

## CGアニメーションのための ナレーション・BGM同期編集システムの開発

多田村 克己† 三好 孝治‡ 中前 栄八郎‡‡

†広島県立大学 経営学部 ‡広島工業大学 環境学部 ‡‡広島工業大学 工学部  
〒727 庄原市七塚町562 〒731-51 広島市佐伯区三宅2-1-1

本論文では、熟練や専門知識を持たない人でも、映像・音声情報の視察によるインタラクティブな操作を用いて、BGM・ナレーションの品質を損なうことなくCGアニメーションに容易に同期編集するためのプロトタイプを提案する。本提案システムは以下の特長をもつ。(1) 時間軸に沿ったBGM、ナレーションの各要素の内容が視覚的に把握でき、さらに、コマ取りCGアニメーションのシナリオも同一時間軸上に表示できる、(2) BGMおよびナレーション情報の短縮・伸長を品質を損なうことなく、インタラクティブに編集ができるインタフェースをもつ。

### Development of a Prototype System for Synchronizing Narration and BGM with CG Animation

Katsumi Tadamura†, Takaharu Miyoshi‡, and Eihachiro Nakamae‡‡

† School of Business, Hiroshima Prefectural Univ.

562 Nanatsuka Shoubara, 727

‡ Faculty of Environment, Hiroshima Institute of Technology

‡‡ Faculty of Engineering, Hiroshima Institute of Technology

2-1-1 Miyake Saeki-ku Hiroshima, 731-51

We propose a prototype authoring system to synchronize among CG animation, speech, and background music (BGM) by means of interactive and easy manipulation on a graphical interface; it is meant to be of assistance to those who lack special knowledge and/or skill in editing them.

Our proposed system provides the following two features: (1) Users can grasp the contents of a BGM and a narration along with the time line by visual information, and can also display a scenario of a CG animation sharing the same time line. (2) This system provides a graphical user interface for interactively editing (reduce and/or expand) both BGM and narration information without any distortion of their quality.

## 1 はじめに

プレゼンテーション,特にサイエンティフィックビジュアルライゼーションを目的とするCGアニメーションには、シナリオが付き物である。このシナリオは、映像(CGアニメーション、ビデオ、etc.)とオーディオ(BGM、肉声、擬音、etc.)それぞれのメディア要素から構成されるのが普通である。これらの各メディアの出力の調和は、プレゼンテーションの良否に密接な関係を持つ。中でも、これらのメディアすべての時間的なマッチング、すなわち各シーンに応じたメディア要素の開始、終了時刻の同期がプレゼンテーション全体の品質に大きく影響する。このため、プレゼンテーションのシナリオのための各要素およびそれらのお互いの関係を容易に記述でき、編集を助ける情報を提供できるツールの開発が望まれる。特に最近、CG作成、ビデオ編集、およびデジタルレコーディングがデスクトップ環境でできるようになってきた。その結果、例えば、internet上でのWWWのホームページ作成に代表されるマルチメディア情報を、これらのメディアについて熟練していない人が作成する機会が多くなってきた。このような人でも、マルチメディアドキュメントのデータベースを構築できるようにすることが望ましい。そのための使いやすいオーサリングツールが必要である。

これまでに、映像とオーディオ情報の編集を同時に行うソフトウェアが多数市販されている[1,2]。また、分散ネットワーク環境におけるマルチメディア情報の時間的な調整およびシナリオの編集に関して、数多くの研究が行われている[3-9]。しかし、従来の手法では、アニメーション、BGM、ナレーションについて熟練していない人がこれらのメディアを品質を損なうことなく編集することは非常に困難であった。他方、話し言葉のみを対象として、その原音質を保った音声の速度変更用のツールが開発されている[10]。これは年齢などにより異なる個人の快適な聴取のための条件に対応した音声情報を提供する。しかし、当然のことながら他のメディア情報との同期を考慮していない。

本論文では、熟練や専門知識を持たない人でも、

視覚・音声情報の視察によるインタラクティブな操作を用いて、BGM(Back Ground Music)とナレーションをその品質を損なうことなくCGアニメーションに同期編集するためのプロトタイプを開発した。

