Motus: A Framework for Human Motion Classification in a Not-controlled Moving Environment

Motus: Marco de trabajo para la clasificación de Captura de Movimiento humano en ambientes no controlados

Joselyn Rodríguez-González[1], María Hernández-López[2], Francisco Siles-Canales[3]

Rodríguez-González, J; Hernández-López, M; Siles-Canales, F. Motus: A Framework for Human Motion Classification in a Not-controlled Moving Environment. *Tecnología en Marcha.* Edición especial 2020. 6th Latin America High Performance Computing Conference (CARLA). Pág 55-59.

https://doi.org/10.18845/tm.v33i5.5076

1 Centro de Investigación de Ingeniería en Sistemas (CIIS), Escuela de Ingeniería en Sistemas, Universidad Latina de Costa Rica, Costa Rica. Email: joselyn.rodriguez@ulatina.cr.
https://orcid.org/0000-0003-4326-5036

2 PRIS-Lab: Pattern Recognition and Intelligent System Laboratory, Department of Electrical Engineering, School of Engineering, Universidad de Costa Rica, Costa Rica. Email: maria.hernandezlopez@ucr.ac.cr.
https://orcid.org/0000-0003-2548-0277

3 PRIS-Lab: Pattern Recognition and Intelligent System Laboratory, Department of Electrical Engineering, School of Engineering, Universidad de Costa Rica, Costa Rica. Email: francisco.siles@ucr.ac.cr.
https://orcid.org/0000-0002-6704-0600

Keywords

MoCap; classification; segmentation; feature selection; data cleansing.

Abstract

This work introduces a framework proposal based on various algorithms, processes, and methods to classify Motion Capture (MoCap) data. To provide a generalized model for MoCap data classification, the approach is defined step by step: data collecting, data cleansing, segmentation, data pre-processing, feature selection, model selection, and validation. For each step, we selected and evaluated algorithms, process and methods have shown good performance in previous studies, all of them were proved and validated in BVH databases, but in not freely moving environment.

Palabras claves

MoCap; clasificación; segmentación; selección de características; limpieza.

Resumen

Este trabajo presenta la propuesta para un marco de trabajo (Motus) basado en varios algoritmos, procesos y métodos para la clasificación de archivos de captura de movimiento. El objetivo es proveer un modelo generalizado para la clasificación de movimiento, el enfoque es defino bajo los siguientes pasos: recolección y limpieza de datos, pre-procesamiento de la información, selección de características, segmentación y la selección del modelo y la validación. Para casa uno de los pasos, se selección y evaluó algoritmos que han demostrado un buen rendimiento, en ambientes controlados, en estudios previos, todos ellos han sido probados en archivos BVH, pero no en ambientes no controlados.

Introduction

The classification of human body motion from MoCap data is a difficult problem [1]. Also, the automatic segmentation of image sequences containing more than one class of motion is particularly challenging [1].

Although comparing two motion sequences is an easy task for a person, an automatic comparison is hard due to enormous numerical differences between two similar motion sequences. Spatial variations are mostly due to almost rigid transformations among similar postures. Temporal variations are due to non-linear differences in the dynamic of one action when performed by different subjects or even by two different performances of the same subject [2]. Because of these issues, it is highly important to define an appropriate procedure for a validating framework. It is then essential to define an automated technique that can segment MoCap data into homogeneous intervals, classify each interval into a motion type and produce minimal errors [3].

Related Work

After the MoCap data has been recorded, there are several essential steps that one must take with it in order to make it into an analyzable format [4].

To generate the continuous MoCap data usable for analyses, it needs to be imported with the video material into a data management program in which it can be segmented into more processable chunks [4].

More than a few methods have been previously investigated in the pursuit of this task; such as the weighted principal component analysis [5] which allows finding a particular motion in a large movement sequence; in addition, image-based reconstruction has also played an important role in motion classification, this may be done by creating a comprehensive set of training data for motion recognition methods with views of motion from all angles, or by converting non-orthogonal views taken from a single camera into orthogonal views for recognition [6].

It is important to mention that the literature reviewed for this investigation is based on very controlled motion capture; as for the present work, it is necessary to implement and compare both algorithms in a more freely moving environment.

Methods

Validation Design

The framework was designed considering every single step to motion capture automatic classification, for collecting data we take bvh files, some of them were controlled environment and human movement samples from the motion capture database of Carnegie Mellon University (CMU), and the others came from our records in a not-controlled sceneries. Our data set need to be processing by methods to cleaning data, reduce noising, and normalize, in this research we propose some procedures to correct it. In order to segment motion, we chose two algorithms for each one of the following process: classify data, feature selection and segmentation, all of them with a high performance.

Subjects and data set

For recording the Motion Capture sequences students were selected between the ages of 20 and 24 years old. A mixture of males and females was generated in order to obtain a more wide-range set of results. A total of three sessions were recorded, each one with its own BVH file set.

Results

Framework proposal

Collecting data

Data were obtained by recording interacting sessions between volunteer subjects. There was a total of 8 volunteers and the sessions were divided as follow: The first session involving an interaction between a male and a female of 21 and 23 years old respectively, with a duration of approximately 30 minutes. The environment was set so that joint attention would be initiated by any one of the subjects. In order to do that, no instructions were given to neither subject, but conversation cards were placed in a table to promote interaction between each other. Session two was developed in a very similar manner except that it involved three males of ages between 20 and 24 years old; while session three was a reiteration of session one except that subjects were replaced with new volunteers.

Cleaning data & preprocess data

This variability suggests that factors such as lighting conditions, camera angle, participant body shape, clothing and stance have a noticeable impact upon the quality of the recording.

These types of ‘Glitches’ are common in motion capture causing legs or arms to twitch rapidly into violently bent angles [7]. A solution is the algorithm (figure 1) for the detection, location and correction of glitches in a BVH file, Castresana [8] implemented in Python an algorithm for glitches reduction and will be used in the validation of the framework.

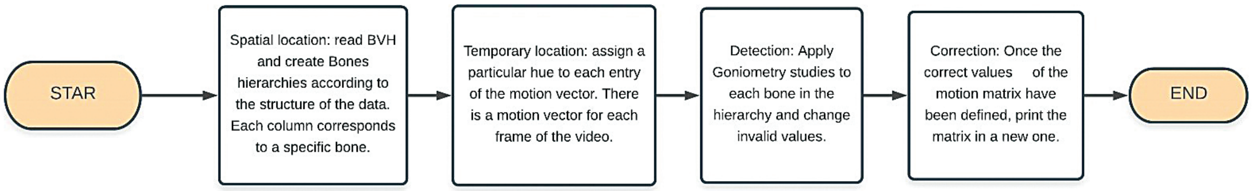

Figure 1. Algorithm Architecture [8]

Segmentation

A genetic algorithm approach to human MoCap data segmentation is applied to obtain the optimal solution, accuracy and efficiency of segmentation in controlled data. They conduct sparse learning on the raw motion data to derive a dictionary of representative postures and convert the raw motion sequence into a symbolic sequence [9]. As well, Quaternion Watershed Transform in Segmentation of Motion Capture Data controlled has a good performance, but in this case, they use the following segmentation hierarchical has still acceptable accuracy, which exceeds 91% [10]. Compare both methods will be the next step.

Feature selection (see figure 2):

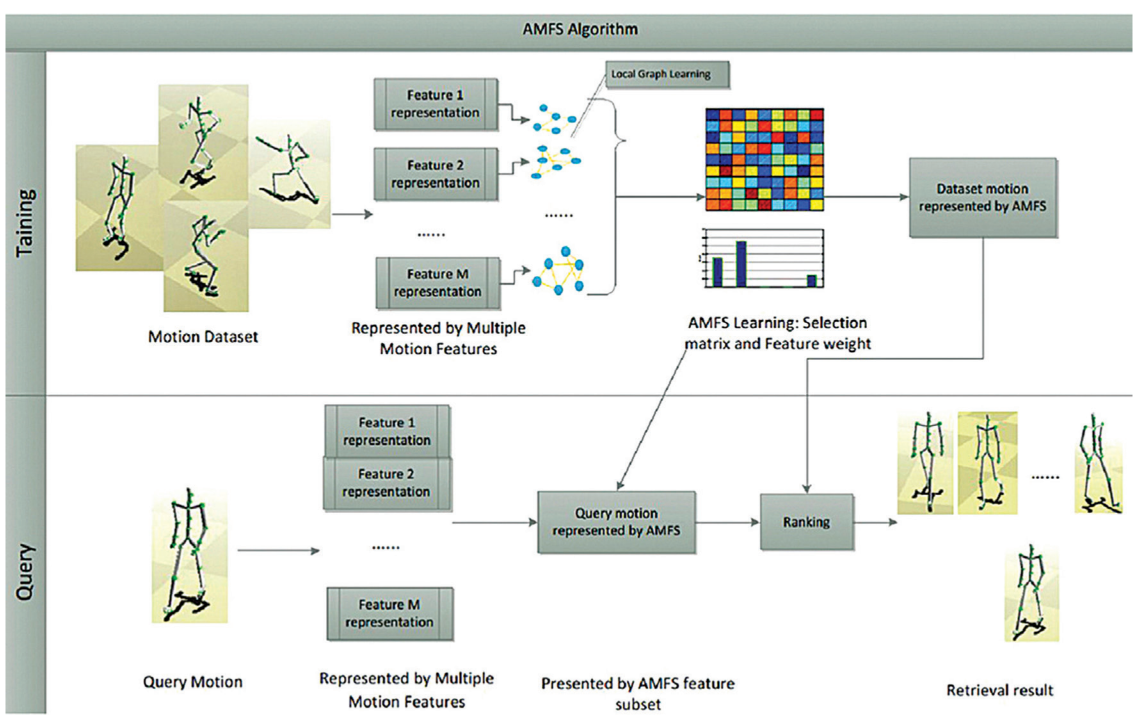

Figure 2. Adaptive multi-view feature selection (AMFS) [11]

Model Selection:

Two-Step SVM Fusion: Combine tree-structured vector quantization and Multiple binary Support Vector Machine classifiers [3]. The proposed algorithms using the 5-fold cross validation procedure and obtained a correct classification rate of 99.6% [3].

Future Work

The framework presented in this work is part of ongoing research, which intends to development a combination of algorithms, processes and methods mentioned in the previous sections and apply it, in order to validate a methodology that helps to classify not controlled motion capture and complex poses.

References

[1] J. Rittscher and A. Blake, "Classification of human body motion," Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 1999, pp. 634-639 vol.1. doi: 10.1109/ICCV.1999.791284

[2] A. W. Vieira, T. Lewiner, W. R. Schwartz and M. Campos, "Distance matrices as invariant features for classifying MoCap data," Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, 2012, pp. 2934-2937.

[3] H. Kadu and C. -. J. Kuo, "Automatic Human Mocap Data Classification," in IEEE Transactions on Multimedia, vol. 16, no. 8, pp. 2191-2202, Dec. 2014. doi: 10.1109/TMM.2014.2360793

[4] Jantunen, T., Buerger, B., De Weerdt, D., Seilola, I., & Wainio, T. Experiences Collecting Motion Capture Data on Continuous Signing. Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions Between Corpus and Lexicon, 75-82. 2012

[5] Forbes, K. & Fiume, E. An efficient search algorithm for motion data using weighted PCA. Eurographics/ACM SIGGRAPH Symposium on Computer Animation. USA, 2005 p.p. 67-76, Doi: 10.1145/1073368.1073377

[6] Bodor, R., Drenner, A., Fehr, D., Masoud, O., & Papanikolopoulos, N. View-independent human motion classification using image-based reconstruction. Image and Vision Computing, 27(8), 1194–1206. July, 2009. Doi: 10.1016/j.imavis.2008.11.008

[7] Gillies, M., Brenton, H., Yee-King, M., Grimalt-Reynes, A., & d' Inverno, M. Sketches vs skeletons. Proceedings of the 2nd International Workshop on Movement and Computing - MOCO '15. Canada, August, 2015, p.p. 104-111 doi:10.1145/2790994.2790995

[8] M. Castresana and F. Siles, "Goniometry-based Glitch-Correction Algorithm for Optical Motion Capture Data," 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), San Carlos, 2018, pp. 1-8. doi: 10.1109/IWOBI.2018.8464198

[9] Lv F., Nevatia R. Recognition and Segmentation of 3-D Human Action Using HMM and Multi-class AdaBoost. In: Leonardis A., Bischof H., Pinz A. (eds) Computer Vision – ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, vol 3954. Belin, 2006- Doi: 10.1007/11744085_28

[10] Świtoński A., Michalczuk A., Josiński H., Wojciechowski K. (2019) Quaternion Watershed Transform in Segmentation of Motion Capture Data. In: Nguyen N., Gaol F., Hong TP., Trawiński B. (eds) Intelligent Information and Database Systems. ACIIDS 2019. Lecture Notes in Computer Science, vol 11432. March, 2019 . Doi: 10.1007/978-3-030-14802-7_49

[11] Zhao Wang, Yinfu Feng, Tian Qi, Xiaosong Yang, Jian J. Zhang, Adaptive multi-view feature selection for human motion retrieval, Signal Processing, Volume 120, 2016, p.p. 691-701, ISSN 0165-1684, doi: 10.1016/j.sigpro.2014.11.015.