

Web3Dにおけるヒューマノイドキャラクターのマルチストリーム転送に基づいたインタラクティブ混合リアルタイム制御手法の提案

5E-7-4

望月 義幸, 中 俊弥, 樋尻 利紀, 西谷 和博, 浅原 重夫

松下電器産業株式会社 マルチメディア開発センター

〒571-8501 大阪府門真市大字門真 1006

本稿では、インターネット上で3次元CGを扱うためのモデリング言語であるVRMLを代表とするWeb3Dにおいて、ヒューマノイドキャラクターをコンテンツ閲覧者がインタラクティブにリアルタイム制御できる手法として、動作データのマルチストリーム転送による自動制御と独自インバースキネマティクスに基づいた手動制御を組み合わせた混合リアルタイム制御手法を提案する。また、本手法を用いたサーバー・クライアントシステムの構成例とそのシステム上でのコンテンツ応用例についても報告する。本手法によれば、動作データのマルチストリーム転送による自動制御の大きな特徴である、「狭帯域のネットワークでもダウンロードの待ち時間が掛からない」という利点を生かした上で、閲覧者のインタラクティブ性を増大させることが可能となり、新たなコンテンツ制作の上でも非常に有効である。

Hybrid Realtime-Controlling Method Based on Multi-Stream Transfer for a Humanoid Character on Web3D

Yoshiyuki MOCHIZUKI, Toshiya NAKA, Toshiki HIJIRI,

Kazuhiro NISHITANI, and Shigeo ASAHARA

Multimedia Development Center, Matsushita Electric Industrial Co., Ltd.

1006, Kadoma, Kadoma-shi, Osaka, 571-8501, Japan

In the present paper, we propose a hybrid realtime-controlling method for a humanoid character on Web3D including VRML that is combined the automatic control by the multi-stream transfer for motion data with the manual control based on the original inverse kinematics called "Potential-based Inverse kinematics" by a user browsing the contents. We also report the server/ client system implemented the proposal method and indicate an example of the contents applied the proposal method on the system. The proposal method can increase the interactivity between a contents and a user without losing the feature of the multi-stream transfer that it is possible to send/ receive motion data in realtime on a network with narrow band width. There is a possibility that the brand-new type of contents can be generated by the proposal method.

1. はじめに

近年、動画や音声のような大容量のデータをインターネット配信する場合、ストリーミング技術が広く利用されるようになった。これは、コンテンツを閲覧する際に、全てのデータをダウンロードする必要が無い場合、インターネット放送における基幹技術の一つとなっている。一方、3次元CG (Computer Graphics) は、動画に比べ、データがコンパクトで閲覧者とのインタラクティブ性を実現できるという点で勝っており、次世代インターネットの主役になり得るデータ形式である。

VRML (Virtual Reality Modeling Language) [1][2] を代表とする、インターネット上で3次元CGを扱うためのモデリング言語は、最近では、Microsoft社の提唱するDirectXやSUN Micro Systems社のJava3Dなどを含めてWeb3Dと総称される。Web3Dにおいて、人間のような骨格構造を持つヒューマノイドキャラクターによる3次元CGアニメーション(ヒューマノイドアニメーション) [3]はコンテンツ内容を豊かにするという意味で非常に重要であり、その骨格構造の定義仕様は、VRML Humanoid Animation Version 1.0[4]として既に標準化されている。我々は当初からこの標準化に関わり、ファイルベースのデータ転送によるヒューマノイドアニメーションを実現するためのサーバー・クライアントシステムや実験的なコンテンツを開発してきた[5][6][7]。しかしながら、コンテンツ規模が大きくなるにつれ、ファイルベースでのデータ転送ではダウンロードの待ち時間が長大になるという問題点が顕著になってきた。この問題点を解決するために、ヒューマノイドキャラクターの動作データや音声データを、マルチストリームとして転送する手法[8][9]を提案した。これにより、一般に用いられるパソコンなどのコンピュータ同士を電話回線などの狭帯域のネットワークで接続するサーバー・クライアントシステムにおいても、ほとんど待ち時間なしにヒューマノイドアニメーションを閲覧することが可能となった。