## 2 提案手法の特徴

本論文では、複数のシーンからなるCGアニメーションの各シーンを基準にして、これにBGMとナレーションを同期させて付加するための、次の特徴を持つインタラクティブなオーサリングツールを提案する。

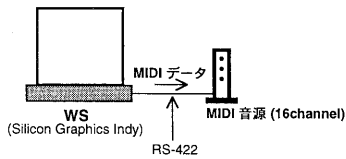
- (1) 時間軸に沿ったBGM、ナレーションの各要素の内容が視覚的に把握でき、さらに、コマ取りCGアニメーションのシナリオも同一時間軸上に表示できる。
- (2) コマ撮りアニメーションのシナリオをインタラクティブに編集できるグラフィカルユーザインタフェースをもつ。
- (3) BGMおよびナレーション情報の圧縮、伸長を品質を損なうことなく、インタラクティブに編集ができるインタフェースをもつ。

なお、音声(human voice)情報処理の分野では、音声情報を有声音区間と非有声音区間に分類し、後者をさらに無声音、無音区間に分類している[11]。しかし、本論文では、音声情報を、ユーザの定義による無音とみなす疑似無音区間(無声音、無音区間、および有声音区間の一部)とそれ以外の有声音区間の2種類に分類する。

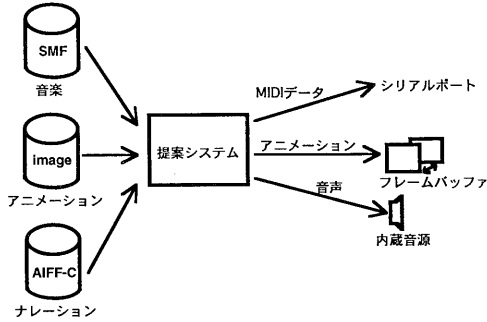
## 3 ユーザインタフェース

### 3.1 データi/o

提案システムの、プロトタイプハードウェア構成を図1(a)に示す。提案システム利用の前提として、アニメーション用データとしてCG画像フレームデータ(rgb,非圧縮)、BGMデータとしてSMF(StandardMIDI File)形式のデータ、ナレーションデータとしてAIFF-C形式のデータ(ステレオ、16bit/サンプル/チャンネル)がすでに作成されているものとする。システムは、これらのデータを利用して編集作業を行う。プレイバック時の出力と



(a) ハードウェア構成



(b) データI/Oの概略

図1. システム構成

して、ナレーションに対してはワークステーションの内蔵音源で出力し、BGMデータに対してはシリアルポートに接続したMIDI音源にMIDIデータを送信する。アニメーションはCG画像の解像度を落としたプレビュー用の画像(例えば、320X240ピクセル)を作成し出力する(図1(b)参照)。

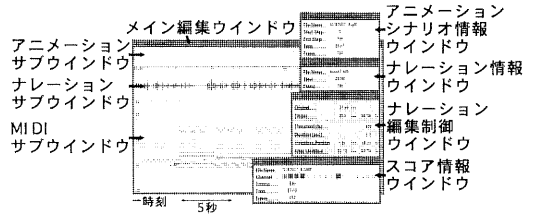
これら、ナレーション、BGM、アニメーションのデータはすべて主記憶上に蓄えておき、リアルタイムで出力可能である。ただし、大きいメモリ容量を必要とするアニメーションについては、プレイバック実行時の表示解像度とプレイバック総フレーム数(プレイバック時間に、1秒間当たりの表示フレーム数を掛けたもの)とがトレード・オフの関係になる。

### 3.2 作業ウィンドウとその役割

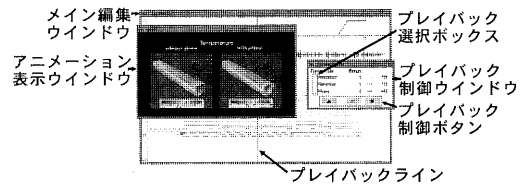
図2に、提案システムの編集およびプレイバック用ウィンドウの例を示す。

図2(a)は、編集モードにおいて、CGアニメーション、BGM、ナレーションの同期編集を行うためのウィンドウを表示したものである。これらは、以下の機能を持つウィンドウから構成される。

