

3次元CGを用いた通信カラオケの開発

白井 豊 林 克彦 徳永 真志 鈴木 誠一 大久保 晴代

shirai@vrl.mei.co.jp

松下電器産業株式会社 AVC 商品開発研究所

パソコンを中心にして、3DCGアニメーションが民生用途へ広がりを見せている。一方で、コンピュータ音楽の標準規格としてMIDIサウンドが定着しつつある。このような背景から、3DCGとMIDIとの融合が今後加速であろうと予想する。そこで、本書は3DCG通信カラオケをターゲットに、以下の項目について報告する。(1) MIDIサウンドと3DCGアニメーションとの同期方式:MIDIと3DCGとの1拍単位の同期により、同期遅延を軽減する。(2) 3DCGカラオケコンテンツ制作環境:CGデータの部品化により、制作期間を短縮する。(3) 3DCGカラオケコンテンツデータ配信方式:(2)を利用してデータ転送量を限定し、転送時間を短縮する。(4) 弊社のM2システムでのプロトタイプ開発。

Server-Based KARAOKE Player with 3-Dimensional Computer Graphics

Yutaka SHIRAI, Katsuhiko HAYASHI, Masashi TOKUNAGA,
Seiichi SUZUKI and Haruyo OHKUBO

AVC PRODUCTS DEVELOPMENT LABORATORY
MATSUSHITA ELECTRIC INDUSTRIAL CO.,LTD.

By using Personal Computer(PC), applicable usage of 3DCG-animation has been large. On the other hand, MIDI-based sounds is formalized as standard of computer music. We expect that combining MIDI and 3DCG has been one of important topics. In this paper, we report the following issues; (1)*CG-animation synchronized with MIDI-based sounds.* By synchronization with MIDI-based sounds on each tempo, CG-animation can recover from AV inconsistency caused by dynamic change of tempo. (2)*Authoring environments for 3DCG KARAOKE software contents.* By making motions parts that is reusable for multiple contents, the average time of authoring for single 3DCG KARAOKE software content is reduced. (3)*Delivery methodology for 3DCG software contents via networks.* By using advantage of (2), the amount of transferred data via networks is restricted. As a result, transfer time is decreased. (4)*Prototype development of "Server-Based KARAOKE Player" on M2 system produced by our company.*

1、はじめに

1990 年前半まで CG(コンピュータグラフィックス)のアプリケーションは、グラフィックス専用ワークステーション上で実装されていた。用途は、映画、各種マージナル映像、CAD 等の業務用途が大半であった。しかし、近年のパソコンの CPU 処理能力の向上や 3Dグラフィックスアクセラレータの普及により、ゲームやインターネットの VRML (Virtual Reality Modeling Language)サイトなど、民生用途へも広がりをみせている。それに伴い 3DCGアニメーションの研究は、従来のキャラクターアニメーションなどの基本的なもの[1]から、運動力学を利用したもの[2]まで種々に及んでいる。

一方で、コンピュータ音楽の標準規格として MIDI(Music Instrument Digital Interface)が定着しつつある [3][4]。MIDI 音楽から 3DCGアニメーションを生成するツールの開発 [5] など、MIDI と 3DCG との融合に関する研究も行われている。3DCGとMIDIとの融合は、今後ますます加速であろうと予想できる。

そこで、著者らは、映像(3DCG)、音響(MIDIサウンド)及び通信の AVCC 融合を目的として、「3DCG 通信カラオケ」の開発に取り組んだ。本書は、その開発内容について以下の通り報告する。

- ・3DCGアニメーションと MIDIサウンドとの AV 同期方式:現状では、各々独立して再生する CGアニメーションと、MIDI サウンドとを1拍単位で同期させることにより、同期遅延を軽減した。
- ・3DCGカラオケコンテンツの制作環境:CGデータを部品化し、複数カラオケコンテンツで共用することにより、制作期間を短縮した。
- ・3DCGカラオケコンテンツデータの配信方式:転送データを限定することにより、転送時間を短縮した。
- ・弊社の M2システムにおける「3DCG 通信カラオケ」のプロトタイプ開発。