以上の、これまで我々が開発してきたシステムやコンテンツでは、ヒューマノイドキャラクターの行動

は、予め規定された定型的な動作、またはその組み合わせによる動作によって自動的に生成される。シナリオ通りのヒューマノイドキャラクターの行動を見せるようなコンテンツではそれでも十分ではあるが、閲覧者とのインタラクティブ性が限定されるという感は否めない。もし、ヒューマノイドキャラクターの動作を閲覧者がリアルタイムに手動制御できるならば、閲覧者とのインタラクティブ性が増大することは明白である。しかし、ヒューマノイドキャラクターの移動を含めた身体部位全ての動作をリアルタイムに手動制御することは、人間等の生物が非常に多数でかつ複雑な筋神経系で制御していることを観ても分かるように、少数の入力系統しかないパソコンなどで用いられる入力機器では実現が困難である。そこで本論文では、ヒューマノイドキャラクターの身体部位の中、閲覧者がリアルタイムに手動制御可能な部位を予め定めておき、マルチストリーム転送された動作データによる自動制御と閲覧者入力のリアルタイム手動制御によって、身体部位全体の動作制御を行う、混合リアルタイム制御手法の提案を行うものである。ヒューマノイドキャラクターの身体部位の中、最も重要で閲覧者とのインタラクティブ性が効果的に増大するのは上肢や下肢と考えられることから、上肢や下肢の動作をリアルタイム手動制御可能とする場合について報告する。また、リアルタイム手動制御を実現する方法として、操り人形の動きを再現できる独自のインバースキネマティクス (Potential-Based Inverse Kinematics) についても説明する。なお、本論文で述べる応用システム例では、閲覧者は一般に用いられるパソコン用の入力機器によってリアルタイム手動制御を実行できる。

2. ヒューマノイドアニメーション

ヒューマノイドキャラクターの骨格構造は、人間の骨格構造に対応する関節 (Joint Node) とセグメント (Segment Node) の階層的な接続関係によって定義される。図1は、本論文の混合リアルタイム制御を実行するヒューマノイドキャラクターの骨格構造と関節名を付記したもので、VRML Humanoid

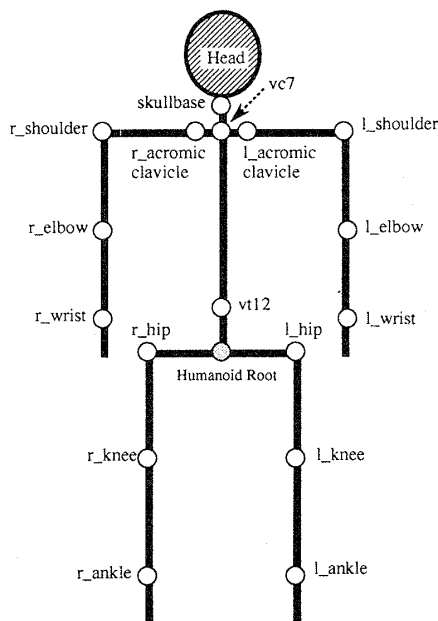


図1 ヒューマノイドキャラクタの骨格構造

Animation Version 1.0 の定義仕様の部分構造となっている。

このように定義した骨格構造に対して、ヒューマノイドキャラクタの身体全体の動作は、ルート (Humanoid Root) の位置ベクトルによって身体全体の位置を制御し、ルートの方向ベクトルとその周りの回転角度によって身体の姿勢を制御する。各関節に対しては、最大で並進 3 自由度と回転 3 自由度の計 6 自由度が与えられている。特に、関節角の制御については軸ベクトルとその周りの回転角度によって行われる。ヒューマノイドキャラクタの各時刻毎の動作データは以上によって構成され、これらの時系列データとして表現できる。

