Assessing Motor Performance with PCA

Nils Y. Hammerla
Culture Lab
School of Computing Science
Newcastle University
n.y.hammerla@ncl.ac.uk

Thomas Plötz
Culture Lab
School of Computing Science
Newcastle University
t.ploetz@ncl.ac.uk

Peter Andras
School of Computing Science
Newcastle University
peter.andras@ncl.ac.uk

Patrick Olivier
Culture Lab
School of Computing Science
Newcastle University
p.l.olivier@ncl.ac.uk

ABSTRACT

Information about the motor performance, i.e. how *well* an activity is performed, is valuable information for a variety of novel applications in Activity Recognition (AR). Its assessment represents a significant challenge, as requirements depend on the specific application. We develop an approach to quantify one aspect that many domains share – the efficiency of motion – that has implications for signals from body-worn or pervasive sensors, as it influences the inherent complexity of the recorded multi-variate time-series. Based on the energy distribution in PCA we infer a single, normalised metric that is intimately linked to signal complexity and allows comparison of (subject-specific) time-series. We evaluate the approach on artificially distorted signals and apply it to a simple kitchen task to show its applicability to real-life data streams.

Keywords

motor skill assessment, motion efficiency, PCA, metric

1. INTRODUCTION

A common setting in Activity Recognition (AR) is that sensors, such as triaxial accelerometers, are worn on the body or embedded into objects of daily use. The recorded multi-variate sensor streams undergo analysis in order to infer the activities that were performed by the subject. Often simple yet effective methods, such as k-NN classification using statistical features, suffice to obtain impressive recognition accuracies. Therefore information about *what* subjects are doing is readily available, rendering activity segmentation a straight-forward followup task. However, so far relatively little work was invested into a further, detailed analysis of these segmented activities, although extracting their characteristics, i.e. *how* these activities were performed, would be beneficial to a variety of applications spanning many domains.

One crucial aspect is the *skill* with which an activity was performed. Especially motor skills are of interest for a variety of applications such as rehabilitation, pain therapy, sports and professional training in tool usage, e.g. for mechanics, among others. They consist of a sequence of specialised movements that particularly emphasise the outcome and it is well understood how these are acquired [9]. Information about the development of this motor performance, i.e. if there is an increase or decline over time, can be very beneficial, particularly in medicine. Here many degenerative conditions such as Parkinson's Disease and Dementia have a significant impact on motor abilities, where an assessment of a decline is a common diagnostic tool.

Different application domains can have diverse requirements towards a skill assessment system. For some it can suffice to evaluate statistical attributes of the sensor signal – others might include high-level aspects such as the sequence of basic activities in a problem-solving task. Although systems exist that assess skill for very narrow domains (e.g., surgery [16], golf [12], or tennis [2]), they do not yet generalise across applications. On the way towards a generic skill assessment framework, many different building blocks have to be developed that assess specific aspects of motor performance, such as smoothness, time constraints, fine-grained segmentation and efficiency. The composition of these building blocks will reflect the requirements of each specific application.

In this paper we develop a method that can be used to assess one of the properties of motor skill: the efficiency of motion. During motor learning this efficiency rapidly increases as unsuitable motions are discarded [9]. The remaining motions will become more and more similar with increasing motor skill, which can be observed as an increased self-similarity on the signal level. The proposed metric infers the intrinsic *complexity* of given sensor data by quantifying the cumulative energy distribution in Principal Component Analsysis (PCA) when applied to frames extracted from the multi-variate time-series. We evaluate the resulting metric on artificial data and apply it to recordings from a sensor-equipped utensil in a simple cooking task to demonstrate how it is linked to motion efficiency. Results indicate that the general approach can be applied to other methods beyond PCA, rendering it a flexible approach to generic data analysis.