2、3DCGアニメーションと MIDIサウンドとの AV 同期方式

2、1、問題点

現状の通信カラオケの映像と演奏との AV 同期に関する問題点を、以下に示す。

- (1)演奏のテンポに合わせた映像を再生することが難しい。現状のカラオケ演奏は、演奏と同期せずに BGM(BackGround Music)として再生しているに過ぎない。「アップビートな演奏に合わせてダンスを踊る」ような高精度な AV 同期を要するカラオケ映像を再生することは困難である。
- (2)ユーザーからのテンポ変更を要求されると、演奏と映像の同期がずれる可能性がある。例えば、ユーザーが演奏中にテンポを速く変更しても映像は同じ速度で描画するので、以降映像は演奏に対し徐々に遅れ始める。

2、2、音源エンジンと CG エンジンのデータフロー

MIDIデータを再生する音源エンジンと、CGアニメーションを描画するCGエンジンのデータフローを図1に示す。

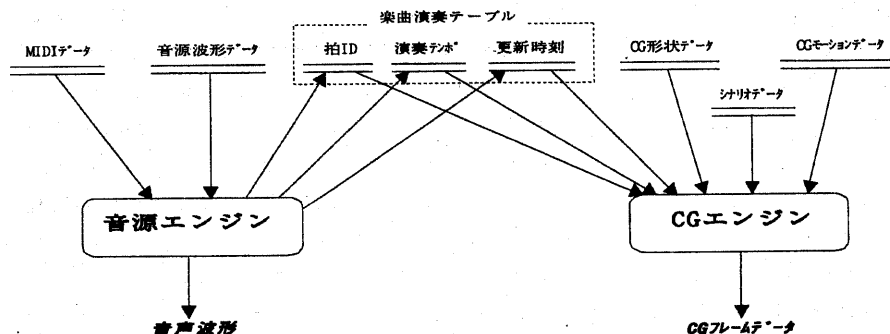


図1 AV同期のデータフロー

・音源エンジン

入力:複数の音源波形データ(楽器音の波形)。MIDIデータ(楽譜データ)。

処理:MIDIデータに基づき音源波形データを合成し、音声波形(音楽)を出力する。現在演奏中の演奏位置と演奏テンポに基づき、1拍単位で楽曲演奏テーブルを更新する。

CGエンジン

入力:CG形状データ(幾何、表面属性)、CGモーションデータ(キールムデータ)、楽曲演奏テーブル、シナリオデータ。

処理:CG形状データの幾何情報を、モーションデータに基づき変換し、CGフレームデータを出力する。

2、2、1、楽曲演奏テーブル:

同期のために、CGエンジン、音源エンジン双方からアクセスされる共有メモリであり、以下のデータから構成する。

- (1)拍ID:楽譜(MIDIデータ)の先頭から4分音符単位で、順々に拍IDを付加する。この拍IDにより、楽曲(楽譜)上の演奏位置を特定することができる。
- (2)演奏テンポ:4分音符の単位時間。MIDI演奏の基準となるパラメータ。曲の途中で、ユーザーからのテンポ変更要求などで動的に変化する。このパラメータに基づき曲のテンポを調節する。
- (3)更新時刻:(1)及び(2)の情報を更新した時の絶対時刻。

2、2、2、シナリオデータ

CGモーションデータと、モーション開始拍IDと終了拍IDとを対応づけたデータである。

2、3、AV同期処理

上述したデータフローに基づきAV同期を行う場合の処理を、図2の拍ID $H_s \sim H_e$ までの演奏中に、モーションデータ $F_s \sim F_e$ までのモーションを再生する場合を例にとり説明する。