(a) メイン編集ウィンドウ: 時間軸とともにCGアニメーション、BGM、ナレーション情報を同時に表示するために、以下の3つのサブウィンドウか



(a) 編集モードで使用するウィンドウ



(b) プレイバックモードで使用するウィンドウ

図2. グラフィカルユーザインタフェース

ら構成される; 同一時間軸上に、アニメーションのシーンごとのチャートを表示するためのアニメーションサブウィンドウ、BGM用のMIDI情報表示のためのMIDIサブウィンドウ、およびデジタルサンプリングされたナレーション情報表示のためのナレーションサブウィンドウ。これらのサブウィンドウは、それぞれ同時にしくは選択的に表示することができる。

(b) アニメーションシナリオ情報ウィンドウ: アニメーションに関連する情報を表示する。すなわち、編集ウィンドウに表示されているアニメーションシナリオに対応するファイル名と再生時間を表示する。

(c) スコア情報ウィンドウ: MIDIフォーマットで記述されたBGMの情報を表示する。すなわち、編集ウィンドウに表示されているMIDI形式のBGMに対応するファイル名、使用チャンネル番号、テンポ、および再生時間を表示する。

(d) ナレーション情報ウィンドウ: ナレーションデータに対応するファイル名と再生時間を表示する。

(e) ナレーション編集制御ウィンドウ: ナレーション編集時のパラメータを入力する。

図2(b)は、プレイバックモードにおいてのみ使用する2つのウィンドウを示したものである。これらは、それぞれ、以下の機能を持つ。

(f) プレイバック制御ウィンドウ: プレイバックする要素とその区間の選択、および再生・停止の制御を行う。

(g) アニメーション表示ウィンドウ: プレイバック要素にアニメーションが選択された際、シナリオにしたがってアニメーションを表示する。

#### 4 アニメーションシナリオ編集

提案システムは、アニメーションの複数のシーンをBGM、ナレーションと同期して合成したり、fade in/outで切り替える際の適切な時間設定などの試行錯誤をインタラクティブに行えるアニメーション用シナリオの編集機能を提供する。例えば、“光の遊び”的なアニメーションを作成するような場合、複数光源を同時に考慮した画像によりアニメーションを作成すると、1フレーム当たりの計算時間が大きくなり、単独の光源について光の強弱、色相の変更を行うことが極めて困難である。それに対して、個々の光源について独立に画像を作成し、合成が容易な編集機能があれば低コストで高品質なアニメーションを容易に作成することができる。また、これらの光源の点灯・消灯のタイミングや色の変化にBGMの特定の楽器を同期させることにより、よりアトラクティブなものが作成可能になる。これを実現するために提案システムは、各シーン(光源)の合成比率の変化を折れ線グラフのタイムチャートで表現し、さらにスペクトル分布を変更できる機能を備えている。これにより、シーンチェンジの際の、fade in/outのタイミングや長さを調節できる。さらに、“光の遊び”アニメーションを作成する際には、各光源毎に単位強度で白色の光源を配した静止画(すべて静止している場合)、一周期分の画像(視点もしくはシーン中の物体が周期性のある動きをする場合)、もしくはシーンに対応するコマ数の画像を、予め生成しておけば、光源の点灯・消灯、明るさ、および色を自由に組み合わせたアニメーションが作成できる。

具体的には、アニメーションシナリオ編集のための以下の機能を備えている。

(1) シーン(光源)毎のタイムチャート作成のためのグラフ描画機能

図3に示すメニューから、線種、線幅、線色を選択

し、シーン毎にタイムチャートを描画できる。また、グラフの節点の編集(追加・削除・移動)も可能である。

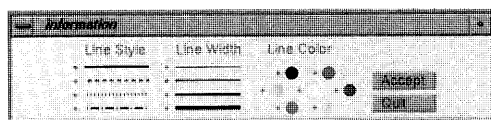


図3. アニメーションタイムチャート描画メニュー (2)複数シーン(光源)のタイムチャートの編集(結合・削除)機能

複数のタイムチャートを一つのタイムチャートにマージしたり、複数のシーン(光源)を含むタイムチャートから特定のシーン(光源)のタイムチャートのみを削除することができる。

なお、作成したシナリオは、任意の時点でプレイバックモードに移り再生可能である。ただし、指定された光源のシーケンスを初めて再生する際には、プレビュー用の縮小サイズの画像を作成し、ディスクのワーク領域に保存する必要がある。そのため、若干の待ち時間を生じる(時間は、画像サイズとフレーム数に依存)。