2. MOTOR SKILLS

The acquisition of new motor abilities, both of basic skills such as walking in childhood and sophisticated skills such as playing an instrument, is a well established field that stretches from psychology to neuroscience. It is common to divide the learning process into three phases [9]: In the *cognitive phase* the main motivation is to discover what actions have to be performed to solve a given task. Here different strategies are explored and the most efficient ones, i.e. those that reach a (sub)goal with the least amount of effort, are retained while inefficient ones are discarded. In this phase a very rapid improvement can be observed. In the subsequent *associative phase* these available, basic strategies are refined by applying small adjustments that further improve performance. The final phase is called *autonomous phase*, in which the learned tasks can be fulfilled without the need for conscious attention. Common examples include walking, speaking or riding a bike.

The opposite process, i.e. the *loss* of motor skills, is often related to neurodegenerative diseases where examples include Parkinson's, Alzheimer's or Huntington's disease [10]. These diseases affect the central nervous system which is responsible, among other things, for coordinating the limbs of the human body. The progress of such a disease is often intimately linked with a decline in motor abilities which can be one of the earliest symptoms. Currently, questionnaires such as the Parkinson's Disease Rating Scale [8], are used to get a subjective estimate of the state of a person's motor abilities. Combined with physical examinations by a clinician, these results form the basis for early diagnosis. An automatic assessment framework, which can be applied by the patient at home, would be very beneficial to improve early intervention and self-assessment. Other applications in medicine lie in rehabilitation and pain therapy. Here the way in which people move, the relationship between single limbs and how that develops over time can give valuable insights in the rehabilitation progress and the success of given medication. So far no objective, quantitative measures are available rendering assessment and diagnosis difficult and time-consuming.

Another example for an area where motor skill assessment is of interest is (professional) education and evaluation of technical skills. For many professions that require (high precision) manual performance such as surgery, students and professionals spend a lot of time in training. In order to improve training progress, immediate feedback on performance, technique and general strategy needs to be supplied. Providing continuous supervision can be very expensive, and – when an evaluation is performed – subjective. An automatic system that can be used to give feedback to people in training and provides an alternative to (manual) assessment frameworks (e.g. [7]) would be very advantageous and help to save money, time and improve the outcome.

Depending on the application, different requirements emerge towards skill assessment. While for many applications an analysis on the signal level can be sufficient, others might need to include higher-level transitions between activities to evaluate strategies. In this paper we want to focus on one crucial aspect which is the efficiency of motion. An efficient motion is one that requires the least amount of physical effort, reaches the goal quickly and therefore represents a shortest path. Especially during the cognitive phase, differences in efficiency between strategies are evaluated by the learner. By preserving just those strategies that require low amounts of effort, the pool of motions that is used to solve a task shrinks with increasing motor skill. When tasks are repeated, i.e. trained, the motions will therefore become more and more similar. This is intimately linked with the signal that captures the performer's movements, as this increasing similarity with increasing skill is also reflected on the signal level. One way to assess efficiency of motion therefore is to judge the self-similarity of the recorded sensor data. This self-similarity is related to the intrinsic complexity of the time-series and hence a measurement thereof will greatly contribute to motor skill assessment.

3. MEASURING COMPLEXITY FOR TIME-SERIES

The complexity of time-series is inverse to its predictability and self-similarity over time. Its measurement is of extensive interest in a variety of fields such as physics, where signals have to be differentiated to be either periodic, chaotic or the result of noise. Many different methods have been developed where the most common examples include Lyapunov-exponents [19], fractal dimensions [1], minimum embedding dimensions [6] or permutation entropy [4]. Most of these methods rely on the phase-space representation for time-series [15] in which chaotic and random movements exhibit very different shapes, the so called (strange) *attractors*. An attractor defines limits on possible orbits through the phase-space that a (deterministic) dynamical system can manifest on and provides a solid theoretical basis for the aforementioned methods. They work very well on time-series that are obtained by simulating low-dimensional dynamical systems but can break down when noise is added to the process [4].

Although providing a solid theoretical foundation, embeddings of real-life time-series data introduces an additional step in the analysis process that requires attention in parameter exploration. Furthermore especially AR data, as it is the result of captured human movement, is naturally exposed to noise on a fundamental level which can prove very difficult to compensate for. However, phase space representations have been used in AR [11] but their application is limited to periodic activity such as walking and running.