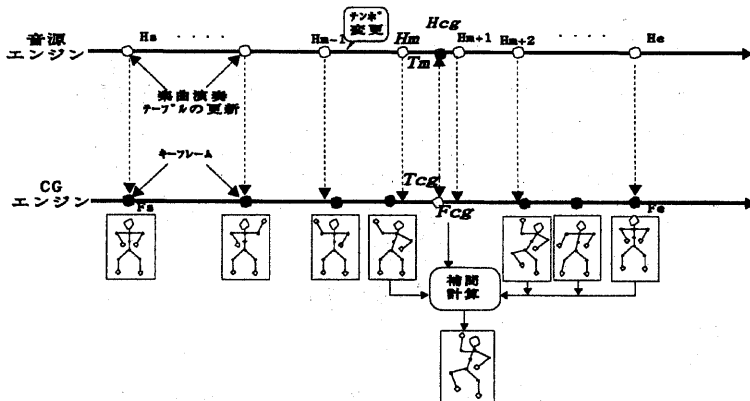


図2 音源エンジンとCGエンジンの同期

(1)音源エンジン:楽曲演奏テーブルの1拍単位で拍ID(H_m)、演奏テンポ(D_m)及び更新時刻(T_m)を更新する。

(2)CGエンジン:以下の処理を、描画フレームマ (F_{cg})が、 F_e になるまで繰り返す。

(a)現在時刻(T_{cg})と楽曲演奏テーブルとから、現在の演奏位置(H_{cg})を計算する。

$$H_{cg} = H_m + (T_{cg} - T_m) / D_m$$

(b)CGエンジン: 描画フレームマ (F_{cg}) を計算する。

$$F_{cg} = F_s + H_{cg} \cdot (F_e - F_s) / (H_e - H_s)$$

(c) F_{cg} に対応した変換行列を、前後のキールム から補間することにより計算する。図2に示すように、その前後のキールムからスプライン補間アルゴリズムにより計算する。

このように1拍(平均約500msec)毎に演奏テンポと拍IDとの同期を行うことにより、演奏に対するCGアニメーションの遅延は、**最大で1拍に抑えることができる。**

3、3DCGカメラコンテンツの制作環境

3、1、現状の3DCGコンテンツ

従来のテレビゲームなどの3DCGコンテンツの制作工程を、以下に示す。

(1)CGデザイナーが、市販のCGモデリングソフトを用いて形状データ、モーションデータ(モーションキャプチャでも作成可)などのCGデータを制作する。

(2)CGプログラマーが、(1)で制作したCGデータを組込んだCGプログラムを開発する。

この工程で、3DCGカラオケコンテンツを制作する場合、1曲単位でコンテンツ制作するため制作期間が長い。これでは、新曲リリースから短期間でカラオケコンテンツをリリースすることができない。

3. 2. モーションピース方式

上記問題点を解決するため、CGデータの部品化による新しい3DCGカラオケコンテンツ制作工程を提案する。モーションデータを、以下の手順で部品化する。

(1)曲全体を同一旋律の繰り返しや休符に着眼することで、2~4小節単位(約8秒~16秒)単位のセグメントに分割する。

(2)同一または類似した旋律(例えば、同じリズム、テンポの部分)に、同一モーションデータを振り付ける。

(3)再生順序が前後するモーションデータを、必ず同一ポジション(ホームポジション)で開始・終了するようにする(このようなモーションデータを、モーションピースと呼ぶ)。

(4)同じような曲調の複数曲で(例えば、同じ作曲家の曲)、モーションピースを共用する。

このように、モーションデータにホームポジションを持たせることにより部品化し、以下のモーションピース方式による3DCGカラオケコンテンツ制作工程を確立した(図3参照)。

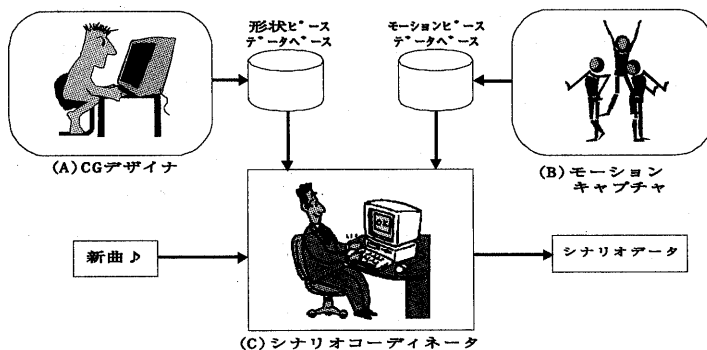


図3 3DCGカラオケコンテンツ制作工程

3. 3. 3DCGカラオケコンテンツ制作工程

(1)CGデザイナーが、CGモデリングソフトを用いて形状データ(これを“形状データピース”と呼ぶ)を制作する。制作したデータを蓄積し、形状ピースデータベースを構築する(図3の(A))。

