *et al.* [35] mentioned that "it is a rough estimation based on a graphical procedure." Although, Sama *et al.* [14] computed τ_{min} using the method proposed by Fraser *et al.* [34], they pointed that the chosen τ_{min} largely depend on the application.

5 Methods

5.1 Artificial Signals

Following the proposal of Hammerla *et al.* [8], artificial signals are created to examine the effects of variability in the precision of motion (additive noise) and in the strategy of motion (structural noise) of activities.

Additive noise is just normalised noise with variance σ_a^2 added to the sinusoid signal S:

$$S^a = S + \mathbf{N}(0, \sigma_a^2) \tag{6}$$

Structural noise is a sinusoid signal distorted with different variance in frequency and amplitude σ_s^2 and window length w_s. Algorithm 1 describes the creation of structural noise. The data is whitened (i.e. data is normalised to have zero mean and unit variance) to make the data less redundant.

Algorithm 1 Structural Noise

Input: time-series S^a, variance σ_s^2, window length w_s
Output: Structurally distorted signal S^s

```
for j = 1 to L, j = j + w_s do
    u' ← N(0, σ_s^2)
    S^a = sinusoid with frequency |u'| and variance σ_a^2 of length w_s
    S^s_{j→j+w_s} = S^s_{j→j+w_s} + S^a × σ_s^2
end for
S^s = whiten(S^s)
return S^s
```

By varying both σ_a^2 and σ_s^2 is possible to simulate and control the additive noise and the structural noise in the structure of the human activity. For example, low values of σ_a^2 are associated with precise movements while low values of σ_s^2 correspond to a well chosen strategy for a motion.

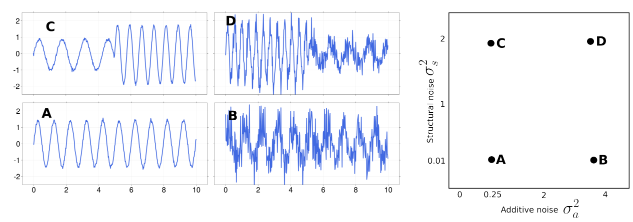

Fig. 5. Graphs (A, B, C and D) present the variability of additive noise (σ_a^2) and structural noise (σ_s^2) on the sinusoid signals with $w_s = 500$ according to the parameters indicated on the left plot.

For graphs in Figure 5, the base frequency of the sinusoid is 1 Hz sampled at 50Hz with an amplitude 1 and window length per of 250.

5.2 TAKEN's-PCA METHOD

6 Publication Plan

Publish my research's results in Measuring Behaviour 2016 and Augmented Human 2016 conferences so as to push myself to submit a journal in the 15th Month of the PhD.

7 Future Work

Acknowledgments

Miguel Pérez-Xochicale gratefully acknowledges the studentship from the National Council for Science and Technology (CONACyT) Mexico to pursue his postgraduate studies at University of Birmingham.

References

[1] O. D. Lara and M. a. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[2] I. Cleland, B. Kikhia, C. Nugent, A. Boytsov, J. Hallberg, K. r. Synnes, S. McClean, and D. Finlay, "Optimal placement of accelerometers for the detection of everyday activities." *Sensors*, vol. 13, no. 7, pp. 9183–9200, 2013.

[3] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, "Activity Recognition Using A Single Accelerometer placed at the Wrist or Ankle," *Med Sci Sports Exerc.*, vol. 45, no. 11, pp. 2193–2203, 2013.

[4] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn intertial sensors," *ACM Computing Surveys*, pp. 1–33, 2014.

[5] K. Grammer, E. Oberzaucher, I. Holzleitner, and S. Atmaca, "Dance: the human body as a dynamic motion system," *The Implications of Embodiment: Cognition and Communication*, pp. 173–192, 2011.

[6] A. Aristidou, E. Stavrakis, and Y. Chrysanthou, "Motion Analysis for Folk Dance Evaluation," *EUROGRAPHICS Workshop on Graphics and Culture Heritage*, p. 20141304, 2014.

[7] D. Iwai, T. Felipe, N. Nagata, and S. Inokuchi, "Identification of motion features affecting perceived rhythmic sense of virtual characters through comparison of latin american and japanese dances," *The Journal of The Institute of Image Information and Television Engineers*, vol. 65, pp. 203–210, 2011.

[8] N. Hammerla, T. Ploetz, P. Andras, and P. Olivier, "Assessing motor performance with pca," *International Workshop on Frontiers in Activity Recognition using Pervasive Sensing*, pp. 18–23, 2011.

[9] E. Velloso, a. Bulling, H. Gellersen, W. Ugulino, and H. Fuks, "Qualitative activity recognition of weight lifting exercises," *Proceedings of the 4th Augmented Human International Conference*, pp. 116–123, 2013.

[10] E. Velloso, A. Bulling, and H. Gellersen, "MotionMA: Motion Modelling and Analysis by Demonstration," *CHI 2013*, 2013.

[11] H. Suzuki and Y. Yamamoto, "Dexterity of Switched Hitting Movement Using Fractal Dimension Analysis," *Japanese Journal of Sport Psychology*, vol. 40, pp. 91–108, 2013.

