

2C-4

デジタル映像制作編集システムの 音声編集機能

斉藤一実、川口尚久、飯島泰裕
富士通研究所

1. はじめに

近年、ビデオカメラやVTRの普及によりビデオによる報告やプレゼンテーションが増えてきた。ビデオはアイデアやイメージなどを直観的に理解でき、受け手の記憶にも残り易い。しかし、内容をだらだらと読みあげるようなビデオでは受け手を退屈させる。効果的に音楽や効果音などの音声を加えることで記憶に残る表現が可能となる。

我々は、効果的なビデオ作品を容易に作成できる、パーソナル用のデジタル映像制作編集システムDPPS (Digital Post-Production System) を開発した。

本報告では、本システムにおいて音声を映像と同期させるための機構と、音声編集ソフトウェアについて述べる。DPPSのシステムについては[1][2]を参照されたい。

2. 従来の音声、映像編集

インパクトのあるビデオを制作するには、効果的な音声が必須である。すなわち、シーンのつなぎ目でBGMを切り換えたり、画面の動きに合わせて効果音を入れる。しかし映像を見ながらCDをかけたり、効果音を出すことは難しい。プロの制作現場では、映像と音声のフレーム単位のタイムチャートを作った上で、映像とは別にさまざまな音楽、効果音等をミキシングした音声テープを作成し、最後に映像と合わせるといった作業を行っている。このように音声編集は、編集した結果をすぐ見て修正することが困難で、熟練した作業でなければ難しい。

すなわち、ビデオ作品の音声編集においては、通常の音量調整、音の混合といったサウンドミキシング機能に加え、音声を視覚的に処理する機能が必要となる

3. AVPUの音声処理機構

我々は、ビデオ作品の制作編集を容易化するために、映像と音声を全てデジタル処理できるデジタル映像制作編集システム DPPSを開発している。DPPSは、AVパソコンFM-TOWNSにAV処理装置AVPU (Audio Video Processing Unit) を付加し、その上で動作する制作編集ソフトウェアから構成される。

AVPUは、映像、音声を全てデジタル化して記録、再生、処理する装置である。AVPUの基本性能は表1のとおりである。

音声品質	16ビットリニア 2チャンネル 48kHzサンプリング
画像品質	360×240ドット 32767色
記録時間	最大約7分 (但し音声は8チャンネル分)

表1. AVPUの基本性能

音声サンプリングレートを48kHzとしたのは、映像のリフレッシュレート 30フレーム/秒の倍数として同期を容易にするためである。これはDATと同じ仕様であり高品質な音声といえる。

図1はAVPUの構成図である。AVパソコンFM-TOWNSをホストとし以下の3種のボードからなる。

- ①コントロールボード—装置全体の同期、制御
- ②画像処理ボード—画像データの記録、再生、処理
- ③音声処理ボード—音声データの記録、再生、処理

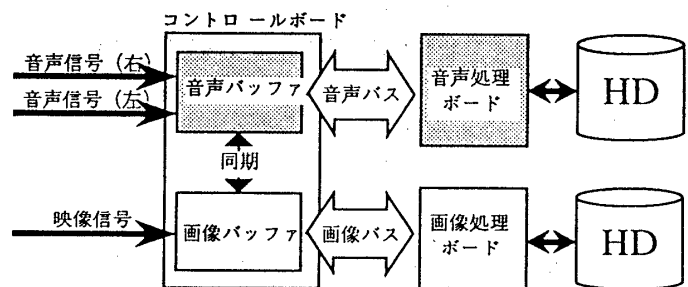


図1 AVPUの構成図

デジタル化した映像、音声信号は別々のハードディスクに記録する。映像、音声ともそれぞれのハードディスク上のデータにランダムアクセスして再生することで、任意の音声と映像を1フレーム単位に確実に同期して再生することが可能となっている。

さらにAVPUでは映像を安定に再生するために画像バッファを設けてTBC機能を持たせた上に、音声にも図2のように同様のバッファを用意し、双方のバッファを同期させた。これにより音声、映像間の同期が確実にとれるようになった。ズレは最大でも音声の1サンプリングタイム分(約21μS)となり可聴範囲で無視できる。