(2)モーションキャプチャにより、モーションピースを蓄積する。モーションピースデータベースを構築する(図3の(B))。

(3)新曲リリース時に、シナリオコーディネータが(1)の形状ピースデータベース、(2)のモーションピースデータベースから、形状ピース、モーションピースを各々取出す。新曲をセグメントに分割し、各セグメントにモーションピースを振り付ける。それにより、曲の各セグメントとモーションピースとを対応づけた“シナリオデータ”を作成する(図3の(C))。

この工程で、形状ピース、モーションピースがある程度蓄積されると複数曲で共用できる。したがって、コンテンツ制作はシナリオデータ制作だけになり、従来工程に比べて1曲当たりの平均制作期間を短縮することができる。

4. 3DCGカラオケコンテンツデータ配信方式

3DCGカラオケコンテンツを端末側で再生するためには以下のデータを、サーバー側から電話回線(36.6Kbps)を通して、限られた時間内(最低でも30~40秒程度)で転送しなければならない。

(1)CGデータ(形状ピースデータ、モーションピースデータ、シナリオデータ):約2 Mバイト/曲。

(2)MIDIデータ:約 50Kバイト/曲。

しかし、これらのデータを全て、電話回線で転送すると約 8 分程度かかってしまう。そこで、図 4 のような構成にし、モーションピースを利用して以下の配信方式を確立する。

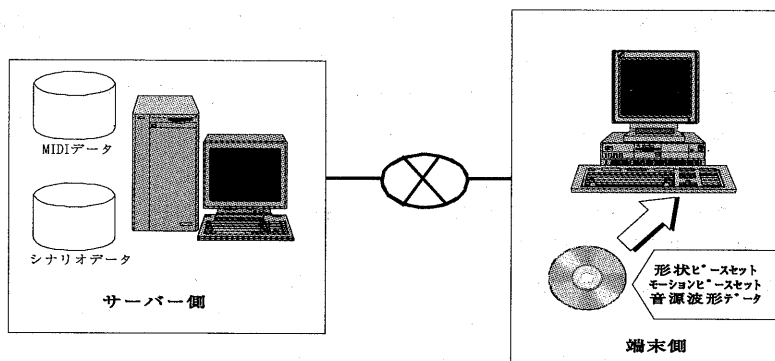


図4 3DCGカラオケコンテンツデータ配信方式

・サーバー側:以下のデータベースを備える。

- (1)MIDIデータベース:各 MIDIデータは曲と対応づけられている。
- (2)シナリオデータベース:各シナリオデータは、曲と対応づけられている。

・端末側:以下の基本データセットを、CD-ROM 等の記録媒体で予め備える。

- (1)形状ピースデータセット:形状データの基本集合。
- (2)モーションピースセット:部品化されたモーションデータの基本集合。
- (3)音源波形データ:複数の楽器音の波形。

この構成で、以下のデータ配信を行う。

- (1)端末:サーバーに対して、演奏曲を指定する。
- (2)サーバー:指定されて曲に対応づけられた MIDI 及びシナリオデータを検索し、端末へ送信する。

この方式を採用すると、データの転送を MIDIデータとシナリオデータに限定できるので、データ転送量を軽減し(約 2MB→60KB)転送時間を短く(約 8 分→15 秒)することができる。

5、M2システムにおける「3DCG 通信カラオケ」

上述した項目に基づき、“M2システム”上で「3DCG 通信カラオケ」のプロトタイプ開発を行った。以下にその構成と性能を示す。

音源エンジン:弊社のインターネットを用いた通信カラオケシステム「ゆめカラ」(<http://town.hi-ho.or.jp/town1/karaoke/index.htm> 上でサービス運用中)を、M2システムへ移植。CPU(PowerPC 602 内部クロック66MHz)で、MIDIデータ(独自フォーマット)から音源波形データを合成し音声波形を出力する。