#### 5 BGMとアニメーションの同期

本論文では、BGMとアニメーションの同期を要求される状況として、以下の2つを想定している。

- (1) 曲全体を一様に短縮・伸長してアニメーションの表示時間にあわせる。
- (2) 一曲の演奏途中の特定の区間を、アニメーションの特定のシーンにあわせ、かつ、曲全体をアニメーションの表示時間にあわせる。

ここで、SMF形式で記述されたBGMの再生時間を変更することを考える。通常、BGMを再生する際、演奏時間はテンポの変更により調節する。SMFで記述された音楽では、この効果は曲のテンポに関連する、4分音符1つ当たりの時間(TQN)、4分音符1つ当たりのデルタタイム数(RQN)値の変更、もしくは各イベントのdelta time値の変更による実現方法が考えられる。ところが、RQNは、曲全体を定義するヘッダ部分のデータであり、曲の途中で変更することができない。したがって、上述の

(2)が実現できない。また、delta time値の変更による方法においては、整数化の際の丸め誤差が累積すると、無視できない大きさになる。さらに、この方法では、変更履歴を残したとしても元のデータに回復できない。したがって、提案手法では、BGM情報の再生時間を調節するため、テンポデータ(TQN)を変更する方法を採用した。

ここで、上述の(2)を実現するためには、BGMの再生時間は1フレームの表示時間、すなわち1/30秒の精度で調節できればよい。通常の楽曲は、100 quarter notes/minute程度であり、このときのTQNは $6 \times 10^5$  (micro sec)である。この場合、TQNの1レベルの変更が1/30secを越えるのは、33分以上の曲に対してである。提案手法では、高々数分のシーンに対して同期を取ることを考えているから、十分な精度を保証できる。

以下に、(1),(2)に対する具体的な実現方法を説明する。

(1)の場合は、BGMの演奏時間を $T_m$ 、アニメーションの時間を $T_a$ 、変更前のテンポを $tmpo$ 、変更後のテンポを $tmpo'$ とすると、 $tmpo'$ は $tmpo' = \text{int}(T_a/T_m * tmpo)$ から容易に求めることができる。

(2)の場合は、以下のようにして実現する。まず、同期を取りたいMIDIのイベント(例えばドラムの発音)とアニメーションのフレームを編集ウィンドウ上でユーザが指定する。図4(a), (b)は、ドラムの発音イベント $B_j$ とアニメーションのシーンが変わるFade in/ outのフレーム $A_i$ を同期させる操作を示したものである。以下に処理手順を説明する。

(a) プレイバックモードで、アニメーションおよびBGMをそれぞれ再生する。再生結果を参考にしながらそれぞれの要素について、編集ウィンドウ上で同期させたいおおよその位置にマウスでマークを付けておく。このとき、ユーザは、同期させたいアニメーションとBGMの位置の時間的な差を視察により容易に認識できる。

(b) 操作メニューから"Music adjust"を選択する。

(c) (a)の情報を基にして、アニメーションサブウィンドウ中で、同期を取りたいアニメーションのフレーム番号を指定(マウスでクリック)する(図4の

例では、 $A_i$ )。

(d)  $A_i$ に同期させたいMIDIのイベント(この例ではドラムの発音)を、MIDIサブウィンドウ中で指定(マウスでクリック)する。(図4の例では、 $B_j$ )。

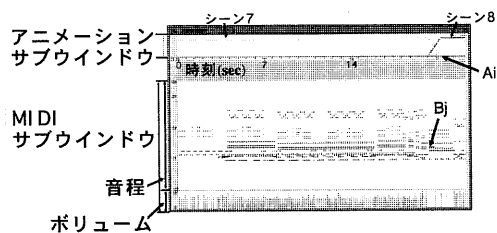
(e) システムは、 $A_i, B_j$ の情報から、以下の(1), (2)式の計算により、新しいテンポをそれぞれの区間に対して求め、その結果に基づいてMIDIサブウィンドウを再描画する。すなわち、イベント $B_j$ を境にして、BGMが2つのセグメントに分割されたと考え、別々のテンポを与える。すなわち、1から $j$ までのイベントに対しては、以下のテンポを与える。