3. マルチストリーム転送

第2章で述べたように、ヒューマノイドキャラクタの動作データは位置ベクトルや回転軸ベクトル、回転角度によって構成される。これらの時系列データを動作ストリームデータとして定義し、音声データ等と共に動作ストリームデータをクライアント・サーバー間で送受信する。ここでは、特に動作ストリームデータの送受信について記述する。

サーバーでは、ヒューマノイドキャラクタの動作

(モーションキャプチャ等で得られたデータ) を解析し、(1) ルートの位置ベクトルの情報、(2) ルートの方向ベクトルと回転角度の情報、(3) ルート以外の関節での回転軸を表す角度ベクトルとその周りの回転角度の情報に分類する。この分類は、(1) と (2) については、身体全体の動きを規定するので精度が必要なため圧縮よりも精度を重視し、(3) については局所的な動作なので精度よりもデータの圧縮を優先するという考え方に基づいている。動作データを構成するこれらのデータをチャンネルと呼ぶことにする。図1の場合、ルートの位置ベクトルを表すチャンネル、ルートの方向ベクトルと回転角度を表すチャンネル、17 関節分の回転軸ベクトルと回転角度データを表す 17 チャンネルがあり、合計 19 チャンネルを持つことになる。また各チャンネルデータの詳細については、(1) は、 x, y, z の 3 成分を持つ 3 次元ベクトル、(2) と (3) は、回転軸となる x, y, z の 3 成分を持つ 3 次元ベクトルと回転角度を表す 1 成分を持った、合計 4 成分を持つデータである。なお、これらの動作データは我々の開発した独自手法により、1/10 程度に圧縮され、また特徴として、データの回線速度や、クライアントのコンピュータの性能により、圧縮率の異なる動作データを適応的に選択するということが可能である [7][8]。

各チャンネルデータは適当に定められた時間長で 1 ブロックとするような、時間軸方向での“ブロック化”を行う。たとえば、1 秒単位でブロック化する場合、1 秒分の動作データが 1 ブロックとなり、その 1 ブロックが 1 Data Packet となる。このブロック化された Data Packet には、ヘッダ情報として、ある基準時からの時刻を表すタイムスタンプを付加する。このタイムスタンプをサーバーとクライアントの双方で管理することによって、サーバーとクライアント間の同期が可能となり、リアルタイムでのヒューマノイドキャラクタの再生を行なえる。なお、動作データの他に、カメラ位置や視線方向等の時系列情報やライティングの位置や輝度等の時系列情報も同様に、Data Packet として定義できる。

図2はマルチストリームデータ転送の仕組みの概

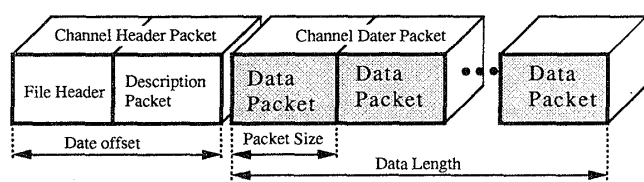


図2 マルチストリームデータ転送の仕組み

念図である。最初に、Channel Header Packet がサーバからクライアントに送られ、どのようなチャンネルが送られてくるかクライアントに知らせる。次に、Description Packet の記述に対応して、Data Packet がクライアントに送られるが、その際には、ネットワークのデータ転送性能やクライアント・マシンの性能を考慮して、Data Packet のサイズを決めておく必要がある。図3はDescription Packet とData Packet の関係を示したものである。ヒューマノイドキャラクターの動作データや、カメラ情報、ライティング情報等のチャンネルの"Type"が最大チャンネル数(TC)分、Description Packet によって規定され、各チャンネルのData Packet は2次元の配列 $CnData[m][n]$ としての内部構造で、 m はチャンネル番号に対応しており、 n は時間方向のデータ数を表す。但し、 $m \leq TC$ でなければならない。

