

Kompression menschlicher Bewegungsdaten: Eine Übersicht

1. Abstrakt:

Bewegungserfassungstechnologien verzeichnen in den letzten Jahren erhebliche Fortschritte, die im Bereich der Erzeugung spezieller 3D Effekte in Videos, der Analyse und Modellierung von menschlichen Bewegungen, sowie der Bewegungssimulation und des Biofeedbacks zur Verbesserung der Leistung im Sport angewendet werden. Eine direkte Nutzung der Messdaten ist unpraktisch und aufwändig. Insbesondere bei hoher Abtastrate und großer Anzahl der Marker ist der Bedarf an Speicher und Bandbreite sehr groß. Um die relevante Information kompakt und schnell zugänglich zu machen, ist Kompression notwendig. Dieser Beitrag liefert eine Zusammenfassung und Taxonomie der aktuellen Kompressionstechniken von menschlichen Bewegungsdaten.

2. Einleitung:

Die menschliche Bewegung wird durch die Aufzeichnung der räumlichen Positionen von Markern erfasst, die normalerweise an den Gelenken des Körpers befestigt sind. Bewegungsdaten sind in konstanten Zeitintervallen abgetastete Signale, die 3D Koordinaten der Position und bei manchen Geräten die Orientierung jedes Markers enthalten. Ein Marker wird mit 1 bis 6 Freiheitsgraden („DOF“) dargestellt. Die Position eines Markers als Funktion der Zeit wird als Bewegungskurve bezeichnet. Die Dimension aller Bewegungskurven ist die Anzahl aller Variablen. Bewegungsdaten werden mit Hilfe eines virtuellen Skeletts abgespielt, um eine unmittelbare oder spätere Analyse und Wiedergabe durchzuführen. Um eine vollständige Bewegung eines Skeletts aufzunehmen, ist es erforderlich, mindestens 40—50 Marker mit einer Abtastrate von 120 bis 240 Hz zu benutzen (Liu2006a). Die Daten können effizient gespeichert werden, weil die menschlichen Bewegungen geeignete Eigenschaften zur Kompression (z. B. räumliche und zeitliche Kohärenz) aufweisen. Die Kompressionsverfahren können entweder bei der Übertragung der Daten oder beim Zugriff auf Datenbanken angewendet werden. Kompressionstechniken werden auch bei der Bewegungsanalyse, Bewegungsklassifikation und beim Retrieval von Bewegungen angewendet (Faloutsos, 2007), um räumliche und zeitliche Strukturen zu extrahieren. In diesem Beitrag werden zunächst die für Datenkompression relevanten Eigenschaften der menschlichen Bewegung gezeigt. Dann wird eine Taxonomie der aktuellen Kompressionstechniken von menschlichen Bewegungen gegeben und danach werden Schlussfolgerungen besprochen.

3. Eigenschaften der menschlichen Bewegungen:

Die menschlichen Daten weisen Eigenschaften auf, die für den Zweck der Kompression nützlich sind.

- Die Positionen und Orientierungen der menschlichen Körperteile als Funktionen der Zeit sind zeitkontinuierlich und differenzierbar.
- Menschliche Bewegungen weisen zeitliche Kohärenz auf, die sich als Folge der starken Korrelationen zwischen zeitlich benachbarten Aufnahmepunkten manifestiert (Liu, 2006a). Außerdem weisen menschliche Bewegungen räumliche Korrelationen auf, weil sie von Natur aus koordiniert sind. Zum Beispiel kann der Körper beim Gehen balancieren, wenn Arme und Beine sich fast periodisch bewegen (Uehara, 2002).
- Eine Sequenz menschlicher Bewegungen ist die Konkatenation von Aktionen (z. B. Gehen, Springen), die durch Key-frames an den schärfsten Änderungen der Bewegung getrennt werden können.
- Bewegungskurven von Markern, die an Körperteilen angebracht sind, die mit der Umwelt in Kontakt kommen, enthalten hohen Frequenzen, die erhalten werden müssen, um die Natürlichkeit der Bewegung zu gewährleisten (Arikan, 2006).