[12] Y. Yamamoto and K. Gohara, "Continuous hitting movements modeled from the perspective of dynamical systems with temporal input," *Human Movement Science*, vol. 19, no. 3, pp. 341–371, Aug. 2000.

[13] D. Gouwanda and A. Senanayake, "Non-linear Time Analysis to Estimate Gait Stability Using Wearable Gyroscopes Network," *Journal of Robotics and Mechatronics*, vol. 24, no. 4, pp. 1–22, 2012.

[14] A. Samà, F. J. Ruiz, N. Agell, C. Pérez-López, A. Català, and J. Cabestany, "Gait identification by means of box approximation geometry of reconstructed attractors in latent space," *Neurocomputing*, vol. 121, pp. 79–88, 2013.

[15] J. Frank, S. Mannor, and D. Precup, "Activity and Gait Recognition with Time-Delay Embeddings Time-Delay Embeddings," *AAAI Conference on Artificial Intelligence*, pp. 407–408, 2010.

[16] C. Caballero, D. Barbado, and F. J. Moreno, "Non-Linear Tools and Methodological Concerns Measuring Human Movement Variability: an Overview," *European Journal of Human Movement*, vol. 32, pp. 61–81, 2014.

[17] A. Miura, S. Fujii, Y. Yamamoto, and K. Kudo, "Motor Control of Rhythmic Dance from a Dynamical Systems Perspective," *Journal of Dance Medicine and Science*, vol. 19, pp. 11–21, 2015.

[18] A. Miura, K. Kudo, T. Ohtsuki, H. Kanehisa, and K. Nakazawa, "Relationship between muscle cocontraction and proficiency in whole-body sensorimotor synchronization: a comparison study of street dancers and nondancers." *Motor control*, vol. 17, no. 1, pp. 18–33, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22964842

[19] D. Crotts, B. Thompson, M. Nahom, S. Ryan, and R. a. Newton, "Balance abilities of professional dancers on select balance tests." *The Journal of orthopaedic and sports physical therapy*, vol. 23, no. 1, pp. 12–17, 1996.

[20] J. R. Ramsay and M. J. Riddoch, "Position-matching in the upper limb: professional ballet dancers perform with outstanding accuracy." *Clinical rehabilitation*, vol. 15, no. 3, pp. 324–330, 2001.

[21] D. S. Alexiadis and P. Daras, "Quaternionic signal processing techniques for automatic evaluation of dance performances from MoCap data," *IEEE Transactions on Multimedia*, vol. 16, no. 5, pp. 1391–1406, 2014.

[22] A. Lynch, B. Majeed, B. O'Flynn, J. Barton, F. Murphy, K. Delaney, and S. C. O'Mathuna, "A wireless inertial measurement system (WIMS) for an interactive dance environment," *Journal of Physics: Conference Series*, vol. 15, pp. 95–100, 2005.

[23] J. Paradiso and E. Hu, "Expressive footwear for computer-augmented dance performance," *Digest of Papers. First International Symposium on Wearable Computers*, no. October, pp. 20–21, 1997.

[24] Y. Wei, H. Yan, R. Bie, S. Wang, and L. Sun, "Performance monitoring and evaluation in dance teaching with mobile sensing technology," *Personal and Ubiquitous Computing*, vol. 18, no. 8, pp. 1929–1939, 2014. [Online]. Available: http://link.springer.com/10.1007/s00779-014-0799-7

[25] N. Griffith and M. Fernström, "LiteFoot: A floor space for recording dance and controlling media," *Proceedings of the 1998 International Computer Music Conference*, pp. 475–481, 1998. [Online]. Available: http://www.ul.ie/ pal/litefoot/Index.htm

[26] C. Park, P. H. Chou, and Y. Sun, "A wearable wireless sensor platform for interactive dance performances," *Proceedings - Fourth Annual IEEE International Conference on Pervasive Computing and Communications, PerCom 2006*, pp. 52–57, 2006.

[27] N. N. Y. Chu, C. M. Yang, and C. C. Wu, "Game interface using digital textile sensors, accelerometer and gyroscope," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 184–189, 2012.

[28] S. J. Guastello and R. Gregson, *Nonlinear Dynamical Systems Analyis for the Behavioral Sciences using Real Data.* CRC Press., 2011.

[29] J. Shlens, "A Tutorial on Principal Component Analysis," 2014.

[30] U. Kruger, J. Zhang, and L. Xie, "Developments and Applications of Nonlinear Principal Component Analysis - a Review," vol. 1, pp. 1–43, 2007. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-73750-6_1

[31] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding." *Science (New York, N.Y.)*, vol. 290, no. 2000, pp. 2323–2326, 2000.

[32] L. C. Uzal, G. L. Grinblat, and P. F. Verdes, "Optimal reconstruction of dynamical systems: A noise amplification approach," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 84, no. 1, 2011.

[33] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, pp. 43–50, 1997.

[34] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," pp. 1134–1140, 1986.

[35] O. Fojt and J. Holcik, "Applying Nonlinear Dynamics to ECG Signal Processing," *IEEE Engineering in Medicine and Biology*, 1998.