Another approach to estimate time-series complexity is to investigate the *goodness-of-fit* for a range of models with increasing model-complexity [5]. Statistical methods to investigate the goodness-of-fit include the Akaike Information Criterion (AIC) [3], the Bayesian Information Criterion [18] and their variants. However, providing a fine-grained range of models for comparison and ensuring their proper estimation (i.e. training) can be very difficult which limits the applicability of these statistical methods for the estimation of time-series complexity.

Although complexity estimates have been applied successfully to EEG and ECG data [14] they have not yet been

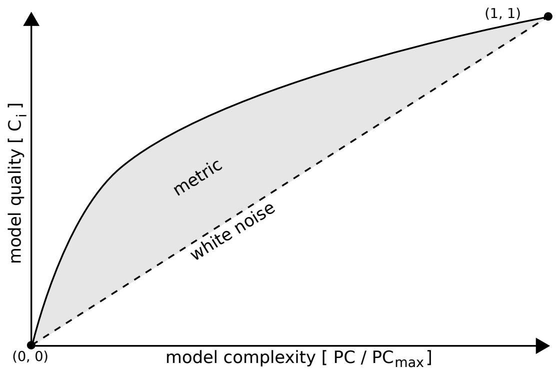

Figure 1: Distribution of energy along the PCs for a hypothetical data-set. The model complexity (fraction of PCs) influences the model quality (fraction of energy preserved).

used in AR. We believe that the field, and especially the assessment of motor performance, can hugely benefit from new methods that can cope with the involved noise, can be applied to all available data and provide simple evaluation.

4. PCA-BASED QUALITY ANALYSIS OF TIME-SERIES DATA

Motor skill assessment, i.e., the fine-grained analysis of the quality of certain movements, can be understood as a classical time-series comparison problem. However, in contrast to standard analysis techniques, which usually focus on qualitative analysis techniques, for example, utilizing clustering or supervised classification, there are certain constraints. First, a quantitative analysis of multi-modal sensor data is required. Ultimately we are interested in a metric of motor skills for an objective judgment of the way certain activities of interest are performed. Second, since we do not want to restrict skill assessment to some specific application, for example, by integrating detailed prior knowledge into the analysis, unsupervised techniques are favorable.

The idea of the proposed motor performance analysis approach is based on the assessment of model complexities. Assuming a well-controlled scenario, i.e., fixing the external variance of an activity of interest, the number of parameters a particular statistical modeling technique requires for capturing a particular sample of domain data represents a reasonable indicator for its complexity. The aforementioned well-controlled scenario is typically easy to achieve by focusing on the analysis of basic movements, which are building blocks of more complex activities and straightforward to segment. Skill analysis would then correspond to the analysis of the model parameters that are required to reasonably capture the signal of interest with a certain pre-defined quality.

Given the aforementioned constraints on skill assessment, and following the idea of model complexity analysis in this paper we concentrate on a Principle Component Analysis (PCA) based approach. PCA is a well established technique used for unsupervised dimensionality reduction be-

Algorithm 1 Estimate the PCA-based complexity metric for multivariate time-series data.

Input: time-series $S = s_1 \dots s_L,\ s_i \in \mathbb{R}^d$ (e.g. $d = 3$ for x, y, z acceleration. Window length f_w.
Output: normalised estimate of the area under energy distribution A_n.

$S' = whiten(S)$ {unit-mean and -variance for each axis}
$\{F\} = \text{extractFrames}(S', \{\text{length}=f_w, \text{stepsize}=1\})$
$(\boldsymbol{\lambda}, \vec{v}) = PCA(\{F\})$
for $i = 1$ **to** $|\boldsymbol{\lambda}|$ **do**
    $C_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{k=1}^{d} \lambda_k}$ {cumulative energy per PC}
end for
$C_0 = 0$
$A = \int_0^{d \times f_w} C$ {integrate cumulative energy}
$A_n = \frac{2A}{d \times f_w}$ {normalise, i.e. get fraction of max. area}
return $\boldsymbol{A_n}$

fore further processing is performed [13]. PCA estimates uncorrelated variables, the principal components (PCs), on the covariance of a (possibly correlated) set of observations. These PCs are sorted by the amount of energy that can be found along the corresponding direction in the input-space in decreasing order:

$$PCA(\{F\}) = (\lambda_j, \vec{v}_j), \quad j = 1 \dots d,\ \lambda_j \geq \lambda_{j+1} \qquad (1)$$

where $\{F\}$ are the input-vectors from $\mathbb{R}^d$ and $(\lambda_j, \vec{v}_j)$ is the sorted list of pairs of eigenvector $\vec{v}_i$ and eigenvalue λ_i of the sample set covariance. $\vec{v}_1$ is commonly referred to as the first PC, $\vec{v}_2$ as the second PC etc. For each PC λ_i we can compute the normalised cumulative energy C_i as

$$C_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{k=1}^{d} \lambda_k}. \qquad (2)$$

Dimensionality reduction is performed by projecting the input-vectors $\{F\}$ onto the subspace spanned by the first k PCs, i.e. $(\vec{v}_1 \dots \vec{v}_k)$, where k is less than or equal to the input dimensionality d. If k equals d, the projection corresponds to a rotation of the input space while all the variance and energy are preserved. If k is smaller, then just the fraction C_k of the total energy is preserved. The way in which the cumulative energy is distributed along the PCs is illustrated in figure 1 (solid line) and is closely linked to certain characteristics of the input data.

If the input vectors $\{N_i\}$ are an *iid* gaussian point-cloud, their energy is distributed equally across each spatial direction. Therefore along each PC the same amount of energy λ_i can be found: $\forall i, j \in \{1 \dots d\} : \lambda_i = \lambda_j$. The cumulative energy then corresponds to

$$C_i = \frac{\sum_{j=1}^{i} \lambda_j}{\sum_{k=1}^{d} \lambda_k} = \frac{\sum_{j=1}^{i} \lambda_1}{\sum_{k=1}^{d} \lambda_1} = \frac{i * \lambda_1}{d * \lambda_1} = \frac{i}{d}. \qquad (3)$$

This linear distribution of energy is illustrated in figure 1, denoted as *white noise*. The only way a distribution can deviate from this lower bound defined by white noise is if correlations and dependencies exist in the input-space. The distance from the lower bound is therefore a measure of how much intrinsic *structure* is contained in the input-space and

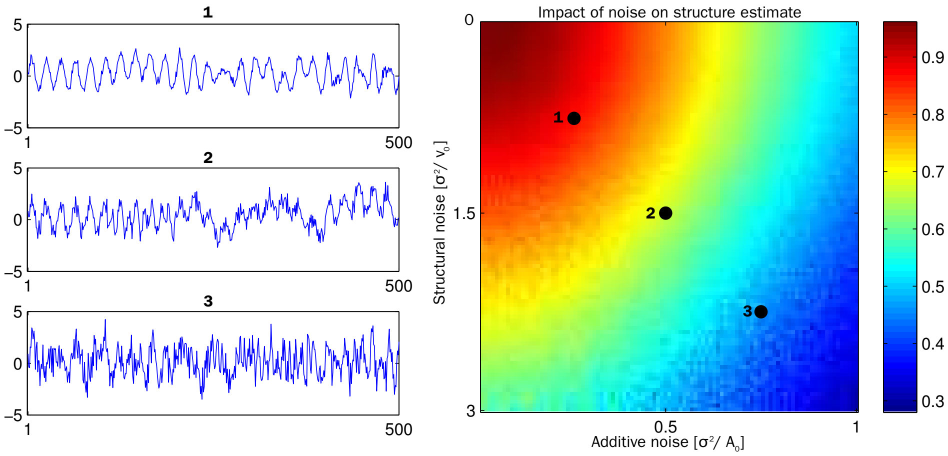

Figure 2: Evaluation of the developed metric on artificially created data. The graphs on the left show single samples of randomly generated signals according to the indicated parameters. The graph on the right shows the impact of both structural (vertical axis) and additive (horizontal axis) noise on the complexity estimate. (Best viewed in colour)

can be easily estimated by integrating along the distribution of energy. The result is a metric, which, normalised by the maximum possible area above the lower bound, can be used to compare different sets of input-vectors independent of their amount and dimensionality. The algorithm that can be used to estimate the metric for multivariate time-series is shown in algorithm 1. Here the data is whitened, i.e. each axis is normalised to have zero mean and unit variance, and frames are extracted with a sliding window procedure.