- ・同時発音数:最大 24 音。
- ・出力サンプリング数:22.05 KHz

CGエンジン:ジオメトリ演算は CPU、レンダリング処理は 3Dグラフィックチップ“BDA”で各々行う。

- ・CG 性能(カラオケ再生時):180,000 (ポリゴン/秒) = 6,000(ポリゴン) × 30(フレーム/秒)

音源計算に CPUリソースの約 40%必要なので、M2システムの最高性能(約 50 万ポリゴン/秒)は出ない。

通信モジュール:PCMCIa/ifiにモデムカード(36.6Kbps)を組込む。PPP 接続し Ftp(TCP/IP)によりデータを転送。

- ・1 曲あたりの転送データ量:約 60Kbyte/曲(MIDIデータ約 50Kbyte、シナリオデータ約 10Kbyte)。
- ・転送時間:約 15 秒/曲。

6、おわりに

「3DCG 通信カラオケ」の開発について、以下のことを報告した。

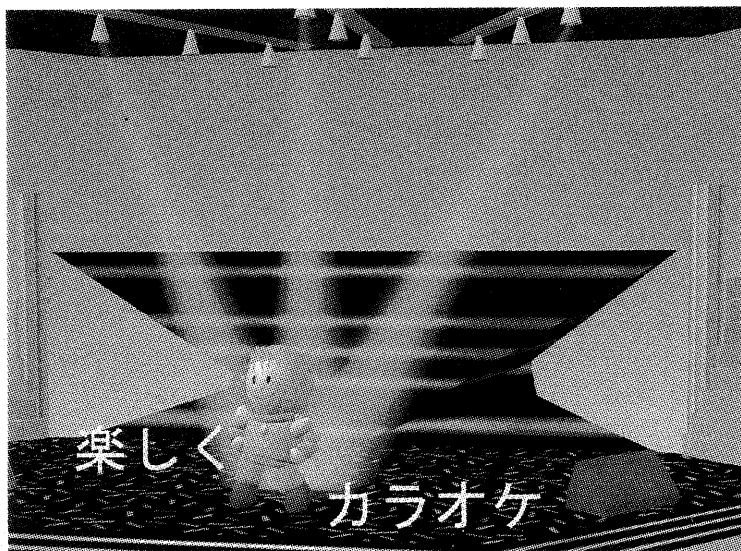
- (1)MIDIサウンドとCGアニメーションのAV同期方式を提案し、演奏テンポが動的に変化した場合の同期遅延を最大1拍(約500msec)程度に留めた。
- (2)モーションデータを部品化することにより3DCGカラオケコンテンツ制作工程を確立し、従来工程に比べ制作期間を短縮した。
- (3)モーションピース方式に基づく3DCGカラオケコンテンツデータ配信方式を確立し、電話回線(36.6Kbps)上でデータ転送時間を15秒程度に抑えた。

今後の課題を、以下に示す。

- (1)3DCG通信カラオケ以外のアプリケーションへの応用展開。
- (2)音源エンジンの3次元音場への拡張。
- (3)ホームポジションを持たないモーション間を滑らかに接続する方式を確立する。

参考文献

- [1]M.O'Rourke:Principles of Three-Dimensional Computer Animation: Modeling Rendering and Animating with 3D Computer Graphics、3次元コンピュータアニメーションの原理(邦訳)、トッパン(1997)
- [2]J.K.Hodgings and N.S.Pollard:Adapting Simulated Behaviors For New Characters、SIGGRAPH'97 pp153-162
- [3]則安:電子楽器とコンピュータ音楽の流れ、日本音響学会誌 48-08(1989)。
- [4]GM with GS HANDBOOK:キーボードマガジン
- [5]村田他:音楽を用いた対話的3次元CGアニメーション〜ミュージックドリフンCGの試作、情報処理学会ヒューマンインターフェース研究会(No.74-1)。



M2システムで実装した3DCGカラオケ