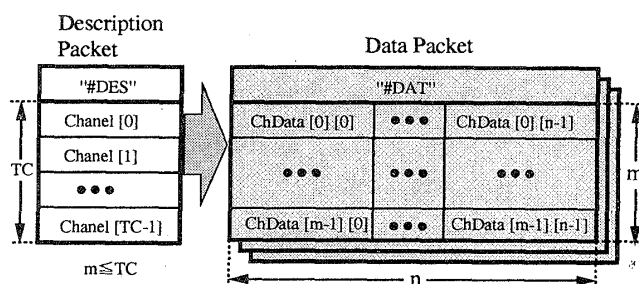


図3 Description Packet とData Packet の関係

4. 混合リアルタイム制御

4.1 手動制御方法

人間のような多関節の物体の動作制御を、少数入力系統の入力機器によって手動で行っている典型的な例として「操り人形 (マリオネット)」があるが、我々はこの操り人形の動作制御をシミュレートすることでヒューマノイドキャラクターのリアルタイム手

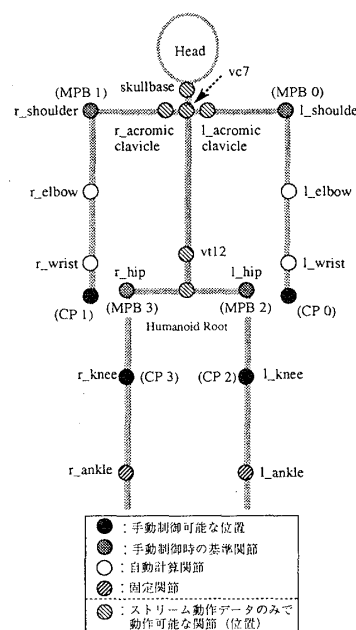


図4 各関節の制御方法

動制御を実現しようと考えた。操り人形の場合、糸の接合部の位置を制御入力とし、ある基準位置とその糸の接合部との中間部の状態は重力作用によって決定される。これは一種のインバースキネマティクス計算であるが、我々はこの重力作用によって中間部位の状態を計算する方法を、特に"Potential-Based Inverse Kinematics"と呼んでいる。

第1章で述べたように、手動で制御を行う部位は両上肢、下肢で、それ以外の部位についてはマルチストリーム転送された動作データによって制御する。両上肢に対しては、先端部 (手先) の空間位置を入力とし、肩関節位置を固定としたときの上肢を構成する3セグメント (上腕部、前腕部、手部)、3ジョイント (肩関節、肘関節、手関節) の状態を自動計算することによって制御を行う。また、両下肢については、膝関節の空間状態 (角度) を入力として、下肢を構成する3セグメント (大腿部、脛部、足部)、2ジョイント (股関節、膝関節) の状態を自動計算し、残りの1ジョイント (足首関節) は固定とすることによって制御を行う。

図4は図1の骨格構造に対して前記の制御部位を示したもので、 $CP0 \sim CP3$ は空間位置の入力を行う関節位置で、 $MPB0 \sim MPB3$ は関節の状態を自動

計算する際の基準位置であり，その間の中間部位については自動計算される。

4. 2 関節状態の自動計算

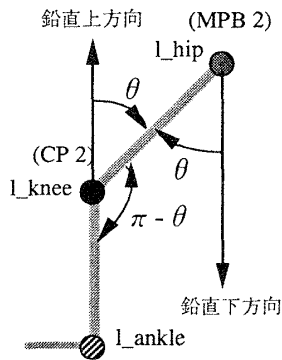


図5 左下肢に対する自動計算

前節の説明からも分かるように，Potential-Based Inverse Kinematics での計算では，重力作用のみを考慮しているのので，入力関節位置と基準関節位置，および中間の関節位置が常に同一平面上に拘束されることに注意しなければならない。下肢の中間状態の自動計算については，膝関節の位置を入力としてもよいが，基本的には大腿部の1セグメントを規定できればよいので入力としては大腿部の鉛直下方向に対する角度 (θ) を入力とする。この際，脛部以下は重力作用によって鉛直方向にぶら下がることに