4. Taxonomie der Kompression von menschlichen Bewegungen

Kompressionsverfahren von menschlichen Bewegungen können in drei Kategorien klassifiziert werden: Raum-, Zeit- und Semantik-basierte Kompression.

4.1. Räumliche Kompressionsverfahren

Räumliche Kompressionsverfahren basieren auf statistischen Analysemethoden und reduzieren die Dimension der Daten durch die Projektion der Datenpunkte in einen niederdimensionalen Unterraum, der einen möglichst großen Anteil der Information enthalten soll. Die bekannteste Methode davon ist PCA („Principal Component Analysis“). Die Anwendung der PCA auf die ganze Datenmenge wird „globale PCA“ genannt. Mit diesem Verfahren sind in der Regel 10-20 Hauptkomponenten erforderlich, um die Bewegung eines typischen menschlichen Skeletts anzunähern, das aus 40-60 DOF besteht (Beaudoin, 2007). Je unterschiedlicher die Aktionen sind, aus denen die Bewegungsdaten bestehen, desto schwieriger ist eine kompakte Repräsentation der gesamten Datenmenge mit der PCA zu finden. Wenn die PCA auf eine Gruppe ähnlicher Bewegungsdaten angewendet wird, kann diese Gruppe mit einer geringeren Anzahl von Hauptkomponenten dargestellt werden (Sattler, 2005). Die Gruppierung kleiner Bewegungsstücke und die Anwendung der PCA auf diesen Gruppen wird „Cluster PCA“ genannt. Typische Bewegungsdaten können gut mit 3-5 Hauptkomponenten angenähert werden (Beaudoin, 2007).

Andere Raumtransformationen suchen nach einer guten Darstellungsmöglichkeit, die nicht unbedingt für eine hohe Kompressionsrate geeignet sind, sondern sich einer besseren Interpretation der Information anbieten. Zum Beispiel kann man durch ICA („Independent Component Analysis“) (Cao, 2003) unabhängige Komponenten zur Beschreibung einer Bewegung finden, die z.B. zwischen bewusst antizipierten und emotional bedingten Anteilen differenzieren.

4.2. Zeitliche Kompressionsverfahren

Zeitliche Kompressionsverfahren segmentieren zunächst einzelne oder ganze Gruppen von Bewegungskurven der Freiheitsgrade. Die Segmentierung wird entweder äquidistant oder durch einen adaptiven Algorithmus durchgeführt z. B. durch Position and Velocity Analysis (Gibet, 2008) oder Probabilistic PCA (Barbivc, 2004).

Zeitliche Kompressionsverfahren können in die folgenden Subkategorien unterteilt werden:

1. Kompressionsverfahren mit Transformation durch Wavelets oder DCT basieren auf der Auswahl von Komponenten, die einen großen Anteil der Energie enthalten. Die Koeffizientenauswahl wird anhand der Länge der Knochen (Beaudoin, 2007) oder einer Rekursion durch dynamische Programmierung (Lee, 2008) getroffen.

2. Kompression durch Abtastung an „Key-frames“ betrachtet die Bewegungsdaten als ein mehrdimensionales Signal. Sie rekonstruiert eine Datensequenz durch Interpolation zwischen zwei Keyframes, die eine kontinuierliche glatte Bewegungskurve erzeugt und deren Natürlichkeit aber von der Interpolationsmethode abhängt (Arikan, 2006). Je größer der zeitliche Abstand zwischen den Datenpunkten ist, desto schlechter ist die Qualität der durch Interpolation rekonstruierten Bewegung.

3. Kompression durch Kurvenbeschreibung nutzt die Glattheit der Bewegungskurven aus. Die Kurvenbeschreibung entspricht einer Kurvenvereinfachung, die nach einem vorgegebenen Fehler eine andere angenäherte Kurve berechnet, die parametrisch modelliert ist. In Bewegungskompression wurden oft Bezierkurven angewendet (Lin, 2007).

4. Kompression mit Hilfe eines Wörterbuches verwendet eine Liste von kurzen Bewegungskurven, die dynamisch aufgebaut wird (Gu, 2009). Die Bewegungsdaten werden eingescannt, und die Kurve aus der Liste, die am besten zum aktuellen Anfangsstück der Bewegung passt, wird gesucht und als Index kodiert. Dies führt zu einer Kompression der Daten. Je größer das Wörterbuch ist, desto kleiner ist der Rekonstruktionsfehler.

