Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 96 (2016) 465 – 474

Procedia Computer Science

20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2016, 5-7 September 2016, York, United Kingdom

Comparing and combining time series trajectories using Dynamic Time Warping

Neil Vaughan[a]*, Bogdan Gabrys[a]

[a]Bournemouth University, Talbot Campus, Fern Barrow, Poole, BH12 5BB, United Kingdom

Abstract

This research proposes the application of dynamic time warping (DTW) algorithm to analyse multivariate data from virtual reality training simulators, to assess the skill level of trainees. We present results of DTW algorithm applied to trajectory data from a virtual reality haptic training simulator for epidural needle insertion. The proposed application of DTW algorithm serves two purposes, to enable (i) two trajectories to be compared as a similarity measure and also enables (ii) two or more trajectories to be combined together to produce a typical or representative average trajectory using a novel hierarchical DTW process.
Our experiments included 100 expert and 100 novice simulator recordings. The data consists of multivariate time series data-streams including multi-dimensional trajectories combined with force and pressure measurements.
Our results show that our proposed application of DTW provides a useful time-independent method for (i) comparing two trajectories by providing a similarity measure and (ii) combining two or more trajectories into one, showing higher performance compared to conventional methods such as linear mean. These results demonstrate that DTW can be useful within virtual reality training simulators to provide a component in an automated scoring and assessment feedback system.

Peer-review under responsibility of KES International

Keywords: Dynamic Time Warping; Virtual Reality; Training Simulator; Trajectory Data; Combining; Comparing;

1. Introduction

Virtual Reality (VR) training simulators are increasingly used to enable trainees to practice tactile procedures with hand-eye coordination. In medical training Epidural is one procedure that requires accurate needle insertion and

* Corresponding author. Tel.: +44-1202-965-038.
E-mail address: nvaughan@bmth.ac.uk

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2016.08.106

VR can provide an ideal platform for training and practice. VR simulators produce data accurately recording the movement, angle and forces on virtual tools, the speed and accelerations of the operators hand motion. This data is multivariate data-streams including trajectories combined with multi-dimensional force and pressure measurements.

Machine learning and data analytics processes are required to analyse the trajectory data and compute a skill level or score to feedback to the trainee. Data analytics could autonomously identify key areas which require practice and give indication of when a trainee is ready to progress to more difficult scenarios.

Two key components which are missing but are required in order to produce an automated scoring system are (i) comparison between two insertions such as a similarity measure and (ii) combination of numerous trajectories into one, to produce an average insertion representing an individual's typical technique.

This research proposes that Dynamic Time Warping (DTW)[1] algorithm could provide a method for achieving both comparison and combination between time series data. These two components can form an important part of an automated scoring and assessment system for virtual reality simulator training. This research presents results and performance assessment from implementations of both DTW and the conventional method using linear mean aiming to identify the optimum method for comparison and combination of trajectory data.

In order to test our implementations and compare performance of DTW with linear mean, we conducted a data collection trial which has recorded time series from 100 expert and 100 novices using a VR simulator for epidural needle insertion. This data revealed numerous problems and difficulties when producing a comparison between two insertions. Insertions commonly take different lengths of time and the landmarks in the insertions therefore occur at different locations in the time series. This makes a mean linear scaling difficult or implausible, as merging two time series with different insertion rates could cause landmarks to become merged, causing blurring and loss of data. To achieve higher accuracy when comparing or combining two trajectory time series, a non-linear approach is required and this research proposes DTW for this purpose.

2. Literature Review

For the analysis of time series or trajectory data, a variety of distance functions have been used to assess the similarity or dissimilarity between two time series[2]. The most popular distance function is Euclidean distance[3]. The longest common subsequence (LCSS) is another distance function, specialized for hierarchical data with high noise[4]. There are over 12 distance functions[5] including edit distance, longest common distance, dynamic time warping (DTW), Sequence Weighted Alignment model (Swale). Recently benefits of DTW were highlighted as a component in an automated clustering system for trajectory data, combined with particle swarm optimisation[2]. No previous studies have applied DTW to combine trajectories, or as a means of analysing data from VR simulators.

There is relatively little literature on comparing two trajectories but one use is for clustering[6]. Methods for clustering time series typically apply methods for clustering static data converted to work with time series. Comparison between multivariate time series trajectories is more complex. Time-series are dynamic data and a current research challenge is to recognize dynamic changes in time series[7]. This justifies our use of dynamic methods such as DTW in this research.

We previously conducted thorough literature surveys of virtual reality simulators[8,9]. Our surveys found that data analysis is lacking in VR simulators but it has potential to offer benefit enabling the autonomous generation of scores and assessment.

2.1. Time Series Data Representation

In our data collection trial (section 4), data was recorded in form of a set of N trajectories where N was 200, comprised of 100 experts and 100 novices: (t): $t_1, t_2, t_3, \ldots, t_i, \ldots, t_N$. Each trajectory t consists of P coordinates, where P is the length of trajectory which is relative to time taken as shown in Eq. 1. Each coordinate consists of 3 values (x, y, z) to represent 3D space and additionally 2 values force (f) and pressure (p) were recorded. The state of an object t_i varies from time step 0 to P and at timestep k it is represented by $(x_{ik}, y_{ik}, z_{ik}, p_{ik}, f_{ik})$.