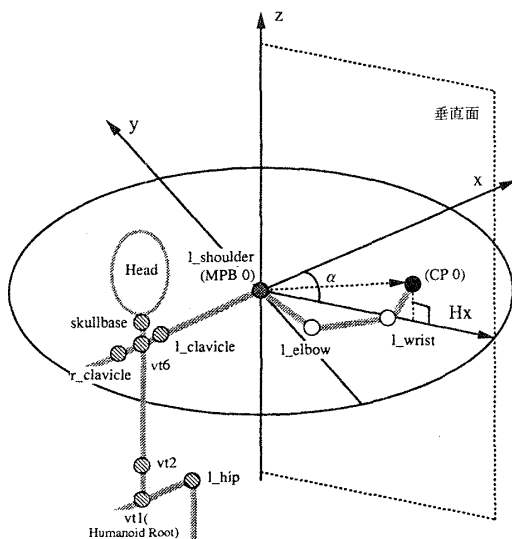


図6 左上肢を拘束する垂直面

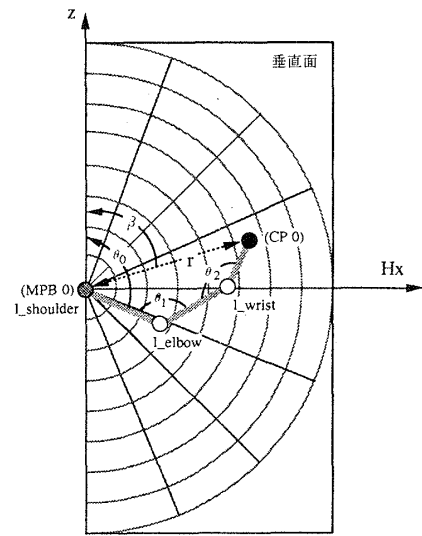


図7 垂直面での左上肢の状態

なるので，脛部は鉛直方向と平行になり，鉛直上方向のベクトルと θ の角度をなす。つまり，膝関節は $\pi - \theta$ となり，中間関節の状態が計算できる。図5に左下肢に対するこの様子を示す。

上肢の場合は下肢に比べ3セグメントの状態を計算しなければならないため複雑で，リアルタイムでの計算は困難である。そこで，代表的な状態を予め計算してテーブル化しておき，入力に対してテーブル値を算出し，それらの値を補間することで算出する。図6に示すように，左上肢の手先の位置は肩関節に規定された局所座標系 xyz の位置ベクトルとして表わせる。この位置ベクトルとその先端から xy 平面に下ろした垂線の足を含む垂直面に，全ての関節位置は含まれる。この垂直面と x 軸のなす角を α とすると，この垂直面は α によって規定される。この垂直面と xy 平面の交線を Hx とし，x 軸を z 軸周りに回転して Hx と重なったときの軸方向を Hx の軸方向とすると，この垂直面上に図7に示す Hx-z 座標系を定義できる。手先の位置は図7の β と r によって規定されるので，ある1平面上で代表的な β と r (例えば図7の極座標系の格子点) に関して図7の $\theta_0 \sim \theta_2$ をテーブル値としてもつようなテーブルを用意しておけばよい。

上腕の長さを L_0 ，質量を M_0 の均質な棒状物体，前腕の長さを L_1 ，質量を M_1 の均質な棒状物体，手

部の長さを L_2 、質量を M_2 均質な棒状物体で近似し、Hx-z 座表系での手関節(L_wrist)の z 座標を Z_w 、手先(CP0)の z 座標を Z_{cp} とすると、重力ポテンシャル E は、

$$E = \{M_0 L_0 \sin \theta_0 + M_1 (L_0 \sin \theta_0 + Z_w) + M_2 \cdot (Z_{cp} + Z_w)\} / 2$$