In order to demonstrate how a decrease in intrinsic structure affects the estimate of the described metric we will investigate artificially created signals. To allow comparison to common activity recognition data we work with 3 dimensional time-series that mimic recordings from triaxial accelerometers. The *clean* signal that we will distort artificially is one regular sinusoid along each of the three dimensions (or axis) of the time-series. We chose this signal for the sake of clarity and repeatability of our experiments. However, similar results are achieved with other signals such as sawtooth or compositions of multiple sigmoids. Recurring signals are a suitable choice here since multiple events of a non-repetitive human activity – such as getting out of bed – can still be concatenated to form a (noisy) repetitive signal.

In order to evaluate the impact of noise on the metric we introduce two different kinds of distortions: structural and additive noise. This relates to real-life time-series captured in AR in the sense that both differences in the strategy behind a motion (motion structure) as well as differences in the performance itself, i.e. the precision of the motion (motion noise), can occur. Structural noise is the distortion of the original sinusoid by introducing variance in frequency, amplitude and local biases. The distortion is performed independently for each axis. Given the base frequency υ_0 of the original signal S, the structurally distorted signal S^s is constructed as described in algorithm 2. With increasing variance σ_s^2 the resulting signal S^s will loose resemblance with the original signal as the introduced random signal gains weight.

Additive noise simply adds normally distributed noise to the signal after it has been created to form the final signal $S = (S_1, S_2, S_3)$:

$$S_i = S_i^s + \mathcal{N}(0, \sigma_a^2), \; i = 1, 2, 3. \tag{4}$$

By varying both variances σ_a^2 and σ_s^2 we can therefore control how much the initial sigmoids are distorted and hence how much intrinsic structure is contained in the signal. In contrast to using real-life data we can easily investigate the impact of noise on the metric. Low levels of additive noise correspond to a precise movements while low levels of structural noise correspond to a well chosen strategy for a motion.

By varying both aspects, mixtures of them (for example well chosen strategy but lack of precision) can be simulated. Figure 2 shows both samples of the signals under consideration and a heat-map that corresponds to the respective metric estimates. Three example signals that were generated according to the indicated parameters are shown in the plots

---

Algorithm 2 Add structural noise to a signal S.

Input: time-series $S = s_1 \dots s_L$, $s_i \in \mathbb{R}^3$, variance σ_s^2. Window length w_s
Output: Structurally distorted signal S^s.

```
for i = 1 to 3 do
    for j = 1 to L, j = j + w_s do
        v' ← N(0, σ_s^2)
        S' = sinusoid with frequency |v'| of length w_s
        S^s_{i,j→j+w_s} = S_{i,j→j+w_s} + S' × σ_s^2
    end for
    S^s_{i,:} = whiten(S^s_{i,:})
end for
return S^s
```

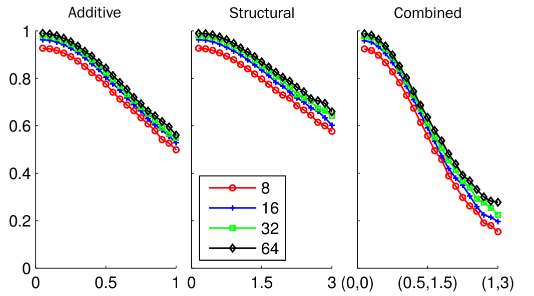

Figure 3: The three plots show the detailed impact of noise for the different disturbances for 4 different window-lengths. From left to right i) just additive noise, ii) just structural noise, iii) both structural noise and additive noise.

on the left. For clarity just one of the three axis and just a part of the sequence is shown. The heat-map on the right is generated by sampling signals of length 5000 samples with increasing structural as well as additive noise and each pixel is the colour-coded average of 10 evaluations. The base frequency of the sinusoids is 5 per 100 samples with an amplitude A_0 of 1 and a window-length f_w of 16 was used.

