Automatic Motion Segment Detection & Tracking
for Predictive Animation Solutions

Ben Kenwright
*

Abstract

The rapid increase in digital media, specifically animation data, from a variety of fields (video games, robotics, medical analysis) has introduced new challenges. This includes, the managing and searching of these huge unstructured animation libraries. In this paper, we present an novel approach for identifying and matching character features within animation libraries for poses or animation-segments to predict motions (e.g., responding to unforeseen disturbances in a virtual environments). We demonstrate simple pose and pattern matching principles before going onto identifying and searching for other feature characteristics, such as, mood and energy. Rhythmic properties are a central component that we use to classify motion segments (e.g., Fourier analysis). We incorporating different search features with pattern matching algorithms to reduce look-up times and provide a helpful tool for artists and practitioners - i.e., enabling the search and comparison of data outside the visual eye.

Keywords: animation, patterns, motion, detection, segments, poses, characters, articulated skeletons

1 Introduction

Why would we need to search and match similar features within animation databases? This could be for motion blending or mixing similar animations to accomplish targeted actions, automatic categorization and organisation to identify faults or issues within a motion. Animation data is used for a wide variety of fields, from video games to medical analysis. The topic covers a diverse set of animation systems. This paper, focuses specifically on character animations - those with articulated rigid skeletons.

Rapid increase in digital character animation solutions. Everyday in computer generated films, video games, robotics, and medical analysis fields, we generate tera-bytes of key-frame animation data. This information comes from motion capture data, artists, or procedural solutions. This is a huge untapped resource that provides valuable information. The challenge is efficiently organising the data so we are able to search, identify and retrieve specific information. With this in mind, motion pattern identification and re-targeting has been an active area of research for the past two decades [Arikan et al. 2003; Müller et al. 2009; Keogh et al. 2004].

Contribution The key contributions of this paper are: (1) search and identify similar motions within a large database of key-frame animations using a variety of feature classifications (e.g., overall centre of mass frequency, energy, and joint displacement); (2) ability to instantly predict future motion patterns based on past behaviours;

*

2 Related Work

The importance of the problem has resulted in a number of solutions that automatically retrieve and classify motion data [Müller et al. 2005; Forbes and Fiume 2005; Keogh et al. 2004; Kovar and Gleicher 2004; Chiu et al. 2004]. The core difficulties evolve around the fact that 'similar' animations may possess a number of differences (e.g., skeletal or numerical). A popular solution in the literature is to use motion representations to match semantically raw data. Problems occur when posses are drastically different or we need to identify sub-motions within an action (e.g., nervous disposition that sits on top of a fundamental animation, such as, walking or climbing). A different approach, as used by Liu et al. [Liu et al. 2005] and Müller et al. [Müller et al. 2005], works by absorbing spatial and temporal variations within the feature level to allow for accurate motion comparison. In our approach, we focus on using a variety of different feature comparison solutions to identify and categorize different motions (e.g., searching for specific features within a motion set, such as, a pose or an embedded emotion).

Pre-recorded animations, using motion capture technologies, introduces an era of realism that is difficult to emulate using other methods. Having said that, data-driven methods are the dominant solution for creating computer animations [Arikan et al. 2003; Cooper et al. 2007; Pullen and Bregler 2002]. Rose et al. [Rose et al. 1998] provided an annotation solution that grouped similar motions into 'verb' classes to search and synthesize new user controlled motions. Notably, this was followed by the popular concept of motion graphics by Kovar et al. [Kovar et al. 2002]. After which a variety of annotation and hybrid methods appeared [Arikan et al. 2003; Müller et al. 2009].

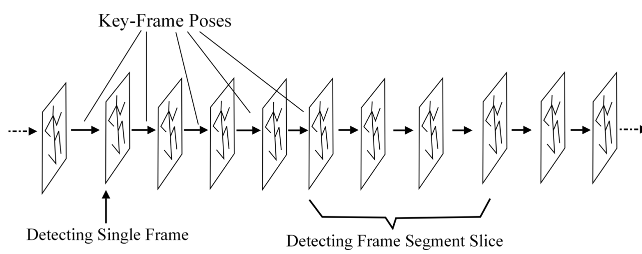

Figure 1: ***Tracking*** *- Detecting individual poses or frame segments in animations.*

3 Method

Feature Identification Articulated animated characters posses a huge amount of flexibility. Different animation libraries may not necessarily contain skeletons with the same physical-characteristics, such as, number of joints, mass, or physical dimensions. Hence, to help identify and categorize different animations, features include:

- joint angles (local and world transforms)
- displacement (end-effectors)

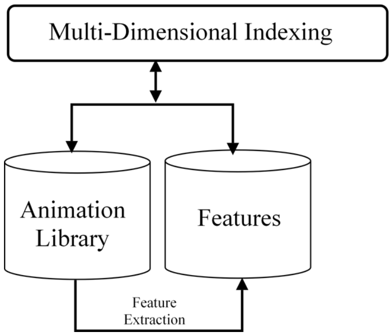

Figure 2: ***High-Level Interconnected Elements*** *- Storing indexes and features for a large animation library allow for faster search times.*

- velocity and energy (overall or specific joints)
- global parameters (**centre-of-mass** and centre-of-pressure)
- frequency domain (e.g., Fourier transform [Unuma et al. 1995])

Number of Features vs Error and Search Times Specifying a larger number of features reduces the classification error. However, an increase number of features makes the algorithm less flexible and requires greater search time. Another key thing to remember, is animation libraries may not always have all the search features. Since animation libraries are diverse and store different information (e.g., number of degrees of freedom and time-space resolution). The feature of spine posture information might not be available within the animation data set, but is important for a specific classification (see Figure 3).

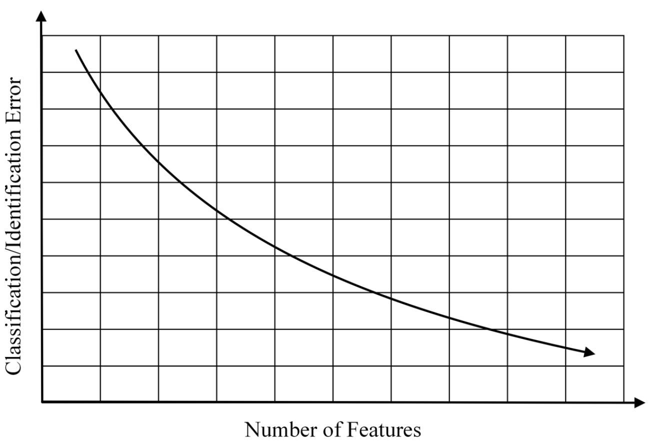

Figure 3: ***Features and Error*** *- As the number of features is increased the test error for identifying and classifying animation segments reduces.*

Training Optimisation or Statistical Methods A variety of techniques are available for categorizing and identifying patterns in multi-dimensional data sets. For example, adaptive solutions, such as, Neural Networks - and signal analysis methods, such as, Fourier transforms [Unuma et al. 1995].

Coarse Search Space Key-frame animation libraries are store discrete form (i.e., incremental time-slices). We might have situations that require us to search for in between frames. Intermediate frames between two poses and gives the appearance that the first pose evolves smoothly into the second pose. This helps increase the search space fidelity by incorporating a simple second order linear predictor, as shown below in Equation 1.

$$\vec{\theta}(t) = \vec{\theta}_n + \left(\vec{\theta}_n - \vec{\theta}_{n+1}\right)\ t \tag{1}$$

where $\vec{\theta}_n$ and $\vec{\theta}_{n+1}$ represent animation frame n and $n+1$, while $\vec{\theta}(t)$ is the approximate linear prediction ($t \in 0 : 1$). For example, joint angles are commonly stored as Euler angles in motion capture format which can be linearly interpolated, while for other formats, such as, quaternions, spherical linear interpolation is available.

Differences and Pose-Frame Matching For two sets of feature data (i.e., one from the reference and one from the search data). We need to determine a measurable value (ψ) between two animation frames. We accomplish this using the inner product, as shown below in Equation 2.

$$\psi = \frac{\vec{q}_r \cdot \vec{q}_n}{|\vec{q}_r|\ |\vec{q}_n|} \tag{2}$$

where $\vec{q}_r$ is the reference array of feature data and $\vec{q}_n$ is the corresponding frame feature data. Equation 2 allows us to see how the features within an animation change over time.

$$\psi_n = \frac{\vec{q}_n \cdot \vec{q}_{n+1}}{|\vec{q}_n|\ |\vec{q}_{n+1}|} \tag{3}$$

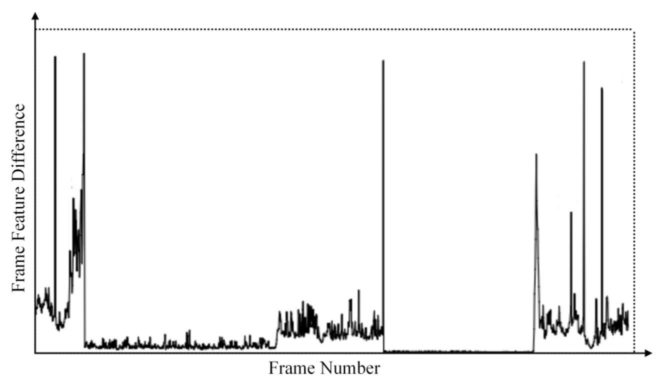

Figure 4: ***Frame Feature Deviation*** *- Plot of ψ from Equation 3 for frame-frame feature changes. The spikes show abrupt pose changes within the animation (e.g., trips and sudden impacts). The information also provides a visual means for searching for changes when browsing an animation libraries.*

Limitations We focused on human topology key-frame libraries. However, since some of the feature classification methods are topology independent, for instance, energy and Fourier frequency decomposition, our approach should be applicable to a diverse set of motion analysis systems, such as, animals like quadrupeds.

4 Conclusion

In this paper, we presented a novel approach for searching and categorizing large unstructured collections of character animation libraries to identify targeted features. Combining different feature searches, such as, Fourier frequency harmonics, centre-of-mass velocity, and frame-frame displacement, reduces the search error. We showed various experiments to demonstrate the practicality of our automated approach. Our solution is generic and flexible in the sense that it allows for expansion and customization to fit a variety of character animation searches. The searching and classifying of different key-frame motion segments goes beyond an artistic tool (i.e., for generating new animations). For example, reducing the error means the technique would allow medical practitioners to search pre-recorded patient movements for ailments. Notwithstanding, key-frame character animations does not just include 'bipeds' but can be extended to a variety of skeleton animation topologies.

Acknowledgements

We would like to thank the reviewers for taking time out of their busy schedules to provide constructive and insightful feedback that has went into making this paper clearer and more concise.

References

ARIKAN, O., FORSYTH, D. A., AND O'BRIEN, J. F. 2003. Motion synthesis from annotations. *ACM Transactions on Graphics (TOG) 22*, 3, 402–408. 1

CHIU, C.-Y., CHAO, S.-P., WU, M.-Y., YANG, S.-N., AND LIN, H.-C. 2004. Content-based retrieval for human motion data. *Journal of Visual Communication and Image Representation 15*, 3, 446–466. 1

COOPER, S., HERTZMANN, A., AND POPOVIĆ, Z. 2007. Active learning for real-time motion controllers. In *ACM Transactions on Graphics (TOG)*, vol. 26, ACM, 5. 1

FORBES, K., AND FIUME, E. 2005. An efficient search algorithm for motion data using weighted pca. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM, 67–76. 1

KEOGH, E., PALPANAS, T., ZORDAN, V. B., GUNOPULOS, D., AND CARDLE, M. 2004. Indexing large human-motion databases. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, VLDB Endowment, 780–791. 1

KOVAR, L., AND GLEICHER, M. 2004. Automated extraction and parameterization of motions in large data sets. In *ACM Transactions on Graphics (TOG)*, vol. 23, ACM, 559–568. 1

KOVAR, L., GLEICHER, M., AND PIGHIN, F. 2002. Motion graphs. In *ACM transactions on graphics (TOG)*, vol. 21, ACM, 473–482. 1

LIU, G., ZHANG, J., WANG, W., AND MCMILLAN, L. 2005. A system for analyzing and indexing human-motion databases. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, ACM, 924–926. 1

MÜLLER, M., RÖDER, T., AND CLAUSEN, M. 2005. Efficient content-based retrieval of motion capture data. In *ACM Transactions on Graphics (TOG)*, vol. 24, ACM, 677–685. 1

MÜLLER, M., BAAK, A., AND SEIDEL, H.-P. 2009. Efficient and robust annotation of motion capture data. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ACM, 17–26. 1

PULLEN, K., AND BREGLER, C. 2002. Motion capture assisted animation: Texturing and synthesis. In *ACM Transactions on Graphics (TOG)*, vol. 21, ACM, 501–508. 1

ROSE, C., COHEN, M. F., AND BODENHEIMER, B. 1998. Verbs and adverbs: Multidimensional motion interpolation. *Computer Graphics and Applications, IEEE 18*, 5, 32–40. 1

UNUMA, M., ANJYO, K., AND TAKEUCHI, R. 1995. Fourier principles for emotion-based human figure animation. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, ACM, 91–96. 2