となるので、各代表点位置にある手先に対して E が最小となるときの $\theta_0 \sim \theta_2$ を求めてテーブル化する。この算出は、例えば勾配法や準 Newton 法などで求めることができる。

4. 3 上肢中間状態の自動計算結果の評価

テーブル値に対する補間として4点による双線形補間を用いた場合、上肢長に対する半円平面を r (長さ方向) については 100 分割、 β (角度) については 18 分割したテーブルを用いて、テーブル検索したデータを補間して関節角を算出し、その結果から算出した手先位置と入力した手先位置を比較することで、計算精度を評価した、評価結果は、平均で上

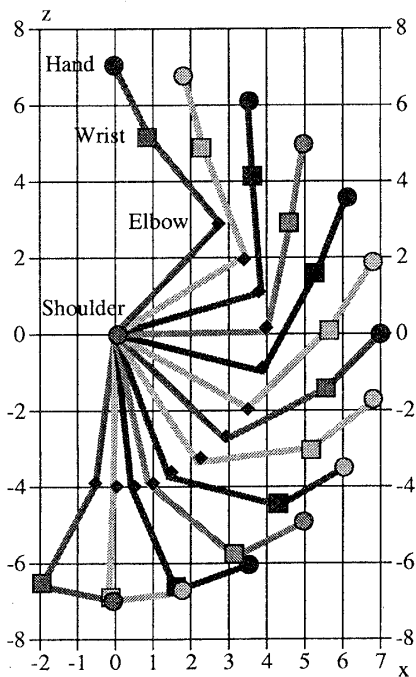


図8 手先の円運動に対する上肢の状態

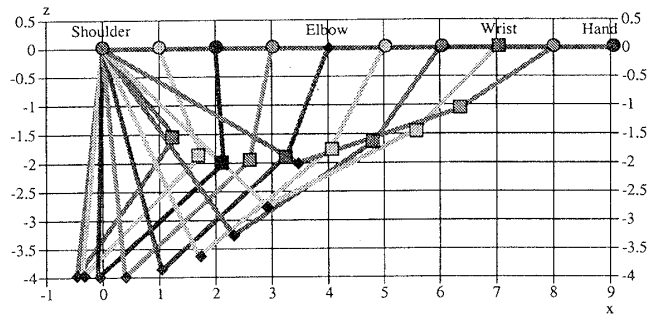


図9 手先の水平移動に対する上肢の状態

肢長の 0.1% で、最大でも 5% 以内に納っていた。なお、テーブルサイズはテーブル値を単精度の浮動小数点で保持した場合、8kB 以内である。

図 8 は手先を回転運動したときの上肢の状態の計算結果を表し、図 9 は手先を水平方向に移動したときの上肢の状態の計算結果である。

5. システム適用例

5. 1 システム構成

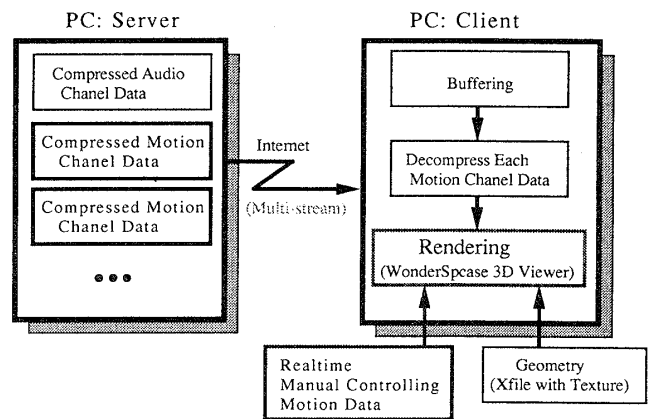


図10 混合リアルタイム制御システムの構成

図 10 に、本システムの全体構成例を示す。本システムは RealNetworks 社のサーバー・クライアントシステムである RealSystem G2 のプラグインとして実現されている。本来、同システムは映像と音声ストリームの転送用のシステムであるが、カスタマイズ機能をもっており、このカスタマイズ機能を利用して動作データのマルチストリーム転送に拡張し