$$tmpo' = \text{int}((A_i/30) / M_j * tmpo) \quad (1)$$

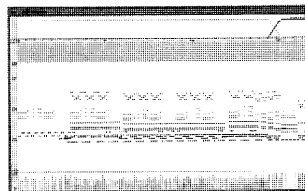
ここで、 $M_j$ はイベント $B_j$ の絶対時間、 $tmpo$ は変更前のテンポである。同様に $j+1$ から最終イベント( $n$ )までのセグメントに対しては、以下のテンポを与える。

$$tmpo'' = \text{int}((M_n - A_i/30)/(M_n - M_j) * tmpo) \quad (2)$$

以上の処理の結果、 $B_j$ のイベントが $A_i$ に移動する(図4(b)参照)。



(a) 同期調整処理前



(b) 同期調整後

図4. BGMとアニメーションとの同期調整

## 6 アニメーションとナレーションとの同期

### 6.1 従来の問題点と提案手法の特長

ナレーションを短縮・伸長するため、単純にデジタルサンプリングされた情報を $a$ 倍( $a$ は正の実数)すると、人声(human voice)の存在する部分(有声部)の周波数が変化するため、著しく音質が劣化する。そこで、音声のない無音の区間を抽出し、

その部分のみを短縮もしくは伸長すれば、音質を損なうことなく所期の目的を実現できる。ここで、ナレーションとアニメーションの同期が必要となるのは、大抵の場合、ナレーションが長すぎる場合である。従来のマルチメディア編集ソフトでは、ナレーションデータの短縮を行う際、デジタル化された波形データを読み、無音区間を探し出し、それをカットするしかない。この場合、明らかにそれとわかる無音区間が十分に存在しないときの編集作業は容易ではない。

他方、品質を損なわずに、無音区間のうちのどの部分をどれくらいカットできるかを判断できるようになるためには、かなりの経験を要する。また、デジタル音声情報処理の手法を利用して、2節で述べたように音声情報を有声音、無声音、無音の3つの区間に分類し、この中の無音区間のみを処理対象にすると、音質を保持した編集処理が行える。しかし、適用するデータ毎の違いを調節するために、上述の3つの区間に分類するためのパラメータ[11](有声音検出パラメータ、零次メルケプストラム係数、零交差数)をユーザが指定する必要がある。これらのパラメータを適切に設定するのは複雑で、かつ、効率よくこれを修正するためには、専門的な知識を要する。

提案システムはこれらの手法とは異なり、音声処理に対して専門的な知識を持たないユーザでも、判りやすい簡単な操作でナレーション情報の短縮(伸長)ができるように、以下の機能を具備している。

- (a) グラフィカルユーザインタフェースを利用した、疑似無音区間の抽出のためのパラメータの設定、および抽出された区間情報のテキストおよびグラフィックスによる表示。
- (b) 上記設定パラメータ下で編集可能な疑似無音区間の自動抽出、およびその区間の自動短縮(伸張)による、有音部を含むナレーション情報の短縮(伸長)。
- (c) 編集途中の情報のプレビューのためのプレイバック。

提案手法は、ユーザが、実際に編集中のナレーションを聞きながら、インタラクティブに疑似無

音区間と認識するパラメータを聞き取りやすいものへと修正することによって、最適な結果を得るものである。提案手法では、以下の理由により疑似無音区間抽出パラメータを2個用いる。まず、有声音区間では、短時間(例えば10msec)内に振幅がある閾値以上になる部分が必ず存在することに着目し、振幅の大きさをパラメータの一つに採用する。ところが、もし振幅の大きさのみで疑似無音区間を定義すると、音声の波形が0と交差する位置付近の区間がすべて対象になる。これらのほとんどは、実際には有声音部分であるので、振幅の大きさのみで疑似無音区間を抽出することはできない。そこで、もう一つのパラメータとして、閾値以下の振幅がある期間以上継続する時間(以下、微声継続時間と呼ぶ)を採用する。この2つのパラメータを用いることにより、実際の有声音区間が編集対象とならないようにすることができる。