5. Kompression durch prädiktive Codierung prognostiziert aktuelle Signalwerte aus einer Kombination der vorangegangenen Werte, damit nur Differenzen zwischen den vorhergesagten Werten und den aktuellen Bewegungsdaten übertragen werden müssen. Je besser die Prediktion ist, desto kleiner ist die Differenz und desto geringer ist die Anzahl der benötigten Bits, um sie zu kodieren.

Prädiktion für menschliche Bewegung wurde im Ortsraum in MPEG BBA (Preda, 2007), im PCA-Raum (Karni, 2001) und mit Hilfe von „Inverse Kinematics“ (Li, 2008) angewendet.

4.3. *Semantische und „Hidden Variables“-basierte Kompressionsverfahren*

Semantische Kompression benutzt Vorwissen über bestimmten Klassen von Bewegungen, um eine bessere Kompressionsrate zu erreichen. Die Semantik lässt sich nicht direkt von menschlichen Bewegungsdaten ablesen, weil sie eine mehrdimensionale Korrelation aus Gelenkpositionen und Zeitreihen aufweisen (Sattler, 2005). Semantische Kompression wird in drei Schritten durchgeführt: (1) Identifikation von primitiven Abschnitten im Datensatz (Primitive Motion (Uehara, 2002)). (2) Klassifikation von primitiven Abschnitten in semantische Gruppen (Motion Motif (Beaudoin, 2008)). (3) Identifikation von Unterschieden zwischen ähnlichen primitiven Abschnitten in jeder semantischen Gruppe.

Zum Beispiel konzentriert sich Gehen vor allem auf die Beine und Arme, die sich in verschiedenen Sequenzen von primitiven Abschnitten bewegen (z.B. ABCD oder XYZ), um das Gleichgewicht zu halten. Diese Sequenzen wiederholen sich nicht genau in einem bestimmten Zeitraum (z.B. ABCD-XYZ-ABCE- XYZ-XYZ-ABCD). Kleine Unterschiede können zwischen zwei Wiederholungen auftreten (ABCD und ABCE) aber die Semantik wird nicht beeinträchtigt. Bei verlustbehafteter Kompression werden semantisch gleiche Symbole mit leichten Unterschieden zusammengefasst und nicht unterschiedlich gespeichert. Die zusammengefassten Sequenzen können durch die Eliminierung der semantisch irrelevanten Unterschiede (Speicherung von ABCD statt ABCE) und durch die Anwendung von Zeit- und Raum-basierten Methoden komprimiert werden. Die Identifikation von semantischen Gruppen basiert u. a. auf künstlicher Intelligenz (Wu, 2009), multimodalen Korrelationen (Mukai, 2007) und Motion Motif (Beaudoin, 2008).

Kompression mit Hidden Variables“ kann die Dynamik bzw. Veränderung mit der Zeit der Variablen durch nichtdeterministische Zustandsmaschinen beschreiben. Ein Zufallsprozess wird durch Zustände, Übergangswahrscheinlichkeiten (zwischen Zuständen) und Emissionswahrscheinlichkeiten (beobachtbare Ausgangssymbole) modelliert. Die vorigen Zustände, die zur Bestimmung des aktuellen Zustands benötigt werden, heißen Kontext. Für jeden Zustandsübergang bewegt sich das Modell aus dem aktuellen Kontext zu einem neuen Zustand, der mit dem aktuellen Kontext einen neuen Kontext bildet. Zur Kompression werden die Übergangswahrscheinlichkeiten und die Emissionswahrscheinlichkeiten als Kompressionsdaten statt der Zufallsvariablen gespeichert.

5. **Schlussfolgerungen**

PCA basierte Verfahren betrachten die Korrelation zwischen Freiheitsgraden der Bewegungsdaten, aber eine Periodizität der Bewegung wird nicht ausgenutzt. Sie sind aufwändig, insbesondere durch die Berechnung der Hauptkomponenten.