た。

図 10 に示すように、サーバーからは、まず音声データと動作データが圧縮された形でクライアント側にマルチストリーム転送される。転送された音声データについては、RealSystem G2 本来の機能によって再生される。動作データは一旦バッファリングされた後、伸長される。伸張された動作データは、DirectX ベースの独自ブラウザである WonderSpace 3D Viewer に入力される。この動作データによってヒューマノイドキャラクターは自動制御される。一方、手動制御については、Microsoft 社製の純正コントローラを入力機器としている。このコントローラの入力部に対して、上肢については、手先の位置制御のために並進移動の 6 自由度の入力を割り付け、下肢については鉛直方向に対する角度の入力を割り付けている。従って、14 系統の入力を行うことになる。このコントローラによって入力されたデータがマルチストリーム転送された動作データ同様、WonderSpace 3D Viewer に入力され、第 4 章で説明した手法で両上下肢の状態が算出される。マルチストリーム転送された動作データによる自動制御と手動制御の切替は、手動制御入力が無いときは自動制御されるというルールで行っている。以上のデータと、予めクライアント側に送られている、または保持している 3 次元 CG データを WonderSpace 3D Viewer でレンダリングを行い閲覧する。

5. 2 コンテンツ応用例

図 11, 12 は第 5.1 節で述べたシステムに対するコンテンツの応用例を示したものである。このコンテンツでは、動作データとしてダンス動作が、その動作に同期した音楽と共にマルチストリーム転送され、その音楽に合わせて、両上下肢を手動制御するというものである。図 11 は両上肢を手動制御したもので、図 12 は両上下肢を手動制御したときのそれぞれの 1 フレームを示したものである。ヒューマノイドキャラクターの身体全体の位置や姿勢、関節動作を、音楽に合わせてダンスさせるということを手動で制御することは非常に困難である。しかし、両上

下肢を手動制御するだけならば、基本的にはマルチストリーム転送されたダンス動作によって自動制御された状態で、両上下肢を局所的に変化させることから、ダンス動作としての不自然さはほとんど感じられない。また、入力操作自体は入力機器の入力系統が 14 あるため、これらの入力系統と両上下肢の動きの対応関係に慣れる必要はあるが、一旦慣れればそれ程の困難はない。