Figure 3 shows how the window-length influences the metric estimate. The impact of just additive noise, just structural noise and the combination (equal portions of additive and structural noise) exhibit a characteristic that is close to a linear decline. The window-length has a very predictable impact of the process: With growing width the whole spectrum is shifted towards 1. The main reason for this behaviour lies in the computation of the integral of the energy distribution. With growing window-length, simply more points are available for a precise interpolation.

The rapid decline of the metric for the case of combined structural and additive noise is a very desirable property, as these signals are a good model for real-life AR data that is subject to sensor noise and differences in performance of the subjects. The metric seems to be quite stable regarding structural noise by itself, which may be due to the construction of that particular disturbance. In [17] different models are introduced that allow fine-grained control of the determinism (or lack thereof) of time-series that will form the basis for future evaluations of the described metric and ease the comparison to other methods. Furthermore it will allow to investigate the suitability of the proposed metric for non-repetitive tasks in activity recognition.

In all our experiments the influence of the window length appears negligible and reasoned by computational precision issues. However, we cannot yet diminish a fundamental influence of it regarding the outcome of the metric, as we believe there is need for some prior knowledge about the relationship of sampling frequency to activity granularity when choosing this parameter. Consider, for example, audio signals and accelerometer signals. While the latter usually involves sampling rates from 30 to 100 Hz, audio signals are usually represented with (at least) 16kHz. However, the granularity, i.e. the amount of atomic activities over time is

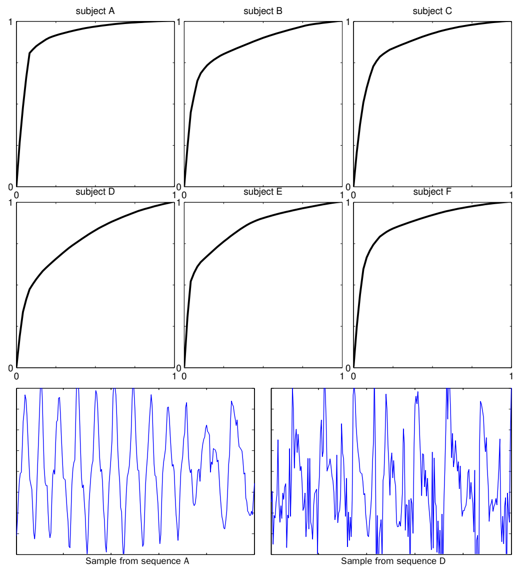

Figure 4: Example results for sequences from a simple experiment (whisking, see text). Both the energy distribution and samples of the raw signal (X-axis) are shown for two subjects.

comparable (e.g. number of phonemes in speech vs. strokes when chopping a carrot). Future experiments will explore the impact of very high resolution signals.

5. CASE STUDY

In order to investigate the applicability of the described metric for real-life AR data we performed a simple cooking experiment, in which we asked six subjects to whisk (single) cream for 5 minutes with a sensor-equipped utensil. The tri-axial accelerometer in the handle of the whisk recorded with a sampling frequency of 100 Hz. After recording, the data was segmented by hand to solely represent the continuous whisking motion as it could also be done automatically by an activity recognition system. The estimated metric values are shown in table 1. Figure 4 shows the energy distributions and two samples of the raw signal. All the subjects had different approaches to the whisking task. While subject A used a regular and circular motion, the movements of subjects D were less coordinated and varied strongly over time. While already few PCs suffice to preserve large amounts

Subject	Score
subject A	**0.82**
subject B	0.69
subject C	0.72
subject D	**0.54**
subject E	0.68
subject F	0.73

Table 1: Evaluation of the metric on a simple (segmented) whisking task.

of the energy present in signal A it takes many more PCs to reach similar levels of model quality for signal B. Hence according to the metric, signal A contains more intrinsic *structure* than signal B, which corresponds to an intuitive interpretation of the raw signal samples in figure 4.

We chose this task because we can safely assume that an increased amount of structure, i.e. a precise and regular motion, is a sign of a superior motor skill of the performing subject. Manual skill assessment is nevertheless very difficult to perform and therefore a quantitative evaluation of the metric (i.e. a comparison with a set of human experts) remains to be done. Nevertheless the estimates of the metric on this experiment conform to the intuitive assessment of the subject's performance.