## 6.2 ユーザインタフェースと概略処理手順

図5は、ナレーション編集時に利用するユーザインタフェースを示したものである。ナレーション編集中は、編集ウィンドウには、アニメーションのシナリオとナレーションの波形情報とを容易に把握できるように、アニメーションサブウィンドウとナレーションサブウィンドウを表示する。また、編集対象のアニメーション、ナレーション情報のファイル名、再生所要時間が判るように、それらをテキスト情報としてアニメーションシナリオ情報ウィンドウおよびナレーション情報ウィンドウにそれぞれ表示する。また、後述するナレーションデータを編集するための疑似無音区間抽出のためのパラメータ(音声振幅と微声継続時間それぞれの閾値)が入力されると、短縮(伸長)により変更された再生時間および短縮(伸長)率の数値情報がナレーション編集制御ウィンドウに表示される。ユーザが、ナレーションサブウィンドウに左右両方のチャンネルの波形データを表示したい場合は、それらは別々の色(図5の例では左(CH1)が緑、右(CH2)が青)で表示される。

ナレーションサブウィンドウでは、有音区間と編集可能な疑似無音区間とが別々の色で表示(図4中では、編集可能な疑似無音区間をオレンジ色で

表示)される。また、閾値の範囲が容易に把握できるように正、負両方にバーが表示(図4中では赤色で表示)される。



図5. ナレーション編集時に使用するグラフィカルユーザインタフェース

アニメーションとナレーションの同期をとるための概略処理手順を以下に示す。

- (1) アニメーションシナリオおよび同期させたいナレーション情報のファイルをユーザが指定する。
- (2) これらの情報は、それぞれ編集ウインドウ中のアニメーションサブウィンドウおよびナレーションサブウィンドウに表示される。アニメーションシナリオ情報ウインドウ、およびナレーション情報ウインドウが自動的に開き、それぞれに対応したファイル名、素材の長さに関する情報が表示される。
- (3) ユーザは、読み出したアニメーションとナレーション情報をそれぞれプレイバックし、再生した内容を参考にして、編集ウインドウ上においてアニメーションとナレーションそれぞれの同期をとりたい位置にマウスで印を付ける。
- (4) ユーザが、「ナレーション編集処理」をメニューから選択すると、ナレーション編集制御ウインドウが開かれ、ナレーションデータの録音時間が「Original」に表示される(図5参照)。
- (5) 処理対象としているナレーションデータを短縮(伸長)して編集したい時間を入力するようにメッセージが表示される。ユーザがその時間を入力すると、それが「Target」(図5参照)に表示され、Originalに対するTargetの比率、デフォルトパラメータ下でのオリジナルのナレーション情報に対する疑似無音区間の長さ、およびそれを利用した修正後のナレーション情報の長さが、ナレーション編

集制御ウインドウ上に表示される。さらに、編集対象となる疑似無音区間を色分けした波形データがナレーションサブウィンドウに再描画される。具体的には、デフォルトパラメータとして、閾値は編集対象の音声データの振幅の絶対値の平均値を、また、微声継続時間は1秒以上が初期設定される。ただし、これら2つのパラメータの組み合わせで抽出した疑似無音区間は、無音区間以外に無声音区間と有声音区間の一部を含んでいる可能性がある。より良好な音質を保持するためには、これらの部分をできるだけ残す必要がある。

(6) ユーザは、デフォルトパラメータ下での短縮(伸長)率を参考にして、疑似無音区間抽出のためのパラメータを変更することができる。この場合は、ナレーション編集制御ウインドウにおいて数値を入力する。

(7) 変更されたパラメータ下の疑似無音区間が抽出され、それに基づきナレーションサブウィンドウが再描画される。同時に、編集結果の時間と、オリジナルのナレーションデータに対する比率が計算され、ナレーション編集制御ウインドウの情報が修正される。

(8) ユーザは必要に応じてプレイバックモードに移って、編集途中のナレーションデータを実際に聞いて適正に編集されているかどうか確認する。

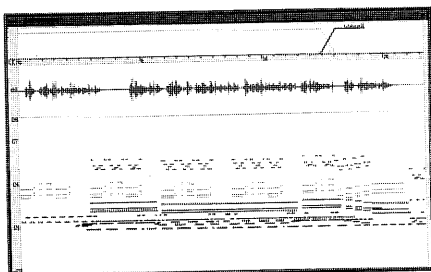
(9) (6)-(8)の処理を、ユーザが満足する同期編集データが得られるまで繰り返す。