図 11 ダンス動作（両上肢のみの制御の場合）

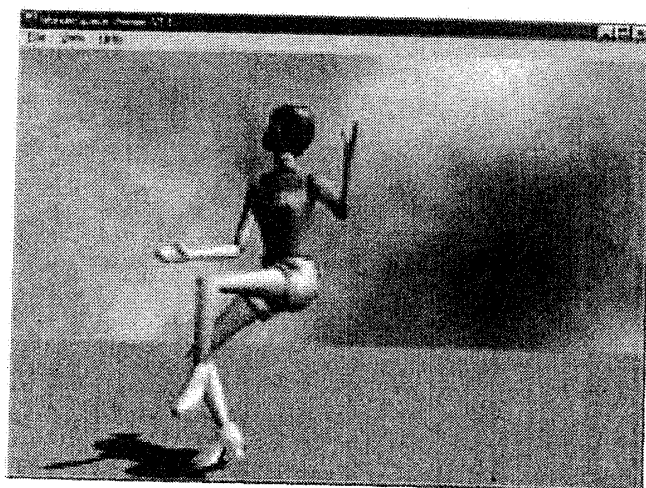


図 12 ダンス動作（両上下肢の制御の場合）

6. まとめ

本稿では、Web3D におけるヒューマノイドキャラクターに対するインタラクティブなリアルタイム動作制御として、動作データのマルチストリーム転送

による自動制御と独自インバースキネマティックス手法に基づいた、コンテンツ閲覧者による手動制御を組み合わせた、混合リアルタイム制御手法について述べた。また、本手法を用いたシステム構成例とそのシステム上でのコンテンツ応用例についても報告を行った。本手法によって、動作データのマルチストリーム転送による自動制御の大きな特徴である、「ダウンロードの待ち時間が掛からない」ということに加え、閲覧者のインタラクティブ性を増大させることが実現できた。本来、ヒューマノイドキャラクターの身体全ての動作制御を手動で行うことが簡単ならば、閲覧者のインタラクティブ性を最も増大させることになる。しかし、通常の人がこのような手動制御を行うためには、特殊な入力装置が必要である（例えば、リアルタイム・モーションキャプチャ・システム）。社会への普及を考えた場合、そのような特殊な入力装置によるよりは、動作自体に多少の制限があっても一般的な入力装置で実現できる手法の方が有効であると考えられる。そのような意味で、本手法の有効性は自明であり、本手法はゲームを含む、新たなコンテンツ制作への活路を与えるものと考えられる。また、インターネットのみならず、今後急速な普及が予測されるデジタル放送においても、同様な手法の適用が可能である。さらに、本手法ではデータの蓄積が起これないため、記憶容量の小さな装置、例えば携帯端末などへの応用も可能であり、今後予定される第3世代移動体通信などで、新たなサービス形態を生むための基幹技術になり得る。

謝辞：(株) 松下ソフトリサーチの勝田氏、および本研究の機会を与えてくれた松下電器産業(株) マルチメディア開発センターに心から感謝を申し上げます。

参考文献

- [1] VRML97 International Standard ISO/IEC 14772-1 2.0 Specification (<http://www.vrml.org>), 1997.
- [2] Jed Hartman, Josie Wernecke, "The VRML 2.0 Handbook", Addison Wesley Developers Press, 1996.
- [3] Norman I. Badler, Cary B. Phillips, Bonnie Lynn Webber, "Simulating Humans: Computer Graphics animation and Control", Oxford University Press, New York, 1993.
- [4] Bernie Roehl, "Specification for a Standard VRML Humanoid Version 1.0", Humanoid Animation WG, August 1997, (<http://ece.uwaterloo.ca:80/~h-anim/>).
- [5] Toshiya Naka and Yoshiyuki Mochizuki, "WonderSpace: Interactive 3D Animation Browser", SIGGRAPH'97 Visual Proceedings, p111, August 1997.
- [6] 中俊弥, 「インターネット上で CG キャラクタを自由に動かす」, 日経 CG11 月号, pp188-192, 1997.
- [7] 西谷和博, 樋尻利紀, 望月義幸, 中俊弥, 浅原重夫, "WonderSpace: VRML ヒューマノイドアニメーション手法", 第 14 回 NICOGRAPH/MULTIMEDIA 論文コンテスト論文集, pp163-172, 1998.
- [8] 樋尻利紀, 西谷和博, 望月義幸, 中俊弥, 浅原重夫, "VRML ヒューマノイドアニメーション用動作ストリーム符合化の一手法", 情報処理学会グラフィックスと CAD 91-8, pp41-46, 1998.
- [9] Toshiya Naka, Yoshiyuki Mochizuki, Toshiki Hijiri, Tim Cornish, and Shigeo Asahara, "A Compression/Decompression Method for Streaming Based Humanoid Animation", VRML99 Fourth Symposium on the Virtual Reality Modeling Language, pp53-70, acm PRESS, 1999, pp53-70, acm PRESS, 1999.