Future experiments with more complex activities, and especially the application of the method to data collected of people with Parkinson's disease will give further insights into the link between the estimated signal complexity and motor skill. Its quantitative evaluation will be supported by the existing classification and rating systems.

6. CONCLUSION

So far, relatively little work was invested into the detailed analysis of segmented human activity data. However, information about the motor performance of a subject, i.e. how *well* an activity was performed, is of interest for a variety of novel applications in AR. This motor skill assessment represent as significant challenge, as requirements can be very diverse across applications and domains. An automatic system is nevertheless much desired in fields such as medicine, where it could improve diagnosis and intervention in neurodegenerative conditions such as Parkinson's disease.

The efficiency of motion is a crucial aspect shared by many applications and has implications for the inherent complexity of data recorded with body-worn or pervasive sensors. We developed a novel method to quantify this signal complexity that is based on the energy distribution in PCA and results in a single, normalised metric which can be used to compare (subject-specific) time-series independent of their length and dimensionality. Using artificially distorted signals we showed the sensitivity of the method to variations in signal complexity and applied it to real-world AR data from a simple kitchen task to demonstrate its applicability. The very few parameters of the method allow its application without the need of detailed expert knowledge about the activities of interest which represents its main advantage over other tools for time-series analysis.

Future experiments will involve a quantitative evaluation of the metric that is based on data collected with people with Parkinson's disease and expert classification provided by clinicians. Furthermore the transferability of the general idea to other methods such as predictive models will be investigated and holds potential for the incorporation of domain-specific knowledge.

7. REFERENCES

[1] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biological Cybernetics*, 77(5):339–350, 1997.

[2] A. Ahmadi, D. Rowlands, and D. James. Towards a wearable device for skill assessment and skill acquisition of a tennis player during the first serve. *Sports Technology*, 2(3-4):129–136, 2009.

[3] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 2002.

[4] C. Bandt and B. Pompe. Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, 88(17):174102, 2002.

[5] R. Brooks and A. Tobias. Choosing the best model: Level of detail, complexity, and model performance* 1. *Mathematical and Computer Modelling*, 24(4):1–14, 1996.

[6] L. Cao. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110(1-2):43–50, 1997.

[7] V. Datta, S. Bann, M. Mandalia, and A. Darzi. The surgical efficiency score: a feasible, reliable, and valid method of skills assessment. *The American Journal of Surgery*, 192(3):372–378, 2006.

[8] S. Fahn, R. Elton, et al. Unified Parkinson's disease rating scale. *Recent developments in ParkinsonâĂŹs disease*, 2:153–163, 1987.

[9] P. Fitts and M. Posner. Human performance. 1967.

[10] M. Forman, J. Trojanowski, and V. Lee. Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs. *Nature medicine*, 10(10):1055–1063, 2004.

[11] J. Frank, S. Mannor, and D. Precup. Activity and Gait Recognition with Time-Delay Embeddings. In *Proceedings of AAAI Conference on Artificial Intelligence*, 2010.

[12] R. Grober. An Accelerometer Based Instrumentation of the Golf Club: Measurement and Signal Analysis. *Arxiv preprint arXiv:1001.0956*, 2010.

[13] I. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer, 1986.

[14] H. Kantz, J. Kurths, and G. Mayer-Kress. *Nonlinear analysis of physiological data*. Springer Berlin etc., 1998.

[15] H. Kantz and T. Schreiber. *Nonlinear time series analysis*. Cambridge Univ Pr, 2004.

[16] R. King, L. Atallah, B. Lo, and G. Yang. Development of a wireless sensor glove for surgical skills assessment. *Information Technology in Biomedicine, IEEE Transactions on*, 13(5):673–679, 2009.

[17] K. Narendra and K. Parthasarathy. Identification and control of dynamical systems using neural networks. *Neural Networks, IEEE Transactions on*, 1(1):4–27, 2002.

[18] D. Spiegelhalter, N. Best, B. Carlin, and A. Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

[19] A. Wolf, J. Swift, H. Swinney, and J. Vastano. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3), 1985.