このようにして、ユーザは、疑似無音区間抽出のためのパラメータのみをインタラクティブに修正することにより、目的とするアニメーションと同期のとれたナレーションデータを作成できる。

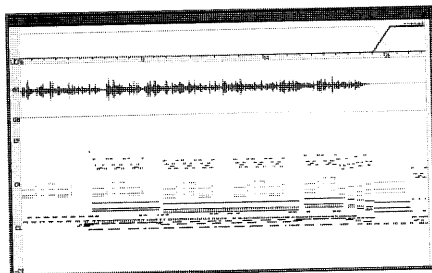
## 7 実験結果

音楽(BGM)、ナレーションおよびCGのマッチングの実験結果について述べる。709フレーム(23.8秒)のハイビジョンCGアニメーション、27.62秒のBGM、および25.89秒のナレーションを独立に作成し(図6参照)、提案システムのプロトタイプ上で、CGアニメーションにBGMとナレーションを同期させ、結果を確認したところ良好な編集結果を得

た。具体的には、サイズが384\*207(オリジナル1920\*1035)のCG画像データ、5チャンネル使用のMIDIデータ、8KHzのナレーションデータを使用して、5フレーム/秒のフレームレートで編集結果のプレイバックを実現できた。CGアニメーション、ナレーション、BGMのデータ量は、それぞれ、37.5Mbyte,828kbyte,4.2kbyteである。ワークステーションは、Silicon Graphics社のIRIS Indigo (R4400 150MHz, 主記憶96MB)を使用した。



(a) 同期調整処理前



(b) 同期調整処理後

図6. 適用例

## 7 結論

本論文では、特別な熟練や専門知識を持たない人でも、BGMとナレーションの品質を損なうことなく、CGアニメーションに同期編集するためのプロトタイプを開発した。

今後、以下の課題を解決することが望まれる。本論文では、(1)アニメーション用のデータをCGによるRGBフレームデータに限定した。適用範囲を広げるためには、movieファイルへの応用が必

要である。(2)グラフィックスワークステーション上で特別なハードウェアを利用することなくプロトタイプを構築したので、メモリの制約により長いシーンの場合、プレイバック時にアニメーションの品質が落ちる問題点がある。(3)本システムと、他のジョブが競合すると、プレイバック時にそれぞれのメディアの同期がずれることがある。(4)アニメーションにナレーションを合わせる場合、提案手法により抽出した短縮可能な疑似無音区間をすべて削除しても、目標時間内にナレーションが入りきらない場合は、ナレーションを作りなおす必要がある。

## 参考文献

- [1] Adobe Premiere 日本語版2.0 ユーザガイド.
- [2] Avid Media Suite Pro User's Guide.
- [3] Ackerman: Direct Manipulation of Temporal Structures in a Multimedia Application Framework, ACM Proc. Multimedia 94, pp.51-58 (1994).
- [4] Anderson and Homsy: A Continuous Media I/O Server and Its Synchronization Mechanism, IEEE Computer 24, 10, pp.51-57 (1991).
- [5] Hardman, Rossum, and Bulterman : Structured Multimedia Authoring, ACM Proc. Multimedia 93, pp.283-289 (1993).
- [6] Drapeau: Synchronization in the MAestro Multimedia Authoring Environment, ACM Proc. Multimedia 93, pp.331-339 (1993).
- [7] Hirzalla, Falchuk, and Karmouch: A Temporal Model for Interactive Multimedia Scenarios, IEEE MultiMedia 2, 3, pp. 24-31 (1995).
- [8] Karmouch and Emery: A Playback Schedule Model for Multimedia Documents, IEEE MultiMedia 3, 1, pp. 50-61 (1996).
- [9] Hamakawa and Rekimoto: Object Composing and Playback Models for Handling Multimedia Data, ACM Proc. Multimedia 93, pp.273-281 (1993).
- [10] Nakamura, et al.: Real Time Speech Rate Converting System for Elderly People, Proc. ICASSP-94, pp.1-4 (1994).
- [11] 今井 聖: 音声認識, 共立出版 (1995).