AVPUでの録音は図2の様にして行う。A/D変換された左右2チャンネル分の音声データは音声バスを通して音声処理ボードへ送る。音声データはチャンネルごとに一旦ウェブメモリにDMA転送して書き込む。ウェブメモリをDMAコントローラによって2チャンネル分のリングバッファとして使うことで、各チャンネルの音声データを1台のハードディスク上でそれぞれ別々のセクタに記録できる。

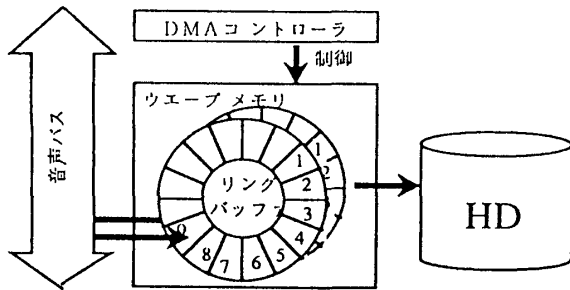


図2 音声の記録方法

図3 音声編集ソフトウェア

音声データは1チャンネル当たり96kB/秒の転送速度で送られてくるが、これに対しハードディスクは約10倍の1MB/秒の転送速度を持っているので、最大10チャンネル程度の同時録音、再生も可能である。この様にしてハードディスクの中を仮想的な音声チャンネルとして区切って使うことでマルチトラックシステムとして使用できる。

4. 音声編集ソフトウェア

図3は本システムの音声編集ソフトウェアの画面である。画面上部には、取り込んだ映像を縮小して1フレーム毎に表示している。その下には、そのフレームが映される時に出す音声を波形で表示する。音声は1つの映像シーンに対し、8チャンネルまで独立に操作できる。

映像と音声のタイミング合わせは、スタイラスペン(本システムではマウスにかわりポインティングデバイスとしてスタイラスペンを使用する)を用いて画面上で音声波形を移動させて指示する。音声のタイミング変更は映像の1フレーム単位に変更できるので正確なアジャストが可能である。

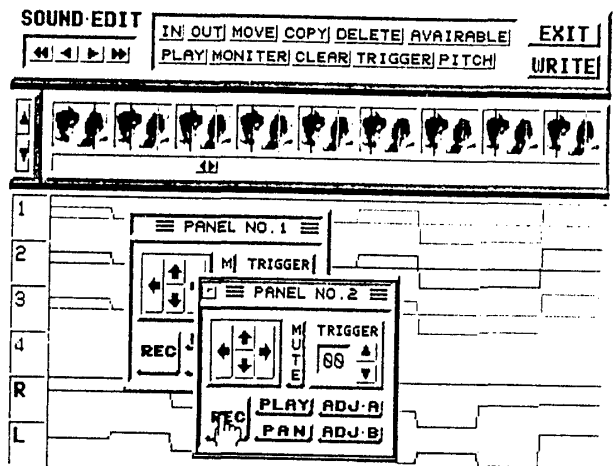


図3 音声編集ソフトウェア

また、画面上の波形の任意の範囲を指定して、削除、移動することもできる。

これらの編集はチャンネル毎に指定可能なので効果音、BGM等をそれぞれ独立に任意の位置にずらせる。編集の結果はテストプレイ機能で直ちに映像を見ながら聞けるため、短いターンアラウンドで何度でも修正が可能である。

ミキシング機能についてもチャンネル毎に音量を増減したり、フェードイン、アウトを指示できる。

また、音声がステレオであることを生かして映像中の左右どの辺りから音声が発しているかを、映像を見ながらその位置をポインティングすることで簡単に指示できる。

5. まとめ

正確な音声、映像同期機構を持つこのDPPSに、映像をフレーム単位で見ながら自由に音声を編集できるソフトウェアを用意することで、簡単に映像と音声の位置合わせや、音声のミキシングができ、完全に同期がとれたビデオが制作できるようになった。今後は、プロのビデオアーティストやビデオ制作者に試用、評価してもらい、改良していく予定である。

[謝辞]

研究の機会を与えてくださった基盤システム研究部門森田部門長代理、マルチメディア研究部藤田部長、山本部長代理に感謝いたします。

[参考文献]

- [1] 飯島、川口、斉藤 " パーソナルなビデオ制作に向けて"、信学画像工学研究会 IE91-7、1991
- [2] 斉藤、川口、飯島 " デジタル映像制作編集システムの並列画像処理機構"、情報処理学会43回全国大会、1991