Diese Berechnungen müssen allerdings nur einmal und können Offline durchgeführt werden. Dagegen betrachten die zeitlichen Kompressionsverfahren die Periodizität jedes eindimensionalen Freiheitsgrades, aber die Analyse der Bewegung als ein mehrdimensionales System wird nicht durchgeführt. Sie ermöglichen jedoch den direkten Zugriff auf ein einzelnes Signal ohne Dekompression von den anderen. Semantik-basierte Kompression (Beaudoin, 2008) und mehrdimensionale Kompression (Li, 2009) betrachten sowohl die räumliche als auch die zeitliche Kohärenz aber die Komplexität der Datenbearbeitung macht solche Kompressionsverfahren schwierig anzuwenden. Es können Kompressionsraten ohne wahrnehmbare Artefakte bis zu 35:1 durch die Kombination von Zeit- und Raum-basierten Kompressionsverfahren (Arikan, 2006; Beaudoin, 2007), und bis zu 49:1 durch die Kombination von Semantik- und Raum-basierten Kompressionsverfahren (Beaudoin, 2007b) erreicht werden.

6. Literaturverzeichnis

- Arikan, O. (2006). *Compression of motion capture databases*. New York, USA: ACM.
- Barbivc, J. et al. (2004). *Segmenting motion capture data into distinct behaviors*. Waterloo: Canadian Human-Computer Communications Society.
- Beaudoin, P. et al. (2008). *Motion-motif graphs*. Aire-la-Ville: Eurographics Association.
- Beaudoin, P., Poulin, P. & van_de_Panne, M. (2007). *Adapting wavelet compression to human motion capture clips*. New York, NY: ACM.
- Beaudoin, P. (2007b). *Compression de donnees d'animation acquises par capture de mouvements*. Montréal: Département d'Informatique et Recherche Opérationnelle, Université de Montréal.
- Cao, Y., Faloutsos, P. & Pighin, F. (2003). *Unsupervised learning for speech motion editing*. Aire-la-Ville: Eurographics Association.
- Faloutsos, C., Hodgins, J. & Pollard, N. (2007). *Database techniques with motion capture*. New York: ACM.
- Gibet, S. & Marteau, P. F. (2008). *Analysis of Human Motion, Based on the Reduction of Multidimensional Captured Data*. Berlin, Germany: Springer.
- Gu, Q., Peng, J. & Deng, Z. (2009). *Compression of Human Motion Capture Data Using Motion Pattern Indexing*. Oxford: Blackwell.
- Karni, Z. & Gotsman, C. (2001). *3D mesh compression using fixed spectral bases*. Toronto: Canadian Information Processing Society.
- Lee, C.H. & Lasenby, J. (2008). *An Efficient Wavelet-Based Framework for Articulated Human Motion Compression*. Las Vegas: Springer.
- Li, L., et al. (2009). *DynaMMo: mining and summarization of coevolving sequences with missing values*. New York: ACM.
- Li, S., Okuda, M. & Takahashi, S.I. (2008). *Compression of human motion animation using the reduction of interjoint correlation*. New York: Hindawi Publishing Corp.
- Lin, Y. & McCool, M. (2007). *Nonuniform Segment-Based Compression of Motion Capture Data*. Berlin: Springer.
- Liu, G. & McMillan, L. (2006). *Segment-based human motion compression*. Aire-la-Ville: Eurographics.
- Mukai, T. & Kuriyama, S. (2007). *Multilinear Motion Synthesis with Level-of-Detail Controls*. Washington: IEEE Computer Society.
- Preda, M., et al. (2007). *Optimized MPEG-4 animation encoder for motion capture data*. New York: ACM.
- Ren, L., et al. (2005). *A Data-Driven Approach to Quantifying Natural Human Motion*. New York: ACM.
- Sattler, M., Sarlette, R. & Klein, R. (2005). *Simple and efficient compression of animation sequences*. Los Angeles: ACM.
- Uehara, K. & Shimada, M. (2002). *Extraction of Primitive Motion and Discovery of Association Rules from Human Motion Data*. London: Springer.
- Wu, S., Wang, Z. & Xia, S. (2009). *Indexing and retrieval of human motion data by a hierarchical tree*